

# **Mediciones acústicas basadas en software**

*Federico Miyara*



# **Mediciones acústicas basadas en software**

*Federico Miyara*

Miyara, Federico  
Mediciones acústicas basadas en software - 1a ed. - Gonnet:  
Asociación de Acústicos Argentinos, 2013  
350 p. ; il.; 26x18 cm

ISBN 978-987-24713-2-3

1.Mediciones. 2. Acústica. 3. Software.  
CCD 534

Fecha de catalogación: 25/07/2013

ISBN 978-987-24713-2-3

© Federico Miyara, 2013  
Hecho el depósito que marca la Ley 11.723

Diseño de tapa: Federico Miyara  
Foto de contratapa: María Victoria Gómez

IMPRESO EN LA ARGENTINA / PRINTED IN ARGENTINA  
Asociación de Acústicos Argentinos, AdAA

*A María Victoria, Francisco y Andrés*

*Un agradecimiento muy especial a mi esposa, María Victoria Gómez, y a mis hijos, Francisco y Andrés, por el permanente apoyo a mis proyectos locos. A mis padres, José Miyara y Julia Elvira Verdeja les agradezco la mejor herencia que me podrían haber dejado, el estímulo intelectual y el haberme enseñado a apreciar el valor del conocimiento, condiciones imprescindibles para que algún día pudiera escribir este libro. Vaya también mi agradecimiento a mis compañeros, becarios, adscriptos y amigos del Grupo Ruido de la Facultad de Arquitectura de la UNR y del Laboratorio de Acústica y Electroacústica de la Facultad de Ingeniería de la UNR, Susana Cabanellas, Vivian Pasch, Marta Yanitelli, Patricia Mosconi, Juan Carlos Rall, Jorge Vazquez, Joan Puigdomènech, Pablo Miechi, David Giardini, Fernando Marengo Rodríguez, Ernesto Accolti, Ezequiel Mignini, Federico Bergero, Mauro Treffiló, Ezequiel Raynaudo y Luciano Boggino, con quienes hemos compartido muchas aventuras metrológicas y de elaboración de software que fueron dando origen y permitiendo depurar los métodos aquí expuestos. Corresponde también un reconocimiento a la Agencia Nacional de Promoción Científica y Tecnológica, organismo que financió proyectos de investigación en el marco de los cuales se desarrollaron y pusieron en práctica varias de las técnicas y protocolos en las que se inspiran partes del libro. Por último, mi agradecimiento a la Asociación de Acústicos Argentinos, la entrañable AdAA, en particular a quien en esta etapa la preside, Nilda Vechiatti, por brindarme la posibilidad de publicar este texto bajo su sello editorial y, así, prestigiarlo.*

## Prólogo

La metrología acústica ha presentado desde siempre el problema del elevado costo de cualquier instrumento de medición capaz de ofrecer mínimas garantías en cuanto a la exactitud de los valores medidos. Dicho costo corresponde esencialmente a dos tipos de componentes: el transductor o sensor, es decir, el micrófono, y la electrónica requerida para acondicionar la señal eléctrica proveniente del transductor, someterla a diferentes procesos (filtrado, análisis, comparaciones) y presentar resultados visibles en un indicador.

En sus orígenes, tanto el micrófono como la electrónica eran intrínsecamente costosos. El micrófono, por ser un elemento electromecánico sumamente delicado, que requiere un complejo proceso de fabricación e insume materiales de gran calidad para garantizar robustez mecánica, estabilidad en el tiempo de sus características y baja distorsión. La electrónica, por requerir una cantidad de componentes costosos con tolerancias estrictas y un aporte de mano de obra altamente especializada tanto en el armado como en la selección y ajuste de los componentes.

Actualmente, si bien la fabricación de micrófonos de instrumentación sigue siendo onerosa, el costo de la electrónica se ha reducido en varios órdenes de magnitud, por lo que los precios de los instrumentos de medición acústica deberían haber bajado considerablemente. ¡Sin embargo, tienden a aumentar! Por ejemplo, hace una década era posible adquirir un sonómetro de precisión de primera marca por alrededor de 2000 dólares. Hoy en día el instrumento más económico que lo reemplaza cuesta el doble, cuando la inflación del dólar a nivel internacional en el mismo periodo alcanzó tan sólo un 30 %, y cuando el costo de fabricación se ha reducido aún más por la tercerización en países como China y por el reemplazo de muchos componentes discretos por circuitos integrados dedicados. Como contrapartida, la industria informática ha mostrado una tendencia contraria: los precios de las computadoras de lo que en cada época se considera de última generación no sólo no han aumentado sino que incluso han bajado con respecto a la inflación del dólar.

Los fabricantes justifican tácitamente el aumento de precios ofreciendo nuevas y atractivas prestaciones que hoy son posibles por la incorporación de sofisticados dispositivos de microcómputo. Sin embargo, el costo real del hardware y del software asociado es una fracción muy pequeña del precio de venta al usuario de los instrumentos. De hecho en muchos casos han adoptado una política heredada de la industria del software consistente en la venta de licencias para módulos de software sobre una misma plataforma de hardware. Desde el punto de vista del fabricante, esto reduce aún más los costos, ya que con un único hardware ofrecen la funcionalidad de diversos instrumentos e incluso la posibilidad de actualizaciones futuras (que por cierto no serán gratuitas). En algunos casos, incluso, los módulos de software ya están preinstalados y sólo deben habilitarse mediante un código (previo pago de la correspondiente licencia).

Ante esta situación surge la idea de explotar las posibilidades, hoy accesibles a cualquier usuario con una mínima experiencia en programación, que ofrece cualquier computadora personal, y sustituir o emular los procesos que en los sonómetros se realizan por software. El prerequisite para ello es disponer de la señal en formato digital, lo cual se logra grabando, generalmente mediante un grabador digital externo, la señal de audio calibrada que la mayoría de los sonómetros comerciales proporcionan como salida de audio analógico. Una vez digitalizada, la señal puede ser sometida a innumerables

procesos, desde la obtención de cualquiera de los parámetros e indicadores clásicos hasta todo lo que la imaginación nos ofrezca como alternativas dignas de ser exploradas e investigadas.

El paradigma tradicional de la medición empieza, así, a cambiar. En lugar de registrar una lectura como traducción abstracta de alguna propiedad de un fenómeno sin posibilidad de verificación ulterior, se conserva el propio objeto a medir con el mínimo grado de abstracción que puede lograrse con el equipamiento disponible. Si bien el objeto original es un campo sonoro de compleja distribución espacial, temporal y espectral, al ser captado con un único micrófono ya se ha ejecutado un proceso de abstracción. Pero es el mismo proceso de abstracción que lleva a cabo el transductor y acondicionador de señal de entrada de cualquier instrumento. Es, por lo tanto, el mismo proceso de abstracción del que partirá el sonómetro que utilicemos, junto con sus costosos módulos opcionales de software, para elaborar y mostrar los resultados de la medición.

La propuesta de este libro es examinar los procesos de medición tradicionales y presentar alternativas a implementar en software de procesamiento matemático, junto con un análisis de su compatibilidad con las normas acústicas. También se mostrarán ejemplos de cálculo de parámetros que en general no se encuentran disponibles en los instrumentos comerciales. En todos los casos la propuesta involucra la mínima inversión posible en hardware.

Dado que por una cuestión de espacio no es posible incluir en el texto más que breves extractos de algunos programas que implementan las ideas presentadas, se irán incluyendo ejemplos ampliados en el sitio web del Laboratorio de Acústica y Electroacústica de la Universidad Nacional de Rosario. Los mismos podrán descargarse de <http://www.fceia.unr.edu.ar/acustica/mediciones>.

Federico Miyara  
Rosario, agosto de 2013



## Contenidos

Capítulo 1	GENERALIDADES	1
1.1	Mediciones acústicas	1
1.2	Valor instantáneo y valor eficaz	1
1.3	Nivel de presión sonora	2
1.4	Nivel equivalente	3
1.4.1	Tiempo de estabilización de $L_{eq}$	5
1.5	Niveles ponderados	7
1.6	Niveles estadísticos	13
1.6.1	Tiempo de estabilización de los niveles estadísticos	14
1.7	Análisis de espectro	17
1.7.1	Analizador de espectro de porcentaje constante	17
1.7.2	Analizador de espectro de líneas	21
1.7.3	Espectro de líneas e incertidumbre tiempo-frecuencia	22
1.7.4	Espectro de líneas y densidad espectral	22
1.7.5	Filtros de banda	26
1.7.6	Respuesta transitoria de los filtros de banda	30
1.7.7	Igualdad de Parseval	34
1.7.8	Efecto de la tolerancia del espectro en la igualdad de Parseval	36
Capítulo 2	INCERTIDUMBRE	39
2.1	Introducción	39
2.2	Resolución, precisión, exactitud	40
2.3	Método y procedimiento de medición	41
2.3.1	Métodos directos e indirectos de medición	41
2.4	Modelo de medición	42
2.5	Incetidumbre	44
2.5.1	Incetidumbre tipo A	45
2.5.2	Incetidumbre tipo B	46
2.5.3	Incetidumbre expandida	47
2.5.4	Incetidumbre estándar combinada	47
2.6	Ejemplos	49
2.6.1	Relación entre las incertidumbres en el nivel y en la presión	50
2.6.2	Incetidumbre en la corrección de condiciones ambientales	50
2.6.3	Incetidumbre en el cálculo de nivel equivalente	51
2.6.4	Incetidumbre en el nivel A a partir del espectro	53
2.6.5	Incetidumbre en la medición del índice de reducción acústica	54
2.7	Incetidumbre y resolución	55
2.8	Incetidumbre y error sistemático	56
2.8.1	Error sistemático aditivo	56
2.8.2	Error sistemático multiplicativo	58
2.8.3	Error sistemático no lineal	62
2.8.4	Errores sistemáticos no corregidos	65
2.9	Cálculos de incertidumbre en cadena	65
Capítulo 3	REGISTRO DIGITAL	67
3.1	Introducción	67
3.2	Audio digital	67
3.3	Muestreo	68
3.4	Digitalización	72
3.5	Relación señal a ruido	74
3.6	Distorsión a bajos niveles	75
3.7	Dither	75
3.8	Jitter	79
3.9	Conversión D/A y A/D	80

3.9.1	Conversión digital / analógica .....	80
3.9.2	Conversión analógica / digital .....	81
3.10	Modulación de código de pulsos (PCM) .....	83
3.11	Modulación de código de pulsos diferencial (DPCM) .....	84
3.12	Modulación delta (DM) .....	87
3.13	Modulación sigma-delta (SDM) .....	89
3.14	Dispositivos de registro digital .....	92
<b>3.15</b>	<b>Formatos de archivos de sonido .....</b>	<b>95</b>
3.15.1	Formato WAV .....	96
<b>3.15.2</b>	<b>Formato FLAC .....</b>	<b>103</b>
3.16	Archivos de proyecto .....	111
3.16.1	Formato .AUP .....	111
3.16.2	Formato .AU .....	117
3.16.3	Formato .AUF .....	118
3.16.4	Acceso externo a los datos de audio .....	118
3.17	Medios de registro y almacenamiento .....	118
3.17.1	Disco duro .....	119
3.17.2	Memoria flash .....	123
3.17.3	Discos ópticos .....	125
3.17.4	Cinta de audio digital (DAT) .....	138
3.18	Sistemas de archivos .....	138
3.18.1	FAT 32 .....	139
3.18.2	NTFS .....	140
3.18.3	EXT4 .....	142
3.18.4	HFS+ .....	143
3.18.5	UDF .....	143
3.19	Preservación a largo plazo .....	144
3.20	Conclusión .....	146
Capítulo 4	EDICIÓN DE AUDIO DIGITAL .....	147
4.1	Introducción .....	147
4.2	Audacity .....	148
4.2.1	Abrir un archivo preexistente .....	149
4.2.2	Grabar sonidos .....	150
4.2.3	Generar señales .....	151
4.2.4	Agregar pistas .....	151
4.2.5	Guardar un proyecto .....	151
4.2.6	Selección .....	152
4.2.7	Marcas y etiquetas .....	153
4.2.8	Selección en presencia de etiquetas .....	154
4.2.9	Tono de calibración .....	154
4.2.10	Filtros FFT .....	156
4.2.11	Espectrograma y análisis de espectro .....	158
4.2.12	Reducción de ruido .....	161
Capítulo 5	TRANSDUCTORES .....	163
5.1	Micrófonos .....	163
5.2	Polarización .....	165
5.3	Preamplificador .....	165
5.4	Campos sonoros .....	165
5.4.1	Campo libre .....	166
5.4.2	Campo de presión .....	166
5.4.3	Campo difuso .....	167
5.4.4	Campo estacionario .....	167
5.5	Micrófonos y campos sonoros .....	167
5.6	Respuesta en frecuencia .....	168

5.7	Respuesta direccional .....	170
5.8	Ruido .....	172
5.9	Distorsión .....	173
5.10	Micrófonos micromaquinados .....	174
5.10.1	Respuesta en frecuencia .....	174
5.11	Auriculares audiométricos .....	175
5.12	Fuentes omnidireccionales .....	178
Capítulo 6	PROCESAMIENTO DIGITAL DE SEÑALES .....	181
6.1	Señales discretas .....	181
6.2	Impulso discreto .....	181
6.3	Señal como convolución con un impulso discreto .....	182
6.4	Sistemas discretos .....	183
6.5	Sistemas con respuesta al impulso finita e infinita .....	184
6.6	Ecuación diferencia de un sistema discreto .....	184
6.7	Respuesta en frecuencia de un sistema discreto .....	186
6.8	Transformada $z$ de una señal discreta .....	188
6.9	Transformada $z$ de una ecuación diferencia .....	189
6.10	Transformada $z$ de una convolución .....	190
6.11	Transformada $z$ y respuesta en frecuencia .....	191
6.12	Soluciones de una ecuación diferencia .....	191
6.13	Polos y estabilidad de un sistema discreto .....	193
6.14	Inversión de una transformada $z$ racional .....	193
6.15	Conversión bilineal de un sistema continuo en discreto .....	194
Capítulo 7	ALGORITMOS BÁSICOS PARA MEDICIONES ACÚSTICAS .....	196
7.1	Introducción .....	196
7.2	Carga de un archivo .wav .....	196
7.3	Promedio energético y nivel equivalente .....	197
7.4	Calibración .....	197
7.5	Envolvente energética .....	199
7.6	Ponderación A .....	200
7.7	Análisis estadístico .....	202
Capítulo 8	ANÁLISIS DE ESPECTRO .....	206
8.1	Introducción .....	206
8.2	Paradigmas para el análisis de espectro .....	206
8.3	Filtros digitales .....	207
8.4	Transformada discreta de Fourier (DFT) .....	211
8.5	Transformada rápida de Fourier (FFT) .....	213
8.6	Análisis de espectro mediante la FFT .....	214
8.6.1	Análisis de espectro de líneas .....	214
8.6.2	El problema de las frecuencias próximas a $F_s/2$ .....	216
8.6.3	El problema de las frecuencias subarmónicas .....	218
8.6.4	Detección precisa de tonos puros .....	218
8.6.5	Ventanas .....	220
8.6.6	Análisis de espectro de bandas con FFT .....	223
8.6.7	Densidad espectral y promediación de espectros .....	226
8.7	Filtros FFT .....	227
8.8	Algunas aplicaciones de la FFT .....	231
8.8.1	Determinación de carácter tonal .....	231
8.8.2	Convolución mediante FFT .....	234
8.8.3	Correlación mediante FFT .....	236
8.8.4	Filtros de bandas críticas .....	238
8.8.5	Determinación de funciones transferencia .....	238
8.9	Contraste de algoritmos para uso en mediciones .....	239
8.9.1	Norma IEC 61260 .....	239

8.9.2	Contraste de algoritmos .....	240
8.9.3	Procedimiento .....	240
8.9.4	Calibración de la placa de sonido .....	241
8.9.5	Constatación de la respuesta del analizador.....	242
8.9.6	Contraste de diversos algoritmos .....	242
8.9.7	Resultados .....	243
8.9.8	Conclusión .....	248
Capítulo 9	ENSAYO DE GRABADORES DIGITALES .....	249
9.1	Introducción .....	249
9.2	Especificaciones del grabador digital.....	249
9.3	Ensayos .....	250
9.3.1	Respuesta en frecuencia.....	250
9.3.2	Ruido.....	252
9.3.3	Linealidad .....	252
9.3.4	Respuesta transitoria .....	254
9.3.5	Incertidumbre.....	254
9.3.6	Conclusiones.....	255
	Referencias .....	256
	Apéndice 1 Glosario y definiciones sobre metrología .....	262
	Apéndice 2 Conceptos fundamentales de Estadística .....	271
	Apéndice 3 Dispersión estadística del valor eficaz de un ruido estacionario en función del tiempo de integración .....	286
	Apéndice 4 Dispersión estadística del valor eficaz de un ruido no estacionario en función del tiempo de integración .....	288
	Apéndice 5 Dispersión estadística de los percentiles.....	291
	Apéndice 6 Envolvente de un ruido filtrado .....	293
	Apéndice 7 Respuesta transitoria de un filtro pasabanda de orden 3 .....	295
	Apéndice 8 Incertidumbre combinada .....	297
	Apéndice 9 Ejemplo de cálculo de incertidumbre en caso de un error sistemático no lineal... 301	
	Apéndice 10 Teorema del muestreo .....	307
	Apéndice 11 Estructura de un archivo FLAC .....	313
	Apéndice 12 Definición de tipo de documento (DTD) para proyectos de Audacity .....	326
	Apéndice 13 Breve descripción de Scilab.....	333
	Apéndice 14 Transformada rápida de Fourier (FFT) .....	340
	Apéndice 15 Igualdad de Parseval y simetría de la DFT .....	343
	Apéndice 16 Espectro de las ventanas de análisis.....	346
	Índice alfabético .....	351

## Capítulo 1

# GENERALIDADES

### 1.1 Mediciones acústicas

A diferencia de la medición de magnitudes físicas esencialmente invariantes, como la longitud de un cuerpo o su masa, el objeto de las mediciones acústicas es extraer información representativa de una señal variable en el tiempo y a menudo correspondiente a un fenómeno irrepetible.<sup>1</sup> Sólo ocasionalmente interesa la historia temporal completa de una señal acústica con fines metrológicos, siendo preferible, en general, informar parámetros globales del fenómeno bajo estudio. En la mayoría de los casos, por cuestiones de economía sólo se mide durante un tiempo relativamente corto, y de la información recogida se busca obtener valores característicos correspondientes a un tiempo mucho mayor.

Por otra parte, dado que el campo sonoro varía de un punto a otro del espacio, la ubicación e incluso la orientación del sensor (micrófono) también tienen una influencia importante en el resultado. En algunos casos, el propio sensor, el instrumento o los accesorios utilizados (y hasta la presencia misma del operador) pueden alterar significativamente la señal a medir.

Lo anterior lleva a que la incertidumbre en la medición involucre tanto al instrumento de medición como al operador, la oportunidad y la posición del sensor, e incluso a las características del propio fenómeno a investigar.

### 1.2 Valor instantáneo y valor eficaz

La magnitud acústica más fácil de medir es la presión sonora  $p$ . Su valor instantáneo  $p(t)$  es, no obstante, poco ilustrativo de los fenómenos acústicos ya que dos sonidos muy similares tanto auditivamente como desde el punto de vista de sus efectos podrían tener valores instantáneos muy diferentes. Desde el punto de vista de los efectos del sonido es más interesante el promedio energético o *valor eficaz* extendido a un intervalo de tiempo  $T$  que depende de la aplicación. Se lo define como

$$P_{ef,T} = \sqrt{\frac{1}{T} \int_0^T p^2(t) dt} . \quad (1.1)$$

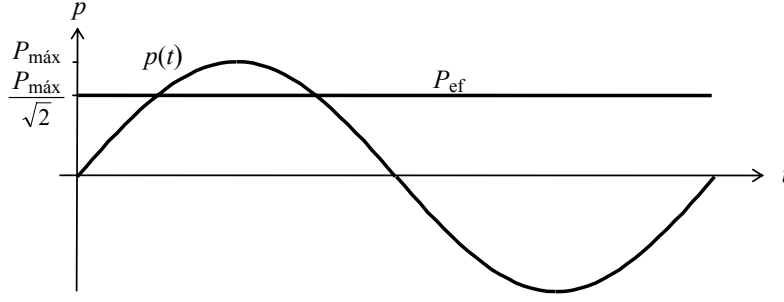
Como ejemplo, en la figura 1 se muestra el valor eficaz para de una onda senoidal.

A veces conviene tener un promedio energético móvil, es decir, que se actualice en forma continua en el tiempo:

$$P_{ef,T}(t) = \sqrt{\frac{1}{T} \int_t^{t+T} p^2(\theta) d\theta} . \quad (1.2)$$

---

<sup>1</sup> En realidad cualquier magnitud física a medir fluctúa en el tiempo. La longitud de un objeto varía según las variaciones aleatorias de temperatura, la masa varía con la deposición de partículas, oxidación, sublimación, etc. La diferencia está más bien en el modelo subyacente de interés para un determinado problema. Así, en la mayoría de los casos la longitud de un objeto se supone constante, a diferencia de la presión sonora, que se considera esencialmente variable.



**Figura 1.** Una onda senoidal y su valor eficaz  $P_{ef}$  extendido a un período

La ecuación (1.2) proporciona la *envolvente energética* de la señal. En este caso, la selección de  $T$  dependerá de la rapidez de variación de la señal, la cual a su vez depende de la duración de los eventos individuales que componen la señal. Así, para un ruido de tránsito,  $T$  puede elegirse grande, pero para un ruido formado por una sucesión de golpes contendrá un valor de  $T$  menor. La envolvente energética puede obtenerse reemplazando la integración por un filtro pasabajos de constante de tiempo  $\tau$ . En ese caso resulta:

$$P_{\tau}(t) = \sqrt{\frac{1}{\tau} \int_0^t p^2(\theta) e^{-\frac{t-\theta}{\tau}} d\theta} . \quad (1.3)$$

Éste es un promedio con ponderación exponencial. Como se puede apreciar, los valores de  $p^2(\theta)$  reciben la mayor ponderación para  $\theta \cong t$ . En otras palabras, se prioriza el pasado reciente de la señal. En la figura 2 se muestra el resultado de aplicar esta promediación a una señal para tres constantes de tiempo  $\tau$  diferentes. La constante de tiempo  $\tau$  se suele denominar *ponderación temporal*.

Las envolventes obtenidas según (1.2) y (1.3) no son iguales, aunque son parecidas. Si bien la ecuación (1.3) parece más complicada, en la práctica el algoritmo con el que se calcula  $P_{\tau}(t)$  es computacionalmente más eficiente.<sup>2</sup>

### 1.3 Nivel de presión sonora

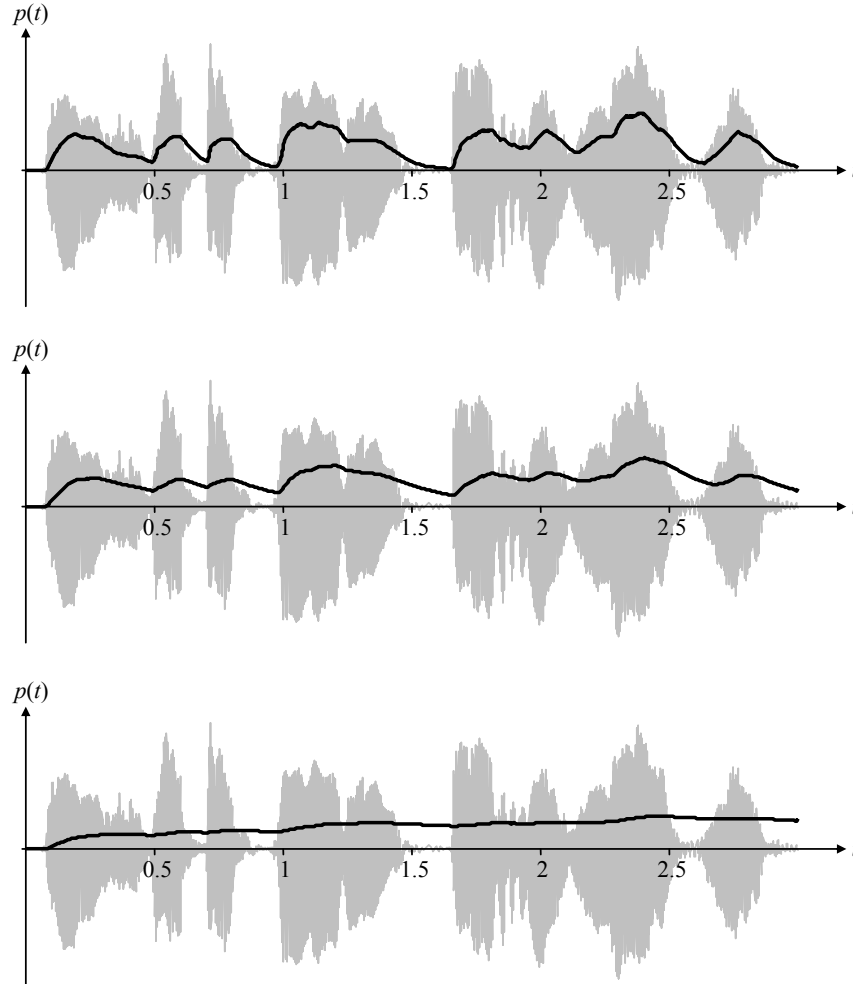
La presión sonora correspondiente a sonidos audibles tiene un rango dinámico muy grande ( $10^6:1$ ), razón por la cual se utiliza en general el *nivel de presión sonora*,  $L_{p,\tau}$ , definido como

$$L_{p,\tau} = 20 \log_{10} \frac{P_{\tau}}{P_{ref}} , \quad (1.4)$$

donde  $P_{ref} = 20 \mu\text{Pa}$  es la presión de referencia, que corresponde aproximadamente al umbral auditivo a 1000 Hz.  $L_{p,\tau}$  se expresa en decibeles, dB. De esta manera el rango de valores queda restringido al intervalo entre 0 dB y 120 dB. El nivel de presión sonora

<sup>2</sup> Dicho algoritmo no consiste en calcular numéricamente la integral sino en aplicar un filtro digital que simula un filtro analógico de primer orden.

es, en general, variable en el tiempo según la envolvente energética. Cuanto mayor sea la constante de tiempo  $\tau$  más lenta será la variación. Se encuentran normalizadas tres constantes de tiempo: 1 s para la respuesta S (lenta), 0,125 s para la respuesta F (rápida) y 0,035 s para la respuesta I (impulsiva). En este último caso, para facilitar la lectura de los instrumentos, la constante de tiempo de bajada se toma igual a 1,5 s. Los valores correspondientes se anotan  $L_{p,S}$ ,  $L_{p,F}$ ,  $L_{p,I}$ .



**Figura 2.** Ejemplo de señal (correspondiente a una emisión vocal) y la envolvente energética para tres constantes de tiempo diferentes. De arriba a abajo,  $\tau = 35$  ms;  $\tau = 125$  ms; y  $\tau = 1$  s. Cuanto menor es la constante de tiempo más detalles temporales se preservan.

#### 1.4 Nivel equivalente

El *nivel equivalente*,  $L_{eq,T}$ , también llamado *nivel de presión sonora continuo equivalente* (o *nivel sonoro continuo equivalente*), es un valor de presión sonora hipotéticamente constante durante un periodo  $T$  que posee la misma energía total que la pre-

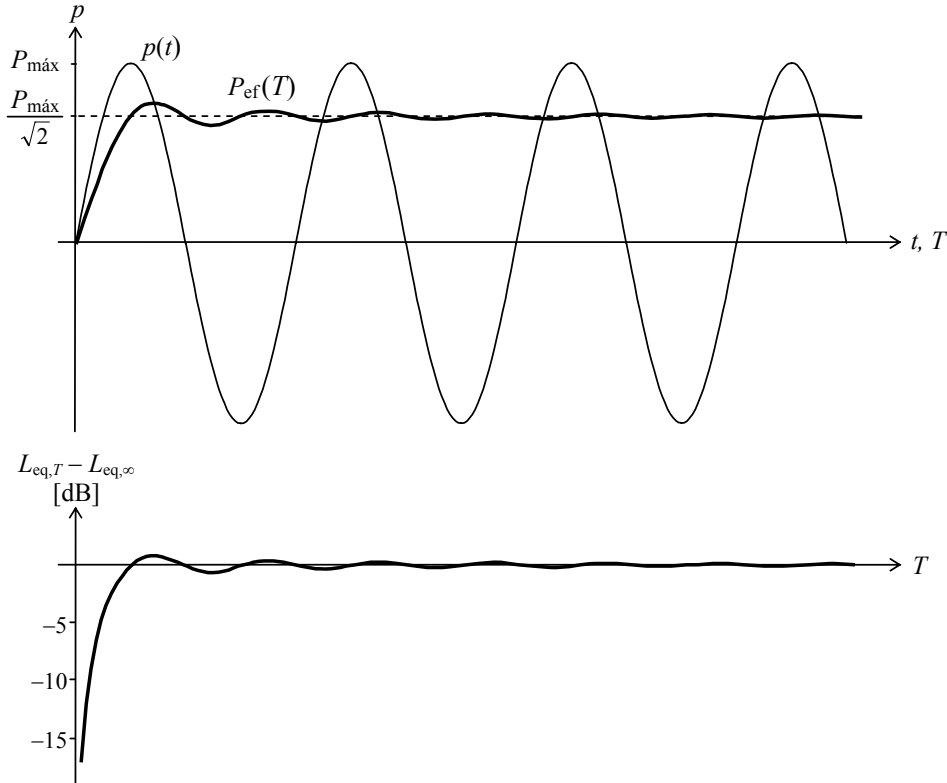
sión sonora variable del sonido a medir durante ese tiempo. Este valor no es otra cosa que el valor eficaz extendido al tiempo  $T$ , obtenido por integración de la presión cuadrática (que es proporcional a la potencia acústica) y expresado en decibeles:

$$L_{\text{eq},T} = 20 \log \frac{\sqrt{\frac{1}{T} \int_0^T p^2(t) dt}}{P_{\text{ref}}} = 10 \log \frac{\frac{1}{T} \int_0^T p^2(t) dt}{P_{\text{ref}}^2}. \quad (1.5)$$

El nivel equivalente depende del periodo  $T$  de medición, por lo que es posible representarlo en función del tiempo. Por ejemplo, en el caso de una onda senoidal de frecuencia angular  $\omega$  resulta

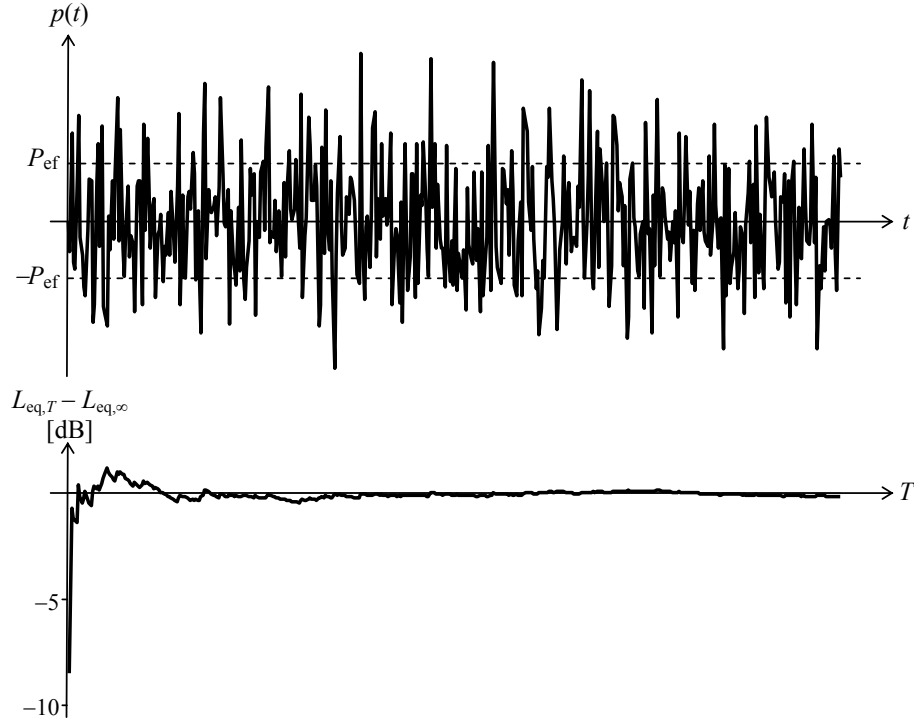
$$L_{\text{eq},T} = L_{\text{eq},\infty} + 10 \log \left( 1 - \frac{1}{2\omega T} \sin 2\omega T \right). \quad (1.6)$$

Este comportamiento se indica en la figura 3. En el caso de un ruido blanco se tiene un comportamiento similar que se muestra en la figura 4.



**Figura 3.** Arriba, valor eficaz  $P_{\text{ef}}$  de una onda senoidal en función del tiempo de integración  $T$ . Abajo, su expresión logarítmica, el nivel equivalente  $L_{\text{eq},T}$ , relativo al nivel equivalente a largo plazo  $L_{\text{eq},\infty}$ .





**Figura 4.** Arriba, oscilograma de un ruido blanco. Abajo, su nivel equivalente  $L_{eq,T}$ , relativo al nivel equivalente a largo plazo  $L_{eq,\infty}$ .

Vemos en estos ejemplos que a medida que se incrementa el tiempo de integración  $T$  el valor eficaz se aproxima cada vez más al valor eficaz a largo plazo (es decir, el que se obtendría integrando durante un tiempo infinito o, en la práctica, durante un tiempo muy largo).

#### 1.4.1 Tiempo de estabilización de $L_{eq}$

En general es posible encontrar un *tiempo de estabilización* de la medición  $T_{e,\Delta L}$  después del cual el nivel equivalente se mantiene dentro de un margen  $\pm\Delta L$  del nivel equivalente a largo plazo. En el caso de un tono puro deberá ser

$$\left| 10 \log \left( 1 - \frac{1}{2\omega T} \sin 2\omega T \right) \right| < \Delta L. \quad (1.7)$$

Aproximando el logaritmo por un polinomio de Taylor de primer grado, resulta

$$\frac{10}{\ln 10} \left| \frac{\sin 2\omega T}{2\omega T} \right| < \Delta L,$$

lo cual se logra si

$$T > \frac{5}{\ln 10} \frac{1}{\omega \Delta L} = \frac{0,346}{f \Delta L}. \quad (1.8)$$

Por ejemplo, para una frecuencia de 100 Hz el tiempo de estabilización para  $\pm 0,1$  dB resulta 34,6 ms.

Para el caso de un ruido blanco gaussiano (distribución normal de amplitudes) el tiempo de estabilización debe expresarse en términos estadísticos. En la situación más común podemos suponer que se trata de un ruido blanco de banda acotada, es decir, cuyo espectro es constante hasta una cierta frecuencia superior  $f_{\text{máx}}$  y, a partir de allí, nulo.<sup>3</sup> Se demuestra en el Apéndice 3 que el desvío estándar de la presión eficaz cuadrática (o valor cuadrático medio) extendida a un tiempo  $T$  está dado por

$$\sigma_{P_{\text{ef},T}^2} \cong \sqrt{\frac{1}{f_{\text{máx}}T}} P_{\text{ef},\infty}^2, \quad (1.9)$$

donde  $P_{\text{ef},\infty}^2$  es la presión eficaz cuadrática a largo plazo. Aproximando la distribución por una normal,<sup>4</sup> podemos, con una confiabilidad del 99,7 %, acotar a un margen de  $\pm 3\sigma$  alrededor del valor esperado (que coincide con el valor a largo plazo):

$$P_{\text{ef},\infty}^2 - 3 \sigma_{P_{\text{ef},T}^2} < P_{\text{ef},T}^2 < P_{\text{ef},\infty}^2 + 3 \sigma_{P_{\text{ef},T}^2}, \quad (1.10)$$

o bien,

$$1 - 3 \frac{\sigma_{P_{\text{ef},T}^2}}{P_{\text{ef},\infty}^2} < \frac{P_{\text{ef},T}^2}{P_{\text{ef},\infty}^2} < 1 + 3 \frac{\sigma_{P_{\text{ef},T}^2}}{P_{\text{ef},\infty}^2} \quad (1.11)$$

Llamando  $\Delta L$  al error entre el nivel equivalente obtenido integrando durante un tiempo  $T$  y el nivel equivalente a largo plazo, es decir

$$\Delta L = L_{\text{eq},T} - L_{\text{eq},\infty}, \quad (1.12)$$

de (1.11) y (1.9) resulta

$$10 \log \left( 1 - \frac{3}{\sqrt{f_{\text{máx}}T}} \right) < \Delta L < 10 \log \left( 1 + \frac{3}{\sqrt{f_{\text{máx}}T}} \right). \quad (1.13)$$

Expandiendo el logaritmo en serie de Taylor de primer grado, lo cual vale para  $T$  elevado, podemos reescribir la anterior como

$$|\Delta L| < \frac{10}{\ln 10} \frac{3}{\sqrt{f_{\text{máx}}T}}. \quad (1.14)$$

<sup>3</sup> No existe un ruido blanco ideal, ya que tendría que tener un ancho de banda infinito. Quizás lo que más se le aproxima es el ruido térmico de una resistencia eléctrica, que contiene energía hasta unos 100 GHz.

<sup>4</sup> Realmente no puede ser normal ya que los valores de la presión cuadrática nunca son negativos, pero si el desvío estándar es pequeño, en las proximidades del valor medio es aproximadamente normal.

Podemos garantizar que  $|\Delta L| < \Delta L_{\text{máx}}$  si adoptamos un tiempo  $T$  de promediación

$$T > \left( \frac{30}{\ln 10} \right)^2 \frac{1}{f_{\text{máx}} \Delta L_{\text{máx}}^2} = \frac{169,8}{f_{\text{máx}} \Delta L_{\text{máx}}^2}. \quad (1.15)$$

Por ejemplo, si  $f_{\text{máx}} = 22050$  Hz y  $\Delta L_{\text{máx}} = 0,1$  dB, entonces el tiempo de estabilización resulta  $T > 0,77$  s.<sup>5</sup> Esto implica que en una integración durante 0,77 s el nivel equivalente obtenido difiere del nivel equivalente a largo plazo en 0,1 dB o menos con una confiabilidad del 99,7 %.

Cuando la señal no es estacionaria el tiempo de estabilización puede ser considerablemente mayor, sobre todo cuando la señal está formada por la superposición de una serie de eventos bastante espaciados en el tiempo. Un ejemplo es el del ruido del tránsito. Se demuestra en el Apéndice 4 que el tiempo de estabilización con un error máximo  $\Delta L_{\text{máx}}$  y una confiabilidad del 99,7 % es

$$T_{\text{estab}} = \frac{9}{Q} \left( \frac{\sigma_{L_{\text{eq},k}}}{\Delta L_{\text{máx}}} \right)^2, \quad (1.16)$$

donde  $Q$  es la cantidad de vehículos por unidad de tiempo y  $\sigma_{L_{\text{eq},k}}$  es el desvío estándar de los niveles equivalentes de los diferentes vehículos individuales.

Por ejemplo, si  $\sigma_{L_{\text{eq},k}} = 1$  dB,  $\Delta L_{\text{máx}} = 0,1$  dB y se tiene un flujo de 600 vehículos por hora, resulta

$$T_{\text{estab}} = \frac{9}{600/3600} \left( \frac{1}{0,1} \right)^2 = 5400 \text{ s}. \quad (1.17)$$

Se observa que el tiempo de estabilización con una confiabilidad del 99,7 % es mucho más largo que en el caso de un ruido estacionario. El tiempo de estabilización se reduce considerablemente si el error admisible es mayor.

El concepto esencial, común a todos los casos analizados, es que la estabilización se logra permitiendo una cantidad de eventos lo suficientemente grande como para reducir la variabilidad. En el caso del tono puro, los eventos son los ciclos, por lo tanto a mayor frecuencia, menor el tiempo necesario para completar la cantidad requerida de eventos. En el caso del ruido filtrado, son las muestras. Debido a la elevada tasa de muestreo, se requiere muy poco tiempo para alcanzar la estabilización. En el caso del ruido del tránsito, los eventos son las pasadas de vehículos individuales. Por estar bastante separados en el tiempo (en el ejemplo, uno cada 6 s), es preciso un tiempo muy largo para alcanzar un número de eventos suficiente.

### 1.5 Niveles ponderados

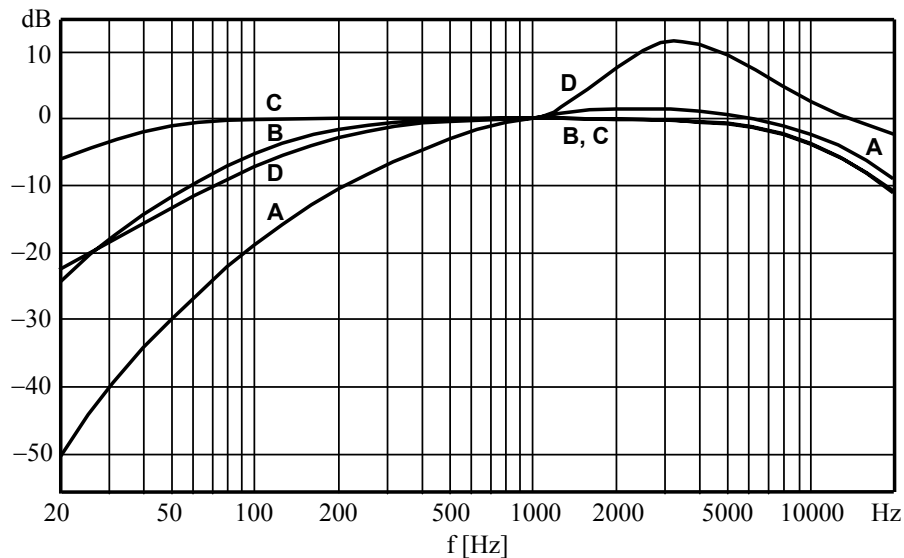
La mayoría de los criterios acústicos vinculados con la percepción o los efectos del ruido sobre el ser humano (daño auditivo, molestia, efectos sobre desempeño) están

<sup>5</sup> La frecuencia máxima  $f_{\text{máx}} = 22050$  Hz se adopta frecuentemente en audio digital debido a que es la máxima frecuencia representable mediante la tasa de muestreo estándar de los discos compactos, es decir, 44100 Hz.

expresados utilizando alguna forma de ponderación (o compensación) frecuencial. El propósito es tener en cuenta que la sensibilidad auditiva cambia con la frecuencia. Las primeras curvas de ponderación que se introdujeron fueron las A, B y C. Aunque originalmente estaban pensadas para tener en cuenta la respuesta del oído a niveles cercanos a 40 dB, 70 dB y 100 dB respectivamente,<sup>6</sup> más adelante se utilizarían, especialmente la A y la C, para evaluar ruidos de cualquier nivel, incluso muy elevados. La razón para ello es que numerosas investigaciones realizadas utilizando ponderación A demostraron correlaciones razonables con efectos tales como la molestia o el daño auditivo, incluso a niveles muy altos. No obstante, en la actualidad el uso de la ponderación A se encuentra en discusión, ya que los efectos subjetivos del ruido dependen de muchas otras cualidades del sonido que van desde el contenido espectral hasta el contenido semántico.

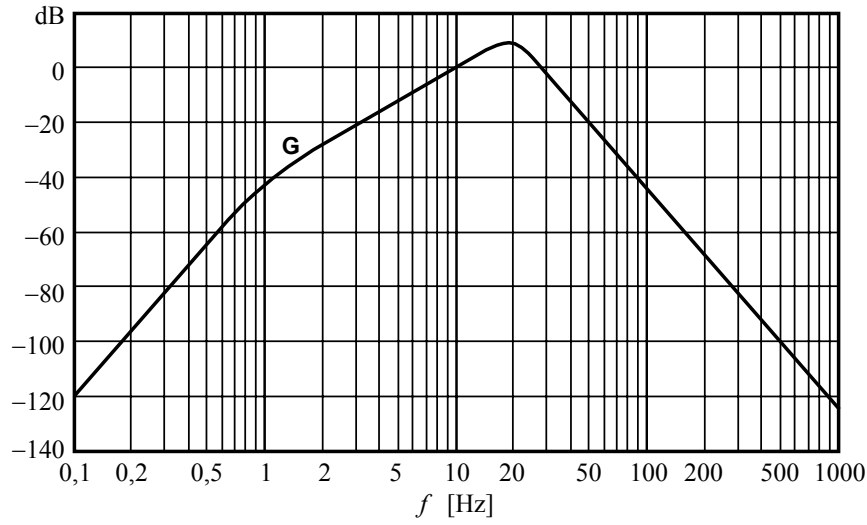
Posteriormente se introdujeron otras ponderaciones aplicables a casos particulares, como por ejemplo la ponderación D utilizada para ruido de aviación, o la ponderación G para ruidos de muy baja frecuencia.

En la figura 5 se muestran las curvas de respuesta en frecuencia correspondientes a las ponderaciones A, B, C y D y en la figura 6, la correspondiente a la ponderación G. Las ecuaciones (1.18) a (1.22) dan las expresiones analíticas de dichas respuestas en frecuencia. En las tablas 1.1 y 1.2 se indican valores numéricos de las respuestas.



**Figura 5.** Curvas de ponderación (o compensación) A, B, C y D. Para cada frecuencia, el valor de la ordenada representa la corrección aditiva a aplicar al nivel de presión sonora de un tono de esa frecuencia para obtener su nivel sonoro. En 1 kHz todas las curvas coinciden en 0 dB.

<sup>6</sup> En forma más precisa, estas curvas se aplicarían a sonidos cuyos niveles de sonoridad fueran de 40 fon, 70 fon y 100 fon. El nivel de sonoridad es numéricamente igual al nivel de presión sonora de un tono de 1 kHz igualmente sonoro que el sonido a evaluar.



**Figura 6.** Curva de ponderación G para medición de ruidos de baja frecuencia. La ordenada representa la corrección aditiva a aplicar al nivel de presión sonora de un tono de esa frecuencia para obtener su nivel ponderado. En 10 Hz la corrección es 0 dB.

$$A(f) = 20 \log \frac{1,2588 \times 12194^2 f^4}{(f^2 + 20,6^2) \sqrt{f^2 + 107,7^2} \sqrt{f^2 + 737,9^2} (f^2 + 12194^2)} \quad (1.18)$$

$$B(f) = 20 \log \frac{1,01972 \times 12194^2 f^3}{(f^2 + 20,6^2) \sqrt{f^2 + 158,5^2} (f^2 + 12194^2)} \quad (1.19)$$

$$C(f) = 20 \log \frac{1,0071 \times 12194^2 f^2}{(f^2 + 20,6^2) (f^2 + 12194^2)} \quad (1.20)$$

$$D(f) = 20 \log \frac{14499,7 f \sqrt{(f^2 - 1018,78^2)^2 + 1039,6^2 f^2}}{\sqrt{f^2 + 282,71^2} \sqrt{f^2 + 1160^2} \sqrt{(f^2 - 3126,4^2)^2 + 3424,1^2 f^2}} \quad (1.21)$$

$$G(f) = 20 \log (630673,8 G_1 G_2 G_3 G_4) \quad (1.22)$$

donde

$$\begin{aligned}
 G_1 &= \frac{f}{\sqrt{(f^2 - 1)^2 + 2f^2}} \\
 G_2 &= \frac{f}{\sqrt{(f^2 - 397,96)^2 + 1485,3f^2}} \\
 G_3 &= \frac{f}{\sqrt{(f^2 - 3981,8)^2 + 796,37f^2}} \\
 G_4 &= \frac{f}{\sqrt{(f^2 - 397,96)^2 + 106,5f^2}}
 \end{aligned}
 \tag{1.23}$$

**Tabla 1.1.** Valores de las correcciones de las curvas de ponderación **A**, **B**, **C** y **D** para la serie de frecuencias estándar de octavas (Normas IEC 61672, IEC 60651/79 e IEC 60537).

Frecuencia [Hz]	Curva A [dB]	Curva B [dB]	Curva C [dB]	Curva D [dB]
10	-70,4	-38,2	-14,3	-26,0
12,5	-63,4	-33,2	-11,2	-24,7
16	-56,7	-28,5	-8,5	-22,6
20	-50,5	-24,2	-6,2	-20,6
25	-44,7	-20,4	-4,4	-18,7
31,5	-39,4	-17,1	-3,0	-16,7
40	-34,6	-14,2	-2,0	-14,7
50	-30,2	-11,6	-1,3	-12,8
63	-26,2	-9,3	-0,8	-10,9
80	-22,5	-7,4	-0,5	-8,9
100	-19,1	-5,6	-0,3	-7,2
125	-16,1	-4,2	-0,2	-5,6
160	-13,4	-3,0	-0,1	-3,9
200	-10,9	-2,0	0,0	-2,6
250	-8,6	-1,3	0,0	-1,6
315	-6,6	-0,8	0,0	-0,8
400	-4,8	-0,5	0,0	-0,4
500	-3,2	-0,3	0,0	-0,3
630	-1,9	-0,1	0,0	-0,5
800	-0,8	-0,0	0,0	-0,6
1000	0,0	0,0	0,0	0,0
1250	0,6	-0,0	0,0	1,9
1600	1,0	-0,0	-0,1	5,0
2000	1,2	-0,1	-0,2	7,9
2500	1,3	-0,2	-0,3	10,3
3150	1,2	-0,4	-0,5	11,5
4000	1,0	-0,7	-0,8	11,1
5000	0,5	-1,2	-1,3	9,6
6300	-0,1	-1,9	-2,0	7,6
8000	-1,1	-2,9	-3,0	5,5
10000	-2,5	-4,3	-4,4	3,4
12500	-4,3	-6,1	-6,2	1,4
16000	-6,6	-8,4	-8,5	-0,8
20000	-9,3	-11,1	-11,2	-2,7

**Tabla 1.2.** Valores de las correcciones de la curva de ponderación **G** para la serie de frecuencias estándar de octavas (Norma ISO 7196).

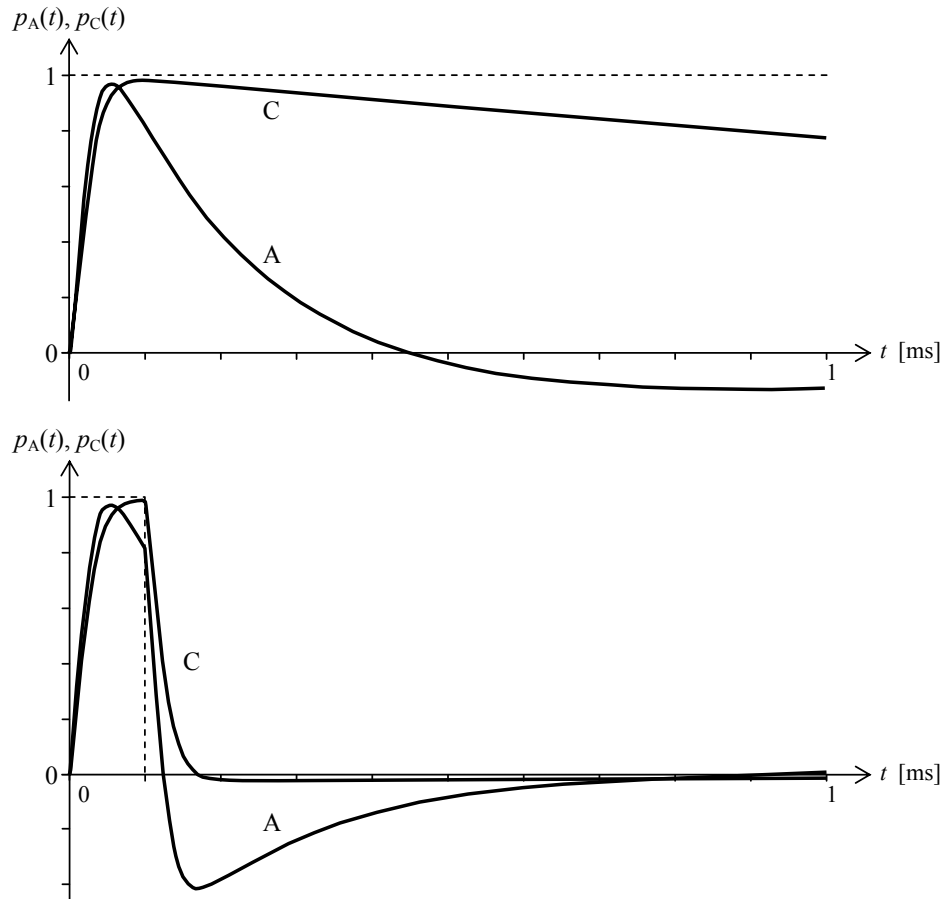
Frecuencia [Hz]	Curva G [dB]
0,25	-88,0
0,315	-80,0
0,4	-72,1
0,5	-64,3
0,63	-56,6
0,8	-49,5
1	-43,0
1,25	-37,5
1,6	-32,6
2	-28,3
2,5	-24,1
3,15	-20,0
4	-16,0
5	-12,0
6,3	-8,0
8	-4,0
10	0,0
12,5	4,0
16	7,7
20	9,0
25	3,7
31,5	-4,0
40	-12,0
50	-20,0
63	-28,0
80	-36,0
100	-44,0
125	-52,0
160	-60,0
200	-68,0
250	-76,0
315	-84,0

Las ponderaciones A y C se encuentran normalizadas por la norma IEC 61672. La B, por la norma obsoleta IEC 60651 y la D por la norma IEC 60537, también obsoleta. La ponderación G está normalizada por la ISO 7196.

De estas ponderaciones la más usada es la A. Se utiliza en general para criterios sobre riesgo auditivo y sobre molestia, lo que motiva que frecuentemente se adopte en la legislación sobre higiene y seguridad laboral y sobre ruido al vecindario. Sin embargo, su validez se encuentra en discusión desde hace muchos años en diversos foros (Kogan, 2004). Prueba de ello es que, en muchos casos, a las mediciones con ponderación A se les aplican correcciones que contemplan características espectrales, temporales o semánticas particulares.

La ponderación C se utiliza para evaluar ruidos con contenido de bajas frecuencias considerable. Para ello se suele evaluar la diferencia entre el nivel ponderado C y el nivel ponderado A. Se considera que el ruido tiene alto contenido de baja frecuencia si dicha diferencia es mayor que 20 dB. Otro uso de la ponderación C es la determinación

de niveles de pico ( $L_{C,peak}$ ). Ello obedece a que su respuesta temporal es más precisa que la de la ponderación A para ruidos de corta duración. En la figura 7 se muestran las respuestas temporales de los filtros A y C ante un escalón unitario y ante un pulso rectangular de 0,1 ms (representativo de un ruido fuertemente impulsivo). Como se observa, el pico alcanzado por la respuesta del filtro C es más cercano al pico real.



**Figura 7.** Arriba, respuestas de las redes de ponderación A y C frente a un escalón unitario. Abajo, respuestas ante un pulso de 0,1 ms de duración. En ambos casos la excitación se indica con línea de trazos.

La ponderación B ha caído en desuso y las normas actuales (por ejemplo la IEC 61672) ya no la contemplan.

La ponderación D se utiliza como método simplificado para evaluar el nivel de ruido percibido efectivo  $L_{EPN}$  (un indicador empleado en ruido de aeropuertos) de aeronaves en sobrevuelo (Yanitelli et al., 2001).

Por último, la ponderación G tiene aplicación en la evaluación de ruidos con fuertes componentes de baja frecuencia, y dado que su respuesta abarca una parte del espec-



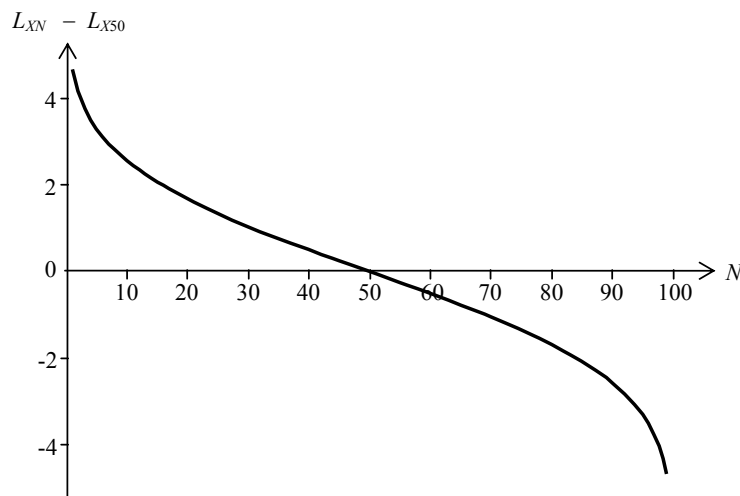
tro que queda virtualmente excluido por la ponderación A, se puede afirmar que es complementaria de ésta.

### 1.6 Niveles estadísticos

Dada una constante de tiempo  $\tau$  y una ponderación frecuencial  $X$  (donde  $X$  puede ser A, B, C, D, Z o cualquier otra), se define el *nivel estadístico*  $L_{XN}$  como aquel valor que es superado un  $N\%$  del tiempo. Así, en un caso muy simple en que el nivel aumenta linealmente entre 60 dBA y 70 dBA durante un lapso de 10 min, podemos afirmar que  $L_{A10} = 69$  dBA, ya que sólo el 10 % del tiempo (correspondiente en este caso al último minuto) se supera dicho valor.

NOTA: Aunque a veces se denomina *niveles percentiles* a los niveles estadísticos, ello es erróneo ya que el percentil  $p$  de una variable aleatoria es el valor de dicha variable que *supera* a un  $p\%$  de los valores de la misma, y no el que *es superado* por dicho porcentaje.

En la figura 8 se muestra como ejemplo la curva de los niveles estadísticos para un ruido cuyo nivel tiene distribución normal con desvío estándar  $\sigma = 2$  dB



**Figura 8.** Niveles estadísticos  $L_{XN}$  referidos a la mediana  $L_{X50}$  para un ruido cuyos niveles  $L_X$  tienen distribución normal con  $\sigma = 2$  dB

Los niveles estadísticos son muy útiles para la aplicación de ciertos criterios. Por ejemplo, para la determinación del ruido de fondo, podría especificarse el valor mínimo absoluto detectado en todo el lapso de la medición. Este valor, sin embargo, no tiene suficiente representatividad, ya que el valor mínimo no se repite entre una medición y otra. En realidad, cuanto más largo sea el periodo de medición, más bajo será el mínimo detectado. En cambio, podemos definir el ruido de fondo como  $L_{A90}$ , es decir, un valor que es superado el 90 % del tiempo. A diferencia del mínimo absoluto, mientras estamos frente a un fenómeno estacionario este valor tiende a estabilizarse,<sup>7</sup> como veremos en la próxima sección. También es posible clasificar el ruido de fondo según su proce-

<sup>7</sup> Son ejemplos de un fenómeno no estacionario la evolución del ruido a lo largo del día. Así, en horas pico el ruido de fondo definido como  $L_{A90}$  será mucho más alto que durante la madrugada.

dencia mediante los niveles estadísticos. Así, generalmente se considera que  $L_{A90}$  es el nivel del ruido de fondo de las fuentes cercanas y lejanas, en tanto que  $L_{A95}$  corresponde al nivel de las fuentes lejanas.

En forma similar, el máximo absoluto no es un valor representativo, pero sí lo es  $L_{A10}$  o  $L_{A5}$ , que también tienden a estabilizarse en mediciones suficientemente prolongadas.

### 1.6.1 Tiempo de estabilización de los niveles estadísticos

Analizaremos ahora el tiempo de estabilización de los niveles estadísticos. Para ello, supongamos que el nivel se muestrea con cierta tasa de muestreo  $F_s$ , por lo cual un tiempo de medición  $T$  equivale a una cantidad de muestras:

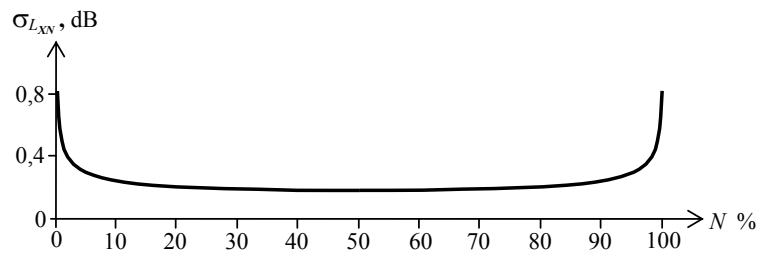
$$n = F_s T \quad (1.24)$$

La estabilización del nivel estadístico  $L_{XN}$  consiste conceptualmente en una reducción de la dispersión estadística por debajo de una cota elegida arbitrariamente con una confiabilidad dada. Por ejemplo, podríamos aceptar que el valor  $L_{XN}$  se ha estabilizado cuando el error respecto al valor final (es decir, el valor para  $T \rightarrow \infty$ ) es menor que 1 dB con una probabilidad del 95 %. Este tipo de problemas se trata, en general, analizando el efecto del número de muestras sobre el desvío estándar de la variable de interés, en nuestro caso,  $L_{XN}$ .

Existe un resultado de este tipo para los percentiles, según se detalla en el Apéndice 5. Tengamos en cuenta que el nivel estadístico  $L_{XN}$  es equivalente al percentil  $p = 100 - N$  del nivel  $L_X$ . Podemos aplicar, entonces, la ecuación (A5.7) del Apéndice 5, resultando:

$$\sigma_{L_{XN}} = \sqrt{\frac{\left(1 - \frac{N}{100}\right) \frac{N}{100}}{n}} \frac{1}{f(L_{XN})}, \quad (1.25)$$

donde  $f$  es la distribución estadística de los valores  $L_X$ . En la figura 9 se muestra un ejemplo para un nivel  $L_X$  con distribución normal. Cerca de los extremos ( $N = 0$  y  $N = 100$ ), si bien el numerador tiende a 0, la función de densidad de probabilidad tiende a 0 más rápidamente, por lo que el desvío estándar crece mucho.



**Figura 9.** Desvío estándar en dB de los niveles estadísticos  $L_N$  para un ruido cuyos niveles tienen distribución normal con  $\sigma = 2$  dB

Si suponemos que la distribución es aproximadamente normal,<sup>8</sup> sabemos que alrededor de un 99,7 % de los posibles valores de  $L_{XN}$  caerán en un intervalo que se extiende de  $\pm 3\sigma_{L_{XN}}$  a partir del valor medio, por lo tanto si admitimos un error  $\Delta L_{XN}$  la cantidad de muestras a tomar deberá ser

$$n \geq \frac{9}{(\Delta L_{XN})^2} \frac{\left(1 - \frac{N}{100}\right) \frac{N}{100}}{f(L_{XN})^2}, \quad (1.26)$$

Teniendo en cuenta (1.24), ello equivale a que el tiempo para lograr la estabilización con error  $\Delta L_{XN}$  y una confiabilidad del 99,7 % debe ser<sup>9</sup>

$$T \geq \frac{1}{F_s} \frac{9}{(\Delta L_{XN})^2} \frac{\left(1 - \frac{N}{100}\right) \frac{N}{100}}{f(L_{XN})^2}, \quad (1.27)$$

Vemos que el tiempo de estabilización aumenta recíprocamente con el cuadrado del error admisible, lo cual significa que para lograr un error muy pequeño el tiempo de medición debe ser muy prolongado. También se observa que a mayor tasa de muestreo el tiempo requerido es menor. Pero ello sólo se aplica a un nivel que varía en forma realmente aleatoria y que resulta además estacionario en el tiempo de medición. Si el ruido no es estacionario (por ejemplo el pasaje de un vehículo frente al punto de observación), podría suceder que, guiándonos por la desigualdad anterior pretendiéramos reducir el tiempo de estabilización  $T$  aumentando para ello  $F_s$ . En ese caso estaríamos tomando una porción no representativa del fenómeno.

Si la distribución de los niveles es normal con desvío estándar  $\sigma_{L_{XN}}$ , entonces, usando la ecuación (A5.6) del Apéndice 4 y teniendo en cuenta que el percentil correspondiente al nivel estadístico  $N$  es  $p = 100 - N$ ,

$$L_{XN} = \sqrt{2} \sigma_{L_X} \operatorname{erf}^{-1}\left(1 - \frac{2N}{100}\right) + \mu_{L_X}, \quad (1.28)$$

donde  $\operatorname{erf}$  es la *función error*, definida por la ecuación (A5.5) del mismo Apéndice.<sup>10</sup> La función de densidad de probabilidad calculada en este valor es

$$f(L_{XN}) = \frac{1}{\sqrt{2\pi} \sigma_{L_X}} e^{-\left(\operatorname{erf}^{-1}\left(1 - \frac{2N}{100}\right)\right)^2}. \quad (1.29)$$

<sup>8</sup> En realidad la distribución no necesariamente es normal, aunque tiende a serlo cuando el ruido a medir está formado por una gran cantidad de eventos independientes.

<sup>9</sup> Aun si la distribución no es normal, la ecuación (1.27) da correctamente el tiempo necesario para cubrir  $\pm 3\sigma$ , aunque según la distribución ello puede implicar una confiabilidad diferente al 99,7 %.

<sup>10</sup> La función error, así como su inversa, se calculan numéricamente. Los paquetes de software matemático las tienen como funciones primitivas.

Como ejemplo, obtengamos el tiempo de estabilización del  $L_{A90}$ , considerado habitualmente como el nivel del ruido de fondo, con un error de 0,5 dB y una confiabilidad del 99,7 % para el caso de un nivel sonoro A con distribución normal y un desvío estándar  $\sigma_{L_{AN}} = 2$  dBA. Para ello, de (1.29)

$$f(L_{A90}) = \frac{1}{\sqrt{2\pi} \sigma_{L_X}} e^{-\left(\text{erf}^{-1}(-0,8)\right)^2} = \frac{0,1755}{\sigma_{L_X}}, \quad (1.30)$$

que, para  $\sigma_{L_{AN}} = 2$  dBA, resulta

$$f(L_{A90}) = 0,087749. \quad (1.31)$$

Sustituyendo en (1.27),

$$T \geq \frac{1}{44100} \frac{9}{(0,5)^2} \frac{0,1 \times 0,9}{0,087749^2} \text{ s} = 9,5 \text{ ms}. \quad (1.32)$$

Debe notarse que este tiempo de estabilización tan pequeño en la práctica no se verifica porque no se cumple la hipótesis señalada anteriormente de que la variación sea realmente aleatoria, es decir, que no haya correlación entre muestras sucesivas. En efecto, los valores de  $L_A$  se obtienen luego de filtrar y obtener la envolvente energética mediante la ecuación (1.3). La señal resultante es lentamente variable, por lo que, cuando se muestrea a una tasa de muestreo alta como la utilizada en el ejemplo anterior ( $F_s = 44100$  Hz), las muestras sucesivas están fuertemente correlacionadas. Conceptualmente, ello implica que si bien se toman numerosas muestras, muchas de ellas representan en realidad casi la misma información muestreada varias veces.

Si, por ejemplo, se utiliza una ponderación temporal rápida ( $\tau = 125$  ms), al muestrear la envolvente de un ruido blanco muy rara vez se obtiene una diferencia entre muestras sucesivas mayor que 0,02 dB. Una forma de tener en cuenta esto es reemplazar en (1.27) la tasa de muestreo real por una frecuencia mucho más baja, del orden de  $10f_c$ , donde  $f_c$  es la frecuencia de corte del filtro pasabajos. En el caso de la respuesta rápida es

$$f_c = \frac{1}{2\pi\tau_F} = \frac{1}{2\pi \times 0,125 \text{ s}} = 1,2732 \text{ Hz}.$$

En este caso, el tiempo de estabilización termina siendo de unos 33 s.

El ejemplo mostrado corresponde a un ruido blanco aleatorio. En el caso de ruidos como el del tránsito, los eventos individuales tienen una envolvente mucho más lenta, por lo cual no son auténticamente estacionarios sino considerando tiempos más prolongados. En este caso podemos reemplazar la frecuencia de muestreo por el flujo del tránsito  $Q$  en vehículos por unidad de tiempo. En el ejemplo de la sección 5 donde  $Q = 600$  veh/h, es decir  $Q = 1/6$  veh/h, resulta un tiempo de estabilización de 2525 s. Si el error  $\Delta L_{A90}$  se incrementa a 1 dB este tiempo se reduce a la cuarta parte, esto es, alrededor de 11 min.

Por último, aclaremos que cuando  $N$  se acerca a 0 o a 100, como ya se mencionó, si bien el numerador tiende a 0, la densidad de probabilidad  $f$  lo hace mucho más rápidamente, lo cual significa que el tiempo de estabilización requerido se vuelve muy

grande. Cuando finalmente  $N=0$  ó  $N=100$ , es decir, cuando se alcanzan respectivamente el máximo y el mínimo de la variable  $L_X$ , el tiempo de estabilización es infinito. Dicho de otra forma, ningún tiempo finito logra capturar realmente los valores mínimo y máximo, de allí que los valores mínimo y máximo de un periodo finito de medición sean muy poco relevantes.

### 1.7 Análisis de espectro

A menudo, el nivel de presión sonora, ya sea ponderado o no, es insuficiente para describir todos los aspectos relevantes de un problema. Por ejemplo, si se conoce sólo el nivel de ruido del lado externo de una fachada, no es posible determinar el nivel que se tendrá en el interior, ya que el aislamiento acústico de la fachada depende de la frecuencia, por lo cual un mismo nivel exterior puede dar origen a niveles muy variados según si el ruido es predominantemente de baja frecuencia o de alta frecuencia. Así, un tono puro de 100 Hz y 90 dB posee un nivel ponderado A de 71 dBA, al igual que un tono de 1600 Hz y 70 dB. Sin embargo, luego de atravesar una fachada típica, el primer tono se habrá atenuado unos 15 dB y el segundo unos 30 dB. El nivel sonoro interno será, por lo tanto, 75 dBA en el primer caso y 40 dBA en el segundo caso.

Otro ejemplo típico es el de identificación de fuentes sonoras. El mecanismo de generación de ruido de muchas fuentes involucra procesos periódicos (rotación, vibraciones, resonancias) con frecuencias definidas. Dichas frecuencias se conservan a lo largo del camino de propagación y hacen posible su detección en puntos remotos. Pero el nivel sonoro no permite discriminarlas.

En estos casos conviene realizar un *análisis de espectro*, es decir, una descomposición del sonido en las diversas frecuencias simultáneas que lo integran. Esto se logra con el *analizador de espectro*. Muchos modelos actuales de sonómetros admiten complementos opcionales de hardware o software que agregan funciones de analizador de espectro. Como veremos más adelante, si se dispone de un registro sonoro digital, es posible realizar el análisis de espectro en cualquier computadora.

Hay tres tipos de analizadores de espectro: de *ancho de banda constante*, de *porcentaje constante*, y de *líneas*. Los de ancho de banda constante poseen típicamente una única banda cuya frecuencia central y su ancho de banda pueden ajustarse en forma independiente. Se utilizan principalmente en radiofrecuencia y muy rara vez en audiofrecuencias, por lo que no los consideraremos aquí. Los de porcentaje constante poseen cierta cantidad de bandas que cubren el espectro completo de audio (20 Hz a 20 000 Hz) de tal manera que el ancho de banda de cada banda es una fracción o porcentaje de la frecuencia central. Expresado de otra manera, el cociente entre la frecuencia superior y la inferior es constante. Los analizadores de líneas presentan una gran cantidad de líneas espectrales (cientos o miles), cada una de las cuales representa un ancho de banda constante y relativamente estrecho.

#### 1.7.1 Analizador de espectro de porcentaje constante

Los analizadores de porcentaje constante se especifican generalmente como *fracción de octava*, por ejemplo, *analizador de octava* o *analizador de tercio de octava*. Llamando  $\alpha$  a la fracción de octava, resulta

$$f_{s,k} = 2^{\alpha/2} f_{o,k} \quad (1.33)$$

$$f_{i,k} = 2^{-\alpha/2} f_{o,k} \quad (1.34)$$

donde  $f_{o,k}$ ,  $f_{i,k}$ ,  $f_{s,k}$  son, respectivamente, las frecuencias central, inferior y superior de la  $k$ -ésima banda. También se cumple que

$$f_{s,k} = 2^\alpha f_{i,k} \quad (1.35)$$

y

$$f_{s,k} f_{i,k} = f_{o,k}^2. \quad (1.36)$$

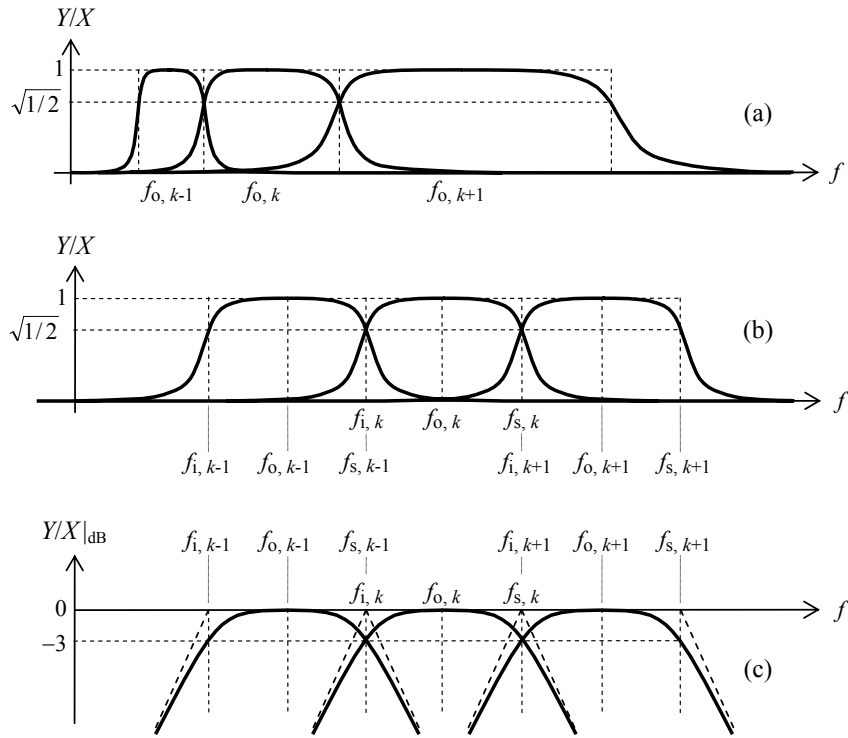
Además existe una condición de adyacencia de bandas por la cual la frecuencia inferior de una banda debe coincidir con la frecuencia superior de la banda anterior:

$$f_{i,k} = f_{s,k-1}. \quad (1.37)$$

Esto implica que

$$f_{o,k} = 2^\alpha f_{o,k-1}, \quad (1.38)$$

de modo que la relación entre frecuencias centrales de bandas sucesivas es la misma que entre las frecuencias superior e inferior de una misma banda. En la figura 10 se muestran tres bandas de octava en tres diferentes presentaciones.



**Figura 10.** Tres bandas consecutivas de un analizador de espectro de bandas de porcentaje constante, representadas: (a) con el eje de frecuencias en escala lineal; (b) con el eje de frecuencias en escala logarítmica; (c) con el eje de frecuencias en escala logarítmica y el de amplitudes en dB .

Se puede apreciar en la escala lineal de frecuencia, que las sucesivas bandas de octava van duplicando su ancho de banda. En una escala logarítmica, en cambio, todas parecen tener el mismo ancho. Una de las razones para elegir este tipo de bandas es que la percepción de la frecuencia por el aparato auditivo humano es prácticamente logarítmica en un amplio rango.

La fórmula (1.38) indica que si se adopta una frecuencia generatriz es posible obtener las sucesivas frecuencias centrales. Luego, a partir de éstas, las ecuaciones (1.33) y (1.34) permiten obtener las respectivas frecuencias inferior y superior de cada banda. Se ha adoptado 1000 Hz como frecuencia generatriz por estar aproximadamente en el centro del rango audible (en escala logarítmica).

Es habitual asociar a los filtros pasabanda un ancho de banda relativo  $B$  definido como el ancho de banda absoluto dividido por la frecuencia central:

$$B_k = \frac{f_{s,k} - f_{i,k}}{f_{o,k}} = 2^{\alpha/2} - 2^{-\alpha/2} . \quad (1.39)$$

Se observa que para una fracción de octava dada el ancho de banda relativo es constante. En la tabla 1.3 se dan los parámetros característicos para filtros de banda de porcentaje constante para varias fracciones de octava.

**Tabla 1.3.** Parámetros característicos de filtros pasabanda de diversas fracciones de octava.

$\alpha$	$f_s / f_i$	$B$
1	2	0,707
$1/2$	1,414	0,348
$1/3$	1,260	0,232
$1/6$	1,122	0,116
$1/10$	1,072	0,0693
$1/12$	1,059	0,0578
$1/24$	1,029	0,0289

Consideremos el caso de las bandas de tercio de octava (uno de los más frecuentes), donde  $\alpha = 1/3$ . Si se aplica (1.38) en forma estricta, las frecuencias exactas resultantes presentan valores con muchas cifras decimales y por lo tanto, difíciles de recordar. Al solo efecto de designar las bandas se ha adoptado una serie de frecuencias nominales que procura respetar en forma exacta las relaciones de década y lo más posible las de octava.<sup>11</sup> En la Tabla 1.4 se muestran los valores nominales y los valores exactos. El error relativo es menor que el 1,6 % en todos los casos.

Si bien durante mucho tiempo los analizadores de porcentaje constante se implementaron con bancos de filtros eléctricos analógicos,<sup>12</sup> todos los diseños actuales utilizan filtros digitales. Esto simplifica mucho el diseño y, sobre todo el ajuste de los fil-

<sup>11</sup> La serie de valores adoptada se conoce como serie R10 de Renard. Las series de Renard buscan distribuir uniformemente en escala logarítmica una cantidad de valores (en este caso, 10) en una década.

<sup>12</sup> Un banco de filtros es un conjunto de filtros que actúan simultáneamente sobre la misma señal de entrada, produciendo varias salidas.

tros, además de mejorar la estabilidad a largo plazo de las características de los filtros (por ejemplo frecuencia central o ancho de banda)

El análisis de espectro por bandas de porcentaje constante es muy útil cuando se aplica a señales de espectro continuo, es decir, sin componentes tonales marcadas, ya que permite evaluar el efecto producido por barreras, tabiques, absorción sonora, propagación a grandes distancias, etc. Las características de muchos materiales acústicos están especificadas por bandas de octava o de tercio de octava.

**Tabla 1.4.** Frecuencias centrales normalizadas de tercio de octava para utilizar en analizadores de espectro para el rango audible. Las señaladas con un asterisco (\*) también son frecuencias centrales de octava.

Frecuencia central nominal [Hz]	Frecuencia inferior [Hz]	Frecuencia central exacta [Hz]	Frecuencia superior [Hz]
20	17,54	19,69	22,10
25	22,10	24,803	27,84
* 31,5	27,84	31,25	35,08
40	35,08	39,37	44,19
50	44,19	49,61	55,68
* 63	55,68	62,50	70,15
80	70,15	78,75	88,39
100	88,39	99,21	111,36
* 125	111,36	125,00	140,31
160	140,31	157,49	176,78
200	176,78	198,43	222,72
* 250	222,72	250,00	280,62
315	280,62	314,98	353,55
400	353,55	396,85	445,45
* 500	445,45	500,00	561,23
630	561,23	629,96	707,11
800	707,11	793,70	890,90
* 1000	890,90	1000,00	1122,50
1250	1122,50	1259,90	1414,20
1600	1414,20	1587,40	1781,80
* 2000	1781,80	2000,00	2244,90
2500	2244,90	2519,80	2828,40
3150	2828,40	3174,80	3563,60
* 4000	3563,60	4000,00	4489,80
5000	4489,80	5039,70	5656,90
6300	5656,90	6349,60	7127,20
* 8000	7127,20	8000,00	8979,70
10000	8979,70	10079,00	11314,00
12500	11314,00	12699,00	14254,00
* 16000	14254,00	16000,00	17959,00
20000	17959,00	20159,00	22627,00

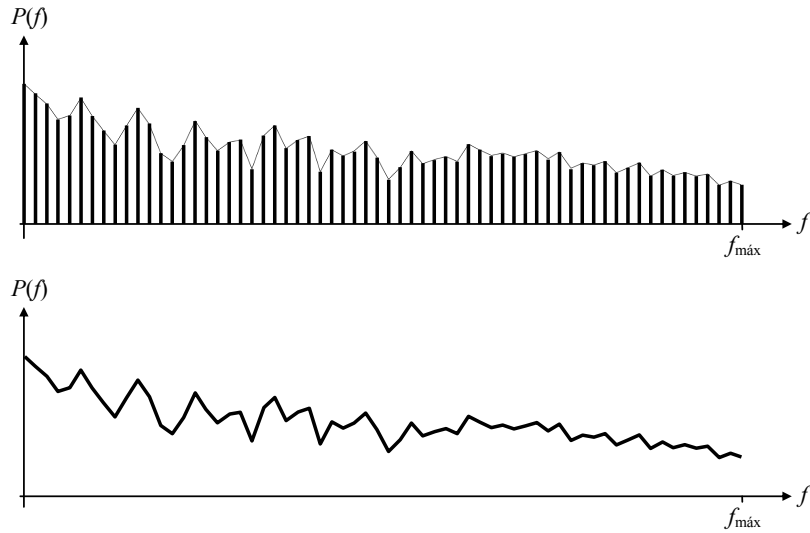


### 1.7.2 Analizador de espectro de líneas

Este tipo de analizadores se popularizó al introducirse el procesamiento digital de señales. Se basa en alguna variante de la Transformada Rápida de Fourier (Fast Fourier Transform), más conocida por sus siglas en inglés, FFT. Si bien luego nos referiremos con mayor detalle a este algoritmo, adelantaremos algunas de las propiedades básicas.

En general la cantidad  $N$  de líneas espectrales es una potencia entera de 2, por ejemplo, 2048, y están distribuidas uniformemente en un rango que va desde 0 (componente de continua)<sup>13</sup> hasta una frecuencia máxima  $f_{\text{máx}}$ , habitualmente un poco mayor que la máxima frecuencia audible, por ejemplo, 22 050 Hz ó 24 000 Hz (ver figura 11)<sup>14</sup>. Estas frecuencias son, por lo tanto,

$$f_k = k \frac{f_{\text{máx}}}{N}. \quad (1.40)$$



**Figura 11.** Arriba, un espectro de líneas (en este caso, 64 líneas). La altura de cada línea representa la energía contenida en una pequeña banda espectral. Abajo, una forma frecuente de representación mediante la envolvente de las líneas. Es útil cuando se utilizan muchas líneas, ya que las mismas no pueden discriminarse al estar tan próximas entre sí.

Esto implica que la resolución en frecuencia es igual a la frecuencia máxima dividida por el número de líneas:

$$\Delta f = \frac{f_{\text{máx}}}{N}. \quad (1.41)$$

<sup>13</sup> La componente de continua en Acústica debería ser siempre 0, ya que la presión sonora es la diferencia entre el valor instantáneo de la presión y la presión media o presión atmosférica. Sin embargo los dispositivos electrónicos agregan pequeños desplazamientos espurios de tensión (llamados *offsets*) que pueden aparecer como una componente de continua en el análisis de espectro.

<sup>14</sup> La máxima frecuencia está dada, como veremos al estudiar los rudimentos del audio digital, por la mitad de la tasa de muestreo.

En el ejemplo particular de una frecuencia máxima  $f_{\text{máx}} = 22\,050$  Hz y  $N = 2048$  líneas espectrales resulta:

$$\Delta f = \frac{22\,050 \text{ Hz}}{2048} = 10,77 \text{ Hz}.$$

Cada línea espectral representa, entonces, el contenido de energía de una banda angosta de frecuencias. Sin embargo, la interpretación es un poco más complicada debido al fenómeno de *derrame espectral*, por el cual una frecuencia única (un tono senoidal puro) repercute en varias o incluso en todas las líneas espectrales (si bien la influencia es menor en las líneas más alejadas). El problema del derrame espectral se puede minimizar mediante técnicas de *enventanado*, como se verá oportunamente.

El espectro de líneas es útil para detectar componentes tonales, así como para identificar y cuantificar la presencia de fuentes específicas. También permite realizar un análisis modal de diversos dispositivos y componentes. Por ejemplo, percutiendo adecuadamente la carcasa de una máquina pueden detectarse las frecuencias normales de vibración, es decir, aquellas que en condiciones de operación serán emitidas con mayor intensidad. También es posible superponer (energéticamente) varias líneas para obtener el nivel en una banda diferente de las bandas normalizadas de octava o tercio de octava.

### 1.7.3 Espectro de líneas e incertidumbre tiempo-frecuencia

Un hecho a tener en cuenta es que la obtención del espectro de líneas mediante la FFT requiere analizar un segmento de señal de duración

$$T = \frac{N}{f_{\text{máx}}}. \quad (1.42)$$

Esto implica que cuantas más líneas se requieran, es decir, cuanto más fina sea la resolución del análisis de espectro, mayor será la duración del segmento a analizar. Esto se conoce como *incertidumbre tiempo-frecuencia*. De hecho, combinando (1.41) y (1.42), resulta

$$T \Delta f = 1. \quad (1.43)$$

Esto se interpreta como que si queremos conocer el espectro con mucho detalle, perdemos información sobre su localización en el tiempo, mientras que si requerimos una buena localización en el tiempo debemos resignar resolución en el espectro. Así, si se requiriera una resolución de análisis de 10 Hz, sería preciso analizar un segmento de 0,1 s, pero si es suficiente una resolución de 100 Hz, bastará analizar sólo 0,01 s.

### 1.7.4 Espectro de líneas y densidad espectral

El espectro de líneas puede utilizarse también para estimar la *densidad espectral* de una señal aleatoria, definida como la presión eficaz cuadrática por unidad de ancho de banda. Con mayor precisión, es el límite de la presión eficaz cuadrática luego de atravesar un filtro pasabanda ideal de ancho de banda  $\Delta f$ , dividida por  $\Delta f$ :

$$\overline{p^2(f)} = \lim_{\Delta f \rightarrow 0} \frac{P_{f, \Delta f, \text{ef}}^2}{\Delta f}. \quad (1.44)$$

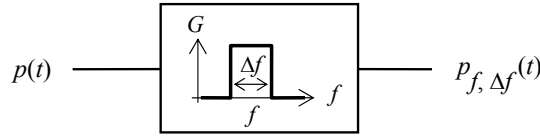
donde

$$P_{f,\Delta f,\text{ef}}^2 = \frac{1}{T} \int_0^T p_{f,\Delta f}^2(t) dt. \quad (1.45)$$

En la figura 12 se muestran esquemáticamente el filtro y su respuesta en frecuencia.

La densidad espectral representa cómo se distribuye la energía de la señal en el espectro de frecuencias. Su integral permite calcular el la presión eficaz en una banda dada:

$$P_{[f_1, f_2], \text{ef}}^2 = \int_{f_1}^{f_2} \overline{p^2(f)} df, \quad (1.46)$$



**Figura 12.** Una señal  $p(t)$  que atraviesa un filtro pasabanda de frecuencia central  $f$  y ancho de banda  $\Delta f$ .

o bien, tomando  $f_1 = 0$  y  $f_2 = \infty$ , el valor eficaz total de la presión sonora:

$$P_{\text{ef}}^2 = \int_0^\infty \overline{p^2(f)} df. \quad (1.47)$$

Para un ancho de banda pequeño  $\Delta f$  alrededor de una frecuencia  $f$  podemos aproximar:

$$\overline{p^2(f)} \cong \frac{P_{f,\Delta f,\text{ef}}^2}{\Delta f}. \quad (1.48)$$

Si elegimos un número de líneas espectrales  $N$  y llamamos  $P_{k,\text{ef}}$  a la presión eficaz de la  $k$ -ésima línea, la cual corresponde a la frecuencia

$$f_k = k \frac{f_{\text{máx}}}{N},$$

entonces podemos estimar la densidad espectral para la frecuencia  $f_k$  como

$$\overline{p^2(f_k)} \cong \frac{P_{k,\text{ef}}^2}{\Delta f} = \frac{P_{k,\text{ef}}^2}{f_{\text{máx}} / N}. \quad (1.49)$$

Muchas veces se desea obtener la densidad espectral media a largo plazo, ya que una señal aleatoria, por su mismo carácter aleatorio, puede variar permanentemente su

espectro. Por ejemplo, si nuestra señal es una larga conversación entre varias personas, el espectro a corto plazo variará no sólo a medida que intervienen personas diferentes con diferente timbre de voz, sino que para una misma persona el espectro variará según qué fonema esté pronunciando.

La ecuación (1.49) sólo proporciona información localizada (dentro de los márgenes compatibles con la indeterminación tiempo-frecuencia). Podemos estimar la densidad espectral media tomando  $M$  segmentos de duración  $T$  y promediando, para cada frecuencia, los valores de densidad espectral estimada para todos los segmentos:

$$\overline{p^2}(f_k) \cong \frac{\frac{1}{M} \sum_{m=1}^M P_{k,m,\text{ef}}^2}{f_{\text{máx}} / N}. \quad (1.50)$$

donde  $P_{k,m,\text{ef}}^2$  es la presión cuadrática correspondiente a la frecuencia  $f_k$  y al  $m$ -ésimo segmento temporal.

Es interesante determinar la cantidad  $M$  de segmentos a promediar para cumplir con una cota de error especificada. Más precisamente, buscamos  $M$  tal que el error se mantenga por debajo de dicha cota con una alta probabilidad, por ejemplo 95 %. Para ello necesitamos conocer la dispersión estadística de las líneas espectrales. Desde luego, dicha dispersión dependerá del tipo de señal. Así, si la señal es un tono puro, no habrá dispersión alguna, mientras que si se trata de un tono con una fuerte modulación de amplitud, el valor variará de acuerdo con la señal modulante.

A efectos de ejemplificar el tratamiento conceptual del problema, supondremos el caso de un ruido blanco aleatorio  $n(t)$  de banda acotada por  $f_{\text{máx}}$ . Podemos suponer que cada línea espectral evoluciona en el tiempo como la envolvente de la salida de un filtro pasabanda de ancho de banda igual a la resolución de análisis, dada por (1.41):

$$\Delta f = \frac{f_{\text{máx}}}{N}.$$

Según se analiza en el Apéndice 6, dicha envolvente tiene una distribución estadística de Rayleigh, cuyo valor medio es

$$\mu_Y = N_{\text{ef}} \sqrt{\frac{\pi \Delta f}{2 f_{\text{máx}}}} = N_{\text{ef}} \sqrt{\frac{\pi}{2 N}} \quad (1.51)$$

donde  $N_{\text{ef}}$  es el valor eficaz del ruido antes de pasar por el filtro. Este resultado indica el valor esperado para las diversas bandas que, desde luego, es constante por tratarse de un ruido blanco.

Asimismo, el desvío estándar es

$$\sigma_Y = N_{\text{ef}} \sqrt{\frac{4 - \pi}{2} \frac{\Delta f}{2 f_{\text{máx}}}} = N_{\text{ef}} \sqrt{\frac{4 - \pi}{4 N}}. \quad (1.52)$$

Este valor indica la dispersión estadística de cada banda. Si deseamos, por ejemplo, que el promedio de  $M$  segmentos tenga un error en dB menor que  $\Delta L$  con una probabilidad del 99,7 %, debemos elegir  $M$  tal que  $\Delta L$  corresponda a 3 desvíos estándar.

Llamando  $Y_{\text{med}}$  al resultado de promediar los valores de una misma línea espectral para  $M$  segmentos, nos interesa que se cumpla

$$\left| 20 \log \frac{Y_{\text{med}}}{\mu_Y} \right| < \Delta L. \quad (1.53)$$

Expresado en términos del logaritmo natural,

$$\left| \ln \frac{Y_{\text{med}}}{\mu_Y} \right| < \frac{\ln 10}{20} \Delta L.$$

Si  $\Delta L$  es pequeño, se puede aproximar el logaritmo por su aproximación lineal,

$$\left| \frac{Y_{\text{med}} - \mu_Y}{\mu_Y} \right| < \frac{\ln 10}{20} \Delta L,$$

es decir,

$$| Y_{\text{med}} - \mu_Y | < \frac{\ln 10}{20} \mu_Y \Delta L$$

Sabemos que para  $M$  relativamente grande  $Y_{\text{med}}$  está distribuida en forma aproximadamente normal, por lo que el 99,7 % de los valores cumplirán

$$| Y_{\text{med}} - \mu_Y | < 3 \sigma_{Y_{\text{med}}} = 3 \frac{\sigma_Y}{\sqrt{M}}.$$

Entonces bastará con que

$$\frac{\ln 10}{20} \mu_Y \Delta L = 3 \frac{\sigma_Y}{\sqrt{M}},$$

es decir

$$M = \left( \frac{60}{\ln 10} \frac{\sigma_Y}{\mu_Y} \right)^2 \frac{1}{\Delta L}. \quad (1.54)$$

Sustituyendo (1.51) y (2.16) en (1.54) resulta

$$M = \frac{4 - \pi}{2\pi} \left( \frac{60}{\ln 10} \right)^2 \frac{1}{\Delta L} \cong \frac{93}{\Delta L}. \quad (1.55)$$

Si, por ejemplo, deseamos que el error en la densidad espectral de un ruido blanco sea menor que 1 dB con una confiabilidad del 99,7 %, se deberán promediar 93 segmentos. Para  $f_{\text{máx}} = 22\,050$  Hz y  $N = 2048$ , ello implica un tiempo total de medición de  $M N / f_{\text{máx}} = 8,3$  s.

Otros tipos de ruido, por ejemplo el del tránsito, pueden requerir la promediación de muchos más segmentos, debido al carácter no estacionario, lo que obliga a tomar tiempos mucho más extensos.

### 1.7.5 Filtros de banda

Los filtros de bandas de octava y tercio de octava están normalizados por la norma IEC 61260, en la que se dan las plantillas que éstos deben cumplir. Las plantillas especifican, para cada frecuencia, los límites mínimo y máximo de la respuesta en frecuencia. Hay tres clases de filtros según cuán estricta sea la tolerancia. La clase 2 es la más permisiva (permite menores atenuaciones fuera de la banda de paso y mayores diferencias respecto a la respuesta ideal en la banda de paso), y la clase 0 la más rigurosa.

Dado que en escala logarítmica de frecuencia todos los filtros de una misma familia son iguales, basta especificar un solo filtro con frecuencia angular normalizada igual a 1. Las plantillas normalizadas para banda de octava se muestran en las figuras 13 y 14 y las correspondientes a banda de tercio de octava en las figuras 15 y 16.

La mayor tolerancia en frecuencias alejadas de la frecuencia central para las clases menos exigentes (1 y 2) en realidad tiene en cuenta la imposibilidad de verificación del cumplimiento debido al mayor ruido eléctrico circuital de la electrónica de bajo costo.

Dado que en el límite entre la banda de paso y la de corte una atenuación de 3 dB resulta aceptable (está entre los márgenes de tolerancia), es natural implementar el filtro mediante una aproximación de Butterworth.<sup>15</sup> Un filtro de Butterworth de orden 3 permite, teóricamente, satisfacer las plantillas correspondientes a todas las clases.<sup>16</sup> El cumplimiento real de cada clase dependerá de cuán ajustadas sean las tolerancias de los componentes utilizados.

Tradicionalmente esta implementación se ha realizado mediante filtros activos analógicos constituidos por tres celdas pasabanda de orden 1, cada una de las cuales se caracteriza por un polinomio cuadrático.<sup>17</sup> La función de transferencia normalizada de tales filtros está dada por las ecuaciones (1.56) para banda de octava y (1.57) para banda de tercio de octava.

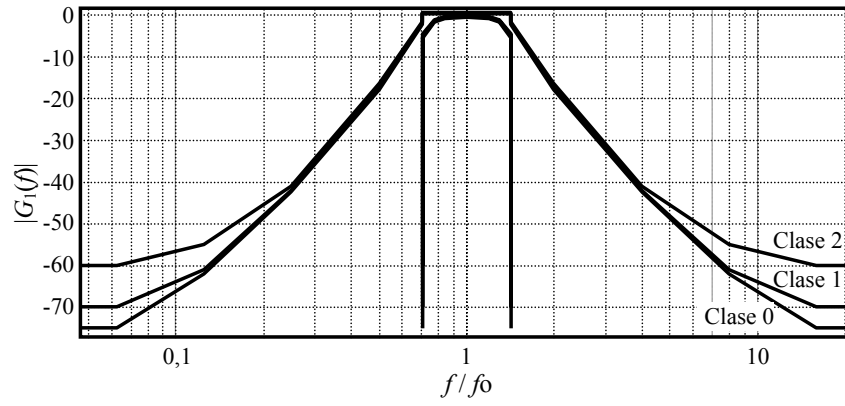
$$G_1(p) = \frac{0,35355 p^3}{(p^2 + 0,45844p + 1,8436)(p^2 + 0,70711p + 1)(p^2 + 0,24867p + 0,54242)} \quad (1.56)$$

$$G_3(p) = \frac{0,12417 p^3}{(p^2 + 0,12735p + 1,2221)(p^2 + 0,23156p + 1)(p^2 + 0,10421p + 0,81829)} \quad (1.57)$$

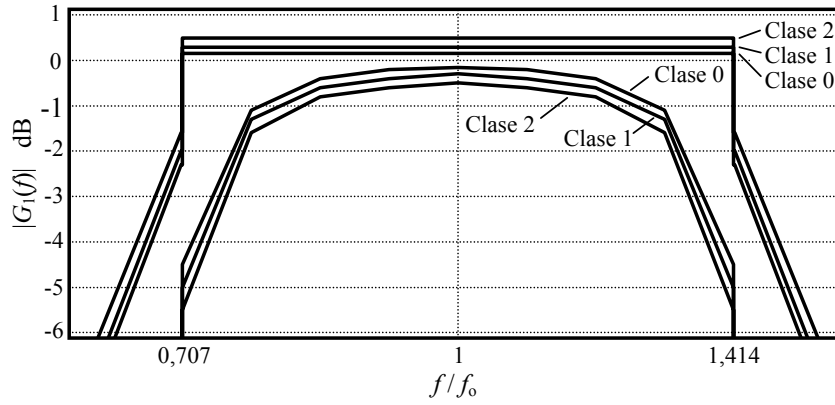
<sup>15</sup> Existen diversas aproximaciones clásicas para la realización de filtros que cumplan con una determinada plantilla, por ejemplo Butterworth, Chebychev o Cauer. Cada una presenta ventajas y desventajas. Por ejemplo la de Chebychev logra la máxima pendiente en la banda de atenuación compatible con un orden dado para un filtro polinomial, pero lo hace a costa de un rizado en la banda de paso y un comportamiento transitorio bastante oscilante. La de Butterworth es máximamente plana en la banda de paso y tiene transitorios aceptables, aunque su respuesta no cae tan rápido como la de Chebychev (Miyara, 2004)

<sup>16</sup> El orden  $n$  de un filtro se relaciona con la rapidez de caída de su respuesta frecuencial en la banda de corte. Concretamente, la pendiente es de  $\pm 20 n$  dB/déc.

<sup>17</sup> Como sistema dinámico estas celdas son de orden 2, de allí que correspondan a una función de transferencia con dos polos, es decir, dos ceros en el denominador. Se lo considera un filtro de orden 1 pues la caída a ambos lados de la frecuencia central es con una pendiente de  $\pm 20$  dB/déc.



**Figura 13.** Plantillas para filtros de banda de octava según la norma IEC 61260. Vista global.



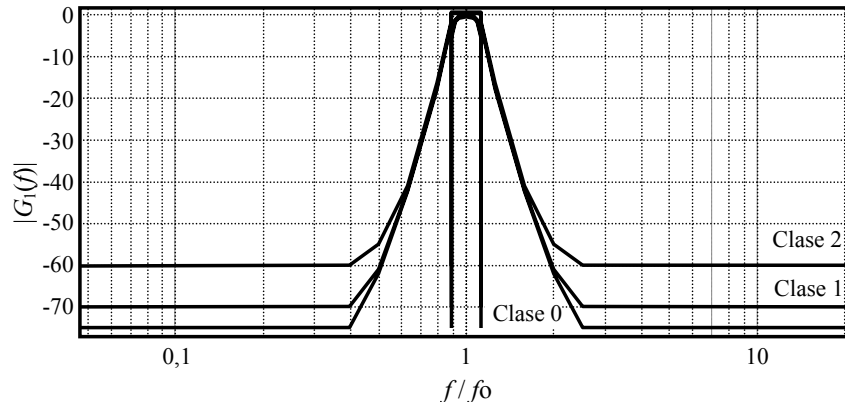
**Figura 14.** Plantillas para filtros de banda de octava según la norma IEC 61260. Detalle de la banda de paso.

Aquí se utiliza la variable de Laplace normalizada  $p$  en lugar de la variable  $s$  pues el filtro está normalizado para tener su frecuencia central en  $\Omega = 1$ .

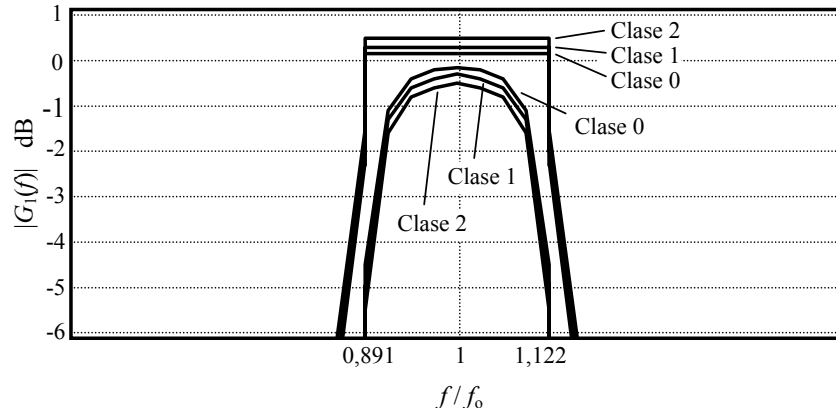
La función de cada celda puede expresarse mediante dos parámetros característicos: la frecuencia angular natural  $\omega_0$  y el factor de calidad  $Q$ . Para una celda genérica,

$$H(s) = \frac{\frac{\omega_0}{Q}s}{s^2 + \frac{\omega_0}{Q}s + \omega_0^2} \quad (1.58)$$

En el caso de funciones normalizadas,



**Figura 15.** Plantillas para filtros de banda de tercio de octava según la norma IEC 61260. Vista global.



**Figura 16.** Plantillas para filtros de banda de tercio de octava según la norma IEC 61260. Detalle de la banda de paso.

$$H(p) = \frac{\frac{\Omega_o}{Q} p}{p^2 + \frac{\Omega_o}{Q} p + \Omega_o^2} \quad (1.59)$$

En las tablas 1.5 y 1.6 se indican los parámetros característicos, es decir, la frecuencia angular normalizada,  $\Omega_o$ , y el factor de calidad  $Q$  para filtros de banda de octava y de tercio de octava respectivamente.

Finalmente, en las figuras 17 y 18 se muestra la respuesta en frecuencia de los filtros de octava y de tercio de octava, junto con las correspondientes respuestas de las tres celdas con que se implementan.

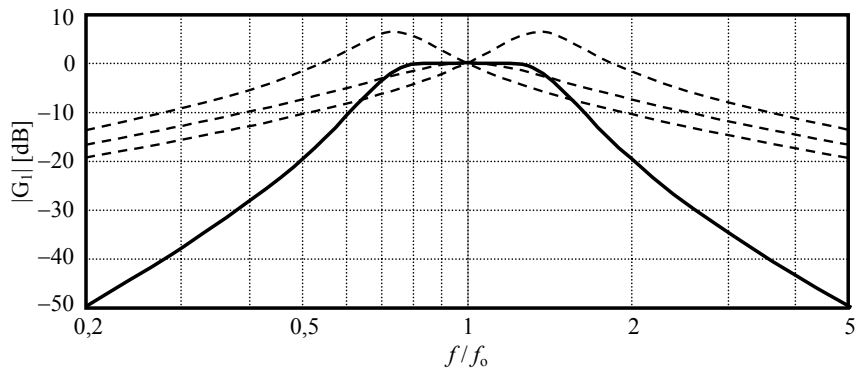


**Tabla 1.5.** Parámetros característicos de las celdas pasabanda normalizadas (frecuencia angular natural normalizada  $\Omega_{ok}$ , y factor de calidad  $Q_k$ ) para filtros de banda de octava.

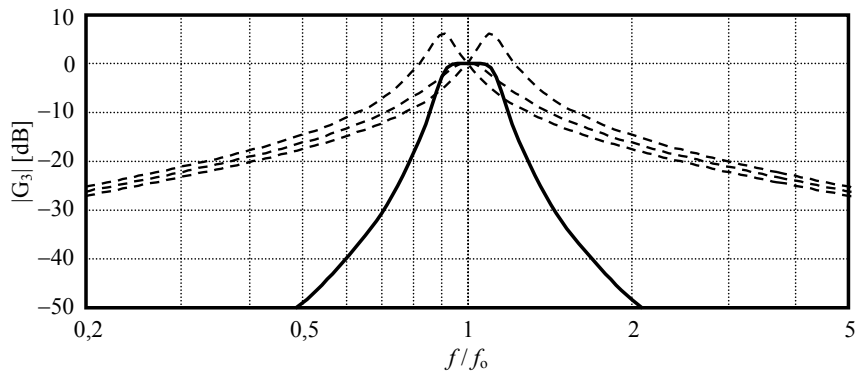
Celda	$\Omega_{ok}$	$Q_k$
1	1,357783939141098	2,961756551889217
2	1	1,414213562373096
3	0,736494202923462	2,961756551889228

**Tabla 1.6.** Parámetros característicos de las celdas pasabanda normalizadas (frecuencia angular natural normalizada  $\Omega_{ok}$ , y factor de calidad  $Q_k$ ) para filtros de banda de tercio de octava

Celda	$\Omega_{ok}$	$Q_k$
1	1,105467721309820	8,680399444876773
2	1	4,318473046963131
3	0,904594481343287	8,680399444876974



**Figura 17.** Respuesta en frecuencia normalizada de un filtro de banda de octava y sus factores de 2º orden (en líneas de trazos).



**Figura 18.** Respuesta en frecuencia normalizada de un filtro de banda de tercio de octava y sus factores de 2º orden (en líneas de trazos).

La respuesta en frecuencia real para una determinada banda se obtiene sustituyendo la variable normalizada  $p$  por

$$p = j \frac{f}{f_0}. \quad (1.60)$$

La función de transferencia de Laplace, por otra parte, se obtiene sustituyendo

$$p = \frac{s}{2\pi f_0}. \quad (1.61)$$

### 1.7.6 Respuesta transitoria de los filtros de banda

La respuesta en frecuencia de un filtro indica en qué medida se atenúa una señal senoidal al atravesar el filtro según su frecuencia. Sin embargo, dicho comportamiento se obtiene recién después de un tiempo, es decir, luego de extinguido el *régimen transitorio* y alcanzado el *régimen permanente*. En efecto, la respuesta temporal está dada por

$$r(t) = A_0 \sin(\omega t + \varphi) + \sum_{k=1}^3 A_k e^{-t/\tau_k} \sin(\omega_k t + \psi_k) \quad (1.62)$$

donde  $\omega_k$  es la frecuencia angular y  $\tau_k$  la constante de tiempo de decaimiento del modo  $k$  (ver Apéndice 7). El primer término es la *respuesta forzada* o régimen permanente y los restantes términos constituyen la *respuesta libre* o régimen transitorio. En las figuras 19 y 20 se ilustra la respuesta completa y sus componentes

Como se puede apreciar, la respuesta del filtro alcanzará una amplitud esencialmente igual a la del régimen permanente tanto más rápido cuanto menores sean las constantes de tiempo  $\tau_k$ . No es posible, sin embargo, reducir dichas constantes de tiempo más allá de un límite que depende del ancho de banda del filtro.

Para verlo tengamos en cuenta la relación (A7.8) del apéndice 7:

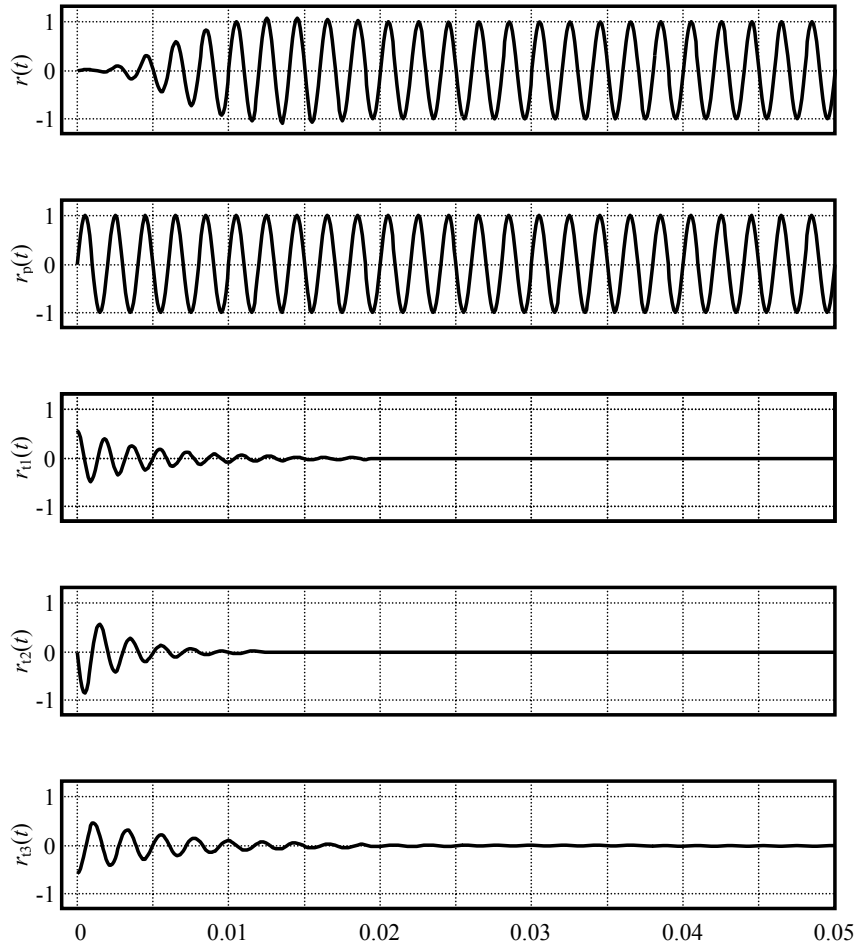
$$\tau_k = \frac{2Q_k}{\omega_{0k}} \quad (1.63)$$

Expresado en términos de la frecuencia normalizada,

$$\tau_k = \frac{2Q_k}{\omega_0 \Omega_{0k}}. \quad (1.64)$$

Podemos observar que a menor frecuencia central  $\omega_0$ , mayores serán las constantes de tiempo y, por lo tanto, mayor será el tiempo requerido para que el transitorio se extinga. Así, un filtro de banda centrado en 31,5 Hz tendrá una respuesta mucho más lenta que uno centrado en 16 kHz.

Asimismo, cuanto más estrecha sea la banda de paso más altos serán los valores de  $Q$ . Si comparamos las tablas 1.5 y 1.6, vemos que para banda de octava  $Q_{\text{máx}} = 2,96$  mientras que para banda de tercio de octava  $Q_{\text{máx}} = 8,68$ . Por consiguiente, a igual frecuencia central, el de tercio de octava será casi 3 veces más lento que el de octava.

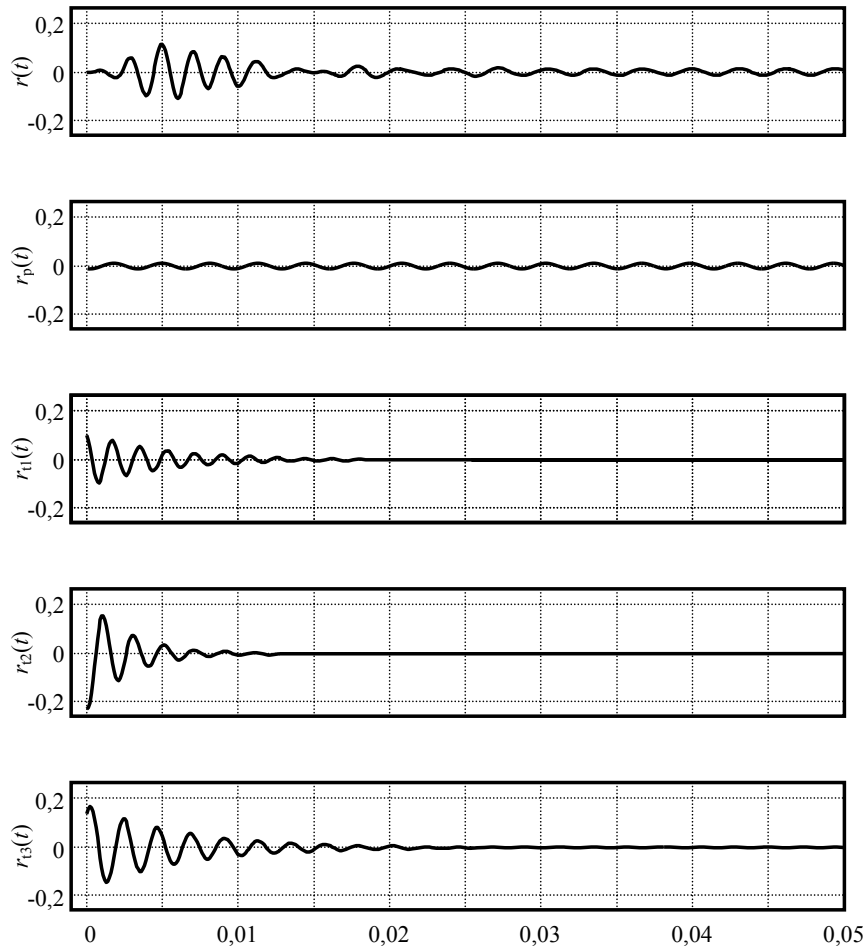


**Figura 19.** Arriba, respuesta  $r(t)$  de un filtro de banda de tercio de octava de Butterworth de tercer orden de  $f_0 = 500$  Hz ante un tono senoidal de 500 Hz. Más abajo, la respuesta en régimen permanente,  $r_p(t)$  y las tres componentes de respuesta transitoria  $r_{t1}(t)$ ,  $r_{t2}(t)$  y  $r_{t3}(t)$ .

Esta limitación es una nueva instancia de la *incertidumbre tiempo-frecuencia*, ya vista anteriormente en relación con el espectro de líneas en la ecuación (1.43). Conceptualmente, cuanto más estrecha sea la banda de análisis, mayor tiempo se requerirá para obtener una respuesta estabilizada (esencial para realizar una medición aceptable).<sup>18</sup>

Un análisis teórico de esto debería fijar un margen de tolerancia expresado en dB con respecto al valor final o de régimen permanente, y estudiar el tiempo requerido para que la señal de salida del filtro ingrese y permanezca en la banda de tolerancia estipulada, como se muestra en la figura 21.

<sup>18</sup> Notemos que para bandas de porcentaje constante, cuanto menor sea la frecuencia central menor será el ancho de banda. Es por esto que las bandas de baja frecuencia requieren mayor tiempo de estabilización



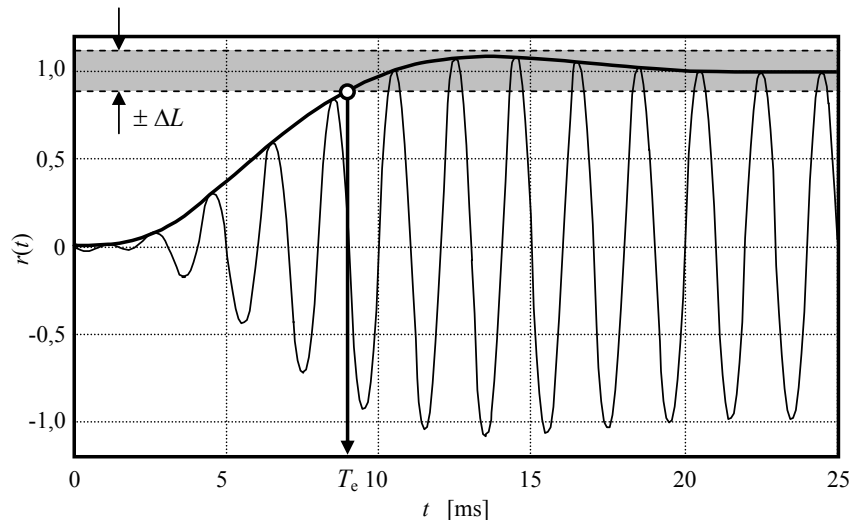
**Figura 20.** Arriba, respuesta  $r(t)$  de un filtro de banda de tercio de octava de Butterworth de tercer orden de  $f_o = 500$  Hz ante un tono senoidal de 315 Hz. Más abajo, la respuesta en régimen permanente,  $r_p(t)$  y las tres componentes de respuesta transitoria  $r_{t1}(t)$ ,  $r_{t2}(t)$  y  $r_{t3}(t)$ .

El análisis se dificulta, sin embargo, debido a la relación complicada que existe entre la amplitud del régimen permanente y las amplitudes iniciales de las componentes del régimen transitorio. Como se ve en las figuras 19 y 20, para una frecuencia de entrada igual a la frecuencia central del filtro la amplitud de los transitorios es menor que la del régimen permanente, pero en el caso de una frecuencia bastante menor, es bastante mayor.

Resulta más operativo simular por software la respuesta para diversas frecuencias y determinar en cada caso el tiempo de estabilización con una tolerancia dada. Los resultados se muestran en las figuras 22 y 23. En la primera se supuso una tolerancia de  $\pm 0,15$  dB, correspondiente a filtros clase 0 en la frecuencia central y en la segunda,  $\pm 0,5$  dB, correspondiente a filtros clase 2.

Tal como era previsible, el tiempo de estabilización es más corto para filtros de banda de octava que para los de banda de tercio de octava. También se observa que para

baja frecuencia el tiempo de estabilización aumenta, pero ello no se debe sólo a la ecuación (1.64) sino también al hecho de que la amplitud final es mucho menor, por lo tanto el transitorio debe caer hasta un valor mucho menor para que en términos relativos a la amplitud final signifique una tolerancia como la especificada en cada caso. Esto explica también el comportamiento en alta frecuencia, donde la amplitud del régimen permanente es, igualmente, muy pequeña.



**Figura 21.** Tiempo de estabilización para  $\Delta L = 1$  dB de un filtro de banda de tercio de octava de Butterworth de orden 3 centrado en 500 Hz. El pequeño círculo muestra el punto en que la envolvente ingresa definitivamente en la banda de tolerancia.

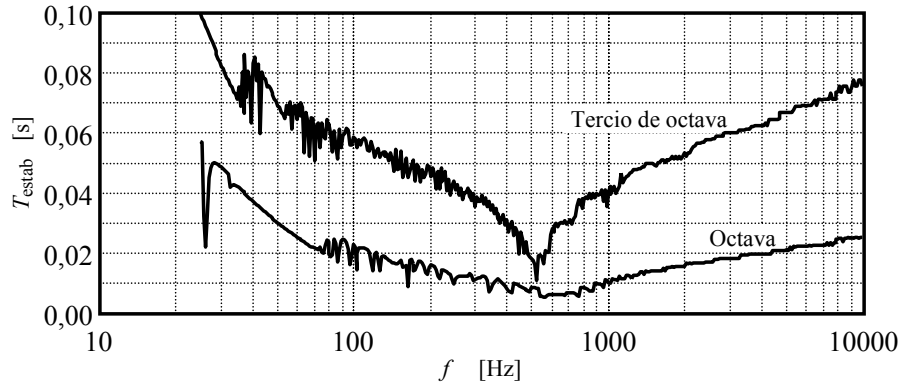
Es interesante constatar el comportamiento irregular del tiempo de estabilización con la frecuencia, si bien se observan las tendencias generales ya comentadas. El motivo es que pequeñas diferencias de frecuencia pueden hacer que el ingreso definitivo en la banda de tolerancia salte de un ciclo al siguiente por un ligero cambio de amplitud del régimen permanente o algún transitorio.

Si bien las gráficas de las figuras 22 y 23 se construyeron para un filtro de frecuencia central 500 Hz, las mismas se pueden aplicar a cualquier otra frecuencia central desplazándolas horizontalmente y luego aplicando una corrección por frecuencia central como la dada en la ecuación (1.64).

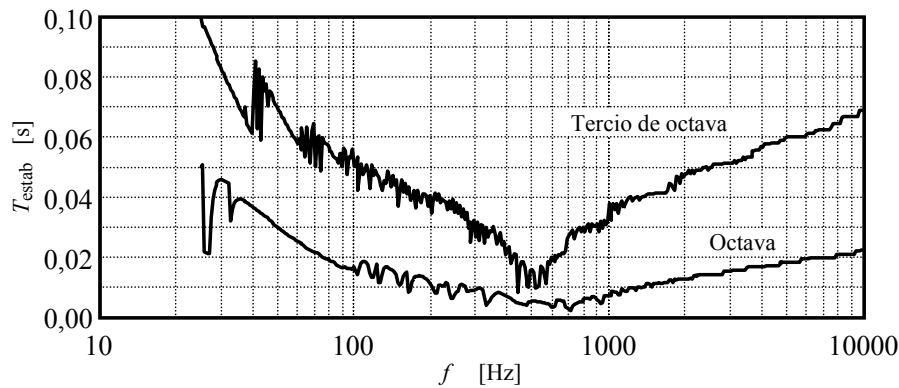
NOTA 1: Cuando la tolerancia  $\Delta L$  es muy pequeña, puede ocurrir que la envolvente de la señal ingrese en la banda de tolerancia y vuelva a salir dos o más veces antes de ingresar definitivamente.

NOTA 2: Podría redefinirse el tiempo de estabilización tomando una tolerancia variable con la frecuencia, teniendo en cuenta las plantillas de las figuras 13 a 16, con lo cual el tiempo de estabilización se reduciría aún más.

NOTA 3: Salvo para casos muy particulares como la medición de tonos puros en laboratorio, el tiempo de estabilización fuera de la banda de paso tiene escasa importancia dado que la respuesta para otras bandas cae rápidamente por debajo del ruido ambiente en la correspondiente banda.



**Figura 22.** Tiempo de estabilización de la respuesta de los filtros de banda de octava y de tercio de octava de Butterworth de orden 3, centrados en 500 Hz, para tonos de diversas frecuencias, bajo una tolerancia de 0,15 dB (corresponde a la tolerancia admisible para clase 0 en el centro de la banda).



**Figura 23.** Tiempo de estabilización de la respuesta de los filtros de banda de octava y de tercio de octava de Butterworth de orden 3, centrados en 500 Hz, para tonos de diversas frecuencias, bajo una tolerancia de 0,5 dB (corresponde a la tolerancia admisible para clase 2 en el centro de la banda).

### 1.7.7 Igualdad de Parseval

Consideremos una señal periódica expresada como una serie trigonométrica<sup>19</sup>

$$p(t) = \sum_{n=1}^{\infty} \sqrt{2}P_{n,\text{ef}} \sin 2\pi n f, \quad (1.65)$$

<sup>19</sup> Hemos omitido el valor medio o valor de continua dado que en el caso de señales acústicas carece en general de interés, salvo en algunos casos muy particulares como la impedanciometría del tímpano, donde se aplica una presión estática positiva o negativa para provocar una deformación de la membrana timpánica.

donde  $P_{n,\text{ef}}$  es el valor eficaz del armónico  $n$ . Se demuestra que el valor eficaz total al cuadrado satisface la siguiente ecuación, denominada *igualdad de Parseval*:

$$P_{\text{ef}}^2 = \sum_{n=1}^{\infty} P_{n,\text{ef}}^2. \quad (1.66)$$

Esta ecuación indica que las energías de los diferentes armónicos se superponen linealmente (a pesar de que la elevación al cuadrado implícita en la definición del valor eficaz sea un proceso inherentemente no lineal). En otras palabras, las energías o potencias de los diferentes armónicos son independientes y no interfieren entre sí. Obsérvese que esto no es cierto para señales que tienen la misma frecuencia. Por ejemplo, la potencia de la superposición de dos tonos de igual amplitud y frecuencia no es el doble de la de cada uno de ellos sino que depende de la fase relativa entre ellos. Así, en el caso extremo en que estén en fase será el cuádruple, y en el caso opuesto, cuando están en contra-fase, será nula.

La igualdad de Parseval tiene varias generalizaciones. La primera es a través de la transformada de Fourier,

$$P(\omega) = \int_{-\infty}^{\infty} p(t) e^{-j\omega t} dt. \quad (1.67)$$

En este caso se cumple la igualdad de Plancherel, análoga a la de Parseval,

$$\int_{-\infty}^{\infty} p^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |P(\omega)|^2 d\omega, \quad (1.68)$$

que también puede expresarse como

$$\int_{-\infty}^{\infty} p^2(t) dt = \int_{-\infty}^{\infty} |P(2\pi f)|^2 df. \quad (1.69)$$

Esta ecuación es válida para señales de cuadrado integrable entre  $-\infty$  e  $\infty$ , es decir, aquellas con energía total finita, por ejemplo señales de corta duración, como pulsos, ráfagas tonales, eventos individuales, etc. Obsérvese que aquí el valor eficaz a largo plazo es 0, por lo que el concepto de potencia media carece de interés práctico.

Para señales estacionarias o a largo plazo cuya energía total es infinita, la igualdad de Parseval adopta la forma ya vista en (1.47),

$$P_{\text{ef}}^2 = \int_0^{\infty} \overline{p^2(f)} df, \quad (1.70)$$

donde  $\overline{p^2(f)}$  es la densidad espectral. En el caso particular de un tono puro la densidad espectral, es una delta de Dirac.

### 1.7.8 Efecto de la tolerancia del espectro en la igualdad de Parseval

Supongamos una señal cuyo espectro está comprendido entre  $f_{\min}$  y  $f_{\max}$ . Si separamos el espectro en una serie de bandas  $[f_{ik}, f_{sk}]$ , donde  $f_{i1} = f_{\min}$  y  $f_{iN} = f_{\max}$ , y aplicamos la ecuación (1.46) junto con las propiedades de la integral, resultará

$$P_{\text{ef}}^2 = \sum_{k=1}^N \int_{f_{ik}}^{f_{sk}} \overline{p^2(f)} df, \quad (1.71)$$

Es decir, la potencia de las bandas disjuntas es aditiva. Podríamos aplicar esto al espectro de bandas de la señal de donde se concluye que el espectro debería cumplir con la igualdad de Parseval.

Sin embargo, la validez de ello está dada bajo el supuesto de que las bandas sean realmente disjuntas entre sí, es decir que ninguna banda acusa contenido espectral proveniente de otras bandas. Si observamos la respuesta en frecuencia de las figuras 17 y 18, o las plantillas de las figuras 13 a 16 veremos que un tono ubicado fuera de una banda dada es capaz de provocar una salida no nula en dicha banda. Este fenómeno se denomina genéricamente *derrame espectral* y constituye otra limitación del análisis de espectro real.

Una consecuencia es provocar errores en el cumplimiento de la igualdad de Parseval. Supongamos, por ejemplo, un tono de frecuencia  $f_0 = 1$  kHz que ingresa a un analizador de espectro por bandas de tercio de octava formado por un banco de filtros de Butterworth de tercer orden cuya respuesta en la frecuencia central es ideal (es decir, es unitaria). La sumatoria de las potencias de las respuestas de las 31 bandas es

$$\sum_{k=1}^{31} P_{k,\text{ef}}^2 = P_{\text{ef}}^2 \sum_{k=1}^{31} G_k^2(f_0). \quad (1.72)$$

Dado que los filtros tienen las mismas respuestas desplazadas de a un tercio de octava, se cumple que

$$G_{k-1}(f) = G_k(2^{1/3} f), \quad (1.73)$$

de donde

$$\sum_{k=1}^{31} P_{k,\text{ef}}^2 = P_{\text{ef}}^2 \sum_{k=1}^{31} G_{18}^2 \left( 2^{\frac{k-18}{3}} f_0 \right). \quad (1.74)$$

El exponente de 2 se eligió para abarcar las frecuencias desde 20 Hz a 20 kHz. El cálculo, realizado computacionalmente en base a la respuesta en frecuencia teórica, da

$$\sum_{k=1}^{31} P_{k,\text{ef}}^2 = 1,030013 P_{\text{ef}}^2. \quad (1.75)$$



Esto implica un error por exceso de 0,13 dB en el cumplimiento de la igualdad de Parseval.

Para el caso de bandas de octava,

$$\sum_{k=1}^{10} P_{k,\text{ef}}^2 = P_{\text{ef}}^2 \sum_{k=1}^{10} G_{18}^2 (2^{k-6} f_0). \quad (1.76)$$

El cálculo da

$$\sum_{k=1}^{10} P_{k,\text{ef}}^2 = 1,021801 P_{\text{ef}}^2. \quad (1.77)$$

que corresponde a un error por exceso de 0,09 dB en el cumplimiento de Parseval.

Supongamos, ahora, el caso de un ruido blanco entre  $f_{\min} = 17,8$  Hz y  $f_{\max} = 22,4$  kHz, es decir entre los límites ideales de la banda de tercio de octava de 20 Hz y 20 kHz. Si la presión eficaz es  $P_{\text{ef}}$ , entonces la densidad espectral, que es constante por tratarse de ruido blanco, puede calcularse como:

$$\overline{p_o^2} = \frac{P_{\text{ef}}^2}{f_{\max} - f_{\min}}. \quad (1.78)$$

El valor eficaz cuadrático de la porción del espectro contenido idealmente en la  $k$ -ésima banda es

$$P_{k,\text{ef,ideal}}^2 = \int_{f_{i,k}}^{f_{s,k}} \overline{p_o^2} df = \overline{p_o^2} (f_{s,k} - f_{i,k}) \quad (1.79)$$

es decir,

$$P_{k,\text{ef,ideal}}^2 = \frac{f_{s,k} - f_{i,k}}{f_{\max} - f_{\min}} P_{\text{ef}}^2 \quad (1.80)$$

Estos valores ideales cumplen con la igualdad de Parseval puesto que el extremo superior de cada banda coincide con el extremo inferior de la siguiente, por lo cual

$$\sum_{k=1}^N (f_{s,k} - f_{i,k}) = f_{\max} - f_{\min}.$$

El valor realmente obtenido incluye la ganancia del filtro, que implica una leve reducción hacia los bordes de la banda pero también un incremento debido al derrame espectral:

$$P_{k,\text{ef}}^2 = \int_{f_{\min}}^{f_{\max}} \overline{p_o^2} |G_k(f)|^2 df = \overline{p_o^2} \int_{f_{\min}}^{f_{\max}} |G_k(f)|^2 df \quad (1.81)$$

Por ejemplo, suponiendo  $P_{\text{ef}} = 1$  Pa, el resultado ideal para la banda de tercio de octava de 1 kHz es 0,10160 Pa, mientras que el resultado real es 0,10397 Pa. Ello representa un error por exceso de 0,2 dB, de donde se concluye que el derrame espectral predomina por sobre la reducción de ganancia hacia los bordes del filtro.

Si se repite el cálculo para todas las bandas y se suman los valores eficaces cuadráticos se obtiene un valor eficaz total estimado  $P_{\text{ef,estimado}} = 1,0193$  Pa. Esto significa que la aplicación de la fórmula de Parseval implica un error por exceso de 0,17 dB.

Podemos concluir, entonces, que debido a fenómeno de derrame espectral, la fórmula de Parseval produce, tanto en el caso de señales espectralmente muy estrechas (tono puro) como en las de banda ancha (ruido blanco), errores de hasta 0,17 dB. Esto constituye un error de coherencia interna de los datos, pero es un error generalmente admisible si se tienen en cuenta las diversas fuentes de incertidumbre en las mediciones acústicas.

## Capítulo 2

# INCERTIDUMBRE

### 2.1 Introducción

Toda medición tiene el propósito de obtener el valor numérico de una cantidad física. Sin embargo, este enunciado en apariencia tan claro y obvio requiere una cuidadosa inspección. En primer lugar la cantidad a medir, denominada *mensurando*, no siempre está completamente definida. Para que la definición fuera exhaustiva deberían especificarse todas las condiciones que podrían influir en el valor del mensurando. En el caso acústico, por ejemplo:

- 1) Condiciones atmosféricas (presión atmosférica, temperatura, humedad, velocidad y dirección del viento, altitud, presencia de material particulado),
- 2) Geometría y materialidad del entorno (presencia de barreras acústicas, superficies absorbentes o reflectoras, etc.),
- 3) Condiciones operativas de la fuente (por ejemplo temperatura, condiciones de carga, tensión de alimentación),
- 4) Oportunidad (horario, tipo de proceso que se lleva a cabo, tiempo transcurrido desde el comienzo del proceso y desde el último mantenimiento),
- 5) Presencia de otras fuentes simultáneas, etc.,
- 6) Ubicación del instrumento de medición y del operador.<sup>20</sup>

En segundo lugar, aun si se pudieran especificar detalladamente estas condiciones, sería imposible garantizar que se cumplan durante la medición, entre otras cosas, por la imposibilidad de conocer los valores exactos de los parámetros asociados a dichas condiciones. Esto significa que en una situación real no hay un único valor posible del mensurando.

A pesar de ello, podemos concebir la existencia de un *valor verdadero*, es decir, el valor que tendría el mensurando si se midiera en condiciones ideales con un instrumento ideal. Por lo dicho anteriormente, no resulta posible determinar exactamente el valor verdadero. Sólo se puede aspirar a obtener una estimación junto con alguna expresión de su *incertidumbre*, es decir una forma de cuantificar cuán lejos puede estar el valor verdadero de la estimación.

La forma de expresar la incertidumbre ha variado a lo largo del tiempo. Inicialmente se suponía que era posible dar un intervalo alrededor del valor estimado dentro del cual se encuentra el valor verdadero. En la actualidad, y desde la publicación por la ISO de la *Guía para la expresión de la incertidumbre en mediciones*, conocida como GUM, por sus siglas en inglés (ISO-GUM, 1993; JCGM, 2008), el paradigma se ha volcado hacia una expresión estadística, especificando la incertidumbre como un intervalo dentro del cual se encuentra el valor verdadero con una determinada probabilidad, por ejemplo, 95 %.

---

<sup>20</sup> El propio instrumento de medición introduce fenómenos de reflexión y difracción que alteran el campo sonoro y, por lo tanto, modifica el mensurando. Lo mismo cabe observar con respecto al operador que realiza la medición

## 2.2 Resolución, precisión, exactitud

Todo instrumento de medición tiene un límite para la mínima diferencia entre dos lecturas no idénticas denominado *resolución* o *apreciación*.<sup>21</sup> En un instrumento con indicación analógica (cada vez menos frecuentes), está dado por la distancia entre dos divisiones sucesivas de la escala. En un instrumento digital, corresponde a una unidad del dígito menos significativo. En la mayoría de los sonómetros la resolución es de 0,1 dB, y en algunos es posible configurarla en 0,01 dB (como veremos, esto puede tener sentido para mediciones de nivel equivalente).

La resolución de un procedimiento de medición podría ser mejor que la del instrumento utilizado, por ejemplo, cuando el valor medido se obtiene como promedio de una cantidad de mediciones. Así, si el procedimiento de medición consiste en promediar 10 mediciones con un instrumento de resolución 0,1 dB, tal procedimiento tendrá una resolución de 0,01 dB.

Cuando se realizan varias mediciones de un mismo mensurando en idénticas condiciones<sup>22</sup> (es decir, se miden varias *realizaciones* del mensurando) no siempre se obtienen lecturas idénticas, sino que las mismas presentan cierta dispersión estadística. Esto implica que es más correcto caracterizar la lectura mediante una variable aleatoria que mediante un valor único en función del valor verdadero.<sup>23</sup> La *precisión* del instrumento se define como la dispersión de las lecturas, y puede expresarse como el desvío estándar  $\sigma$  de las lecturas o como la semilongitud de un intervalo de confianza dentro del cual se encuentra un porcentaje especificado de las lecturas (típicamente, un 95 %; para una distribución normal corresponde aproximadamente a  $2\sigma$ ).

Un concepto más controversial es el de *exactitud*, definido como el grado de proximidad entre cualquier lectura obtenida y el valor verdadero del mensurando. Según la GUM, es un concepto cualitativo, y no cabría asignarle un valor numérico, ya que el valor verdadero no es conocible. Sin embargo, la propia GUM y otras normas y documentos metrológicos utilizan frecuentemente el término *exactitud* de un modo característico de las especificaciones cuantitativas. Por ejemplo, en la subcláusula 3.1.3 expresa:

“En la práctica, la especificación o definición del mensurando está dictada por la exactitud de la medición requerida.”

Cabe preguntarse de qué manera podría especificarse la exactitud requerida que no sea en forma numérica. Una expresión no cuantitativa de la exactitud requerida no permitiría verificar si se ha alcanzado o no.

Pese a lo que indica la GUM, sería posible caracterizar cuantitativamente la exactitud como una dispersión, ya sea a través del desvío estándar de la diferencia entre el valor medido y el valor verdadero o de un intervalo de confianza para dicha diferencia.

<sup>21</sup> El término *apreciación* es más antiguo y era común en la instrumentación analógica, dando cuenta del hecho de que el operador humano interviene activamente en la toma visual de la lectura. En mediciones digitales hasta podría no intervenir un operador humano, tomándose las decisiones que motive la lectura en forma automática, por software.

<sup>22</sup> *Idénticas condiciones* significa controlar todo aquello que sea posible, por ejemplo, el mismo sistema físico, el mismo operador, el mismo procedimiento de medición, las mismas condiciones ambientales y el mínimo transcurso de tiempo entre una medición y la siguiente para evitar derivas en las características del elemento o sistema físico donde se manifiesta el mensurando o del propio instrumento de medición.

<sup>23</sup> Decimos “en función de” en lugar de “igual a” porque podría haber una descalibración que desplazara la lectura en un valor constante o incluso una alinealidad en el método de medición que desplazara la lectura en forma diferente según el valor del mensurando.

### 2.3 Método y procedimiento de medición

Toda medición se realiza siguiendo un procedimiento de medición detallado que implementa un método de medición y obtiene el resultado ya sea en forma directa o a través de un modelo de medición.

El método de medición describe en forma genérica la estructura del proceso de medición. En su versión más simple se reduce a la lectura directa de la indicación de un instrumento de medición. En casos más elaborados puede involucrar el uso de varios transductores, el procesamiento de las magnitudes que éstos generan y la aplicación de correcciones. El método de medición se apoya en un *principio de medición*, es decir el o los fenómenos físicos en que se basa la medición (por ejemplo, la transducción electroacústica en un micrófono capacitivo), y se materializa en un *procedimiento de medición* que describe en detalle todos los pasos de la medición, abarcando la selección del instrumental y de las muestras, el relevamiento de variables ambientales o de otra naturaleza que pudieran influir en el resultado de la medición, la especificación de las condiciones de medición, las recomendaciones y precauciones, las fórmulas y cálculos a realizar y la estructura y contenido del *informe de la medición*.

Para mediciones típicas y de realización frecuente, el procedimiento de medición suele estar contenido en normas de alcance nacional o internacional. En casos más específicos pueden constituir documentos internos de un laboratorio o simplemente incluirse en el informe.

**EJEMPLO:** La medición de la potencia acústica radiada por una máquina en una cámara semianecoica se basa en la transducción electroacústica de la presión sonora en tensión eléctrica en un micrófono capacitivo y en la relación simple que existe entre la intensidad y la presión en un campo libre. El método de medición consiste en realizar mediciones en diversos puntos de una superficie imaginaria alrededor de la máquina, convertir las presiones eficaces en intensidades, realizar una suma ponderada y multiplicarla por el área de la superficie. El procedimiento de medición explicita las características del ambiente de ensayo, el tipo de instrumento a utilizar, las condiciones ambientales de referencia, las condiciones operativas de la máquina (por ejemplo, modo de fijación, tensión de alimentación, tiempo de calentamiento), las coordenadas de los puntos de medición, el tipo, la cantidad y duración de las mediciones a realizar en cada punto (por ejemplo mediciones en bandas de tercio de octava) y la orientación del micrófono en cada una, las fórmulas a utilizar en el cálculo, las correcciones según las condiciones ambientales, el efecto del ruido ambiente, y la forma de evaluar la incertidumbre.

#### 2.3.1 Métodos directos e indirectos de medición

Los métodos de medición se pueden clasificar en directos e indirectos. En los directos el mensurando se determina directamente de la indicación del instrumento de medición. Un ejemplo es la medición del nivel equivalente con un instrumento integrador.

Los métodos de medición indirectos realizan una transformación de los valores recabados antes de obtener el resultado deseado. Por ejemplo, a partir de la medición del nivel de presión sonora en un punto de una onda plana progresiva (campo libre) se puede determinar la velocidad de las partículas por aplicación de la ecuación

$$U_{\text{ef}} = \frac{1}{\rho_0 c} P_{\text{ref}} 10^{\frac{L_p}{20}}. \quad (2.1)$$

donde  $\rho_0$  es la densidad del aire en equilibrio y  $c$  es la velocidad del sonido.

## 2.4 Modelo de medición

Salvo en casos muy sencillos, en general el resultado de la medición no es simplemente la indicación del instrumento sino el resultado de una serie de operaciones, que muchas veces puede ser una relación funcional explícita o implícita entre dos o más variables o inclusive un algoritmo implementado en software.

Si llamamos  $Y$  al mensurando y  $X_n$  a las cantidades obtenidas por lectura directa o comparación, podemos plantear genéricamente la forma explícita

$$Y = f(X_1, \dots, X_N), \quad (2.2)$$

o la forma implícita,

$$h(Y, X_1, \dots, X_N) = 0. \quad (2.3)$$

En algunos casos será posible despejar directamente  $Y$  pasando a una forma explícita como la (2.2) y en otros podrá intentarse una resolución numérica por software, por ejemplo mediante el método de Newton-Raphson.

EJEMPLO 1: Si  $L_1, \dots, L_{10}$  son los niveles de las bandas de octava correspondientes a las frecuencias  $f_1 = 31,5$  Hz, ...,  $f_{10} = 16$  kHz, entonces puede calcularse el nivel ponderado A mediante la expresión:

$$L_{pA} = 10 \log \sum_{k=1}^{10} 10^{\frac{L_k + A_k}{10}}, \quad (2.4)$$

donde  $A_k$  es la ponderación A correspondiente a la banda centrada en  $f_{10}$ .

EJEMPLO 2: Si se conocen los niveles equivalentes ponderados A de un ruido,  $L_{Aeq, T_n}$ , correspondientes a  $N$  intervalos de tiempo de duración  $T_n$  que cubren la duración total  $T$  del periodo de interés (por ejemplo una jornada laboral) puede calcularse el nivel equivalente extendido a todo el período como

$$L_{Aeq, T} = 10 \log \left( \frac{1}{T} \sum_{n=1}^N T_n 10^{\frac{L_{Aeq, T_n}}{10}} \right). \quad (2.5)$$

EJEMPLO 3: En la determinación en cámara de transmisión horizontal del índice de reducción acústica  $R$  de un tabique, se miden los niveles de presión sonora en cada banda de tercio de octava en los recintos emisor (proveniente de una fuente adecuada) y receptor (proveniente de la transmisión a través del tabique y, eventualmente, la transmisión por flancos), y el tiempo de reverberación del recinto receptor. Si llamamos  $L_{k1}$  y  $L_{k2}$  a estos niveles y  $T_k$  al tiempo de reverberación en la  $k$ -ésima banda, puede calcularse el índice de reducción acústica en dicha banda como

$$R_k = L_{k1} - L_{k2} + 10 \log_{10} \frac{T_k S_{12}}{0,161 V_2}, \quad (2.6)$$

donde  $S_{12}$  es el área ocupada por el tabique y  $V_2$  es el volumen del recinto receptor.

EJEMPLO 4: Supongamos que queremos determinar el nivel de presión sonora en determinado punto, en unas condiciones de referencia caracterizadas por una temperatura de 23 °C, una humedad relativa ambiente de 50 % y una presión atmosférica de 1013,25 hPa, pero en realidad medimos a 27 °C, con una humedad del 75 % a una presión atmosférica de 1005,3 hPa. La medición se realiza utilizando un micrófono Brüel & Kjaer Type 4189.

Será necesario realizar una corrección, y debido a que las variaciones con respecto a las condiciones deseadas son pequeñas, podemos suponerla lineal:

$$L_p(T_1, h_{r1}, P_{a1}) = L_p(T_0, h_{r0}, P_{a0}) + K_T \Delta T + K_{h_r} \Delta h_r + K_{P_a} \Delta P_a \quad (2.7)$$

Para un micrófono Brüel & Kjaer Type 4189 se tiene (Brüel & Kjaer, 1995; Brüel & Kjaer, 1996):

$$K_T = -0,001 \text{ dB/}^\circ\text{C}$$

$$K_{h_r} = 0,001 \text{ dB/\%}$$

$$K_{P_a} = -0,001 \text{ dB/hPa}$$

El efecto de las condiciones ambientales resulta de 0,02895 dB, por lo que la corrección a efectuar al valor leído es de -0,02895 dB

EJEMPLO 5: Consideremos el siguiente método para determinar la frecuencia de resonancia  $f_0$  de un sistema. Se mide la respuesta  $Y_k$  en tres frecuencias  $f_k$  próximas a la frecuencia de resonancia (una manera puede ser seleccionar las frecuencias de modo que  $Y_1, Y_3 < Y_2$ ). Luego se obtiene una interpolación parabólica<sup>24</sup> y de allí se determina analíticamente a qué frecuencia se da el máximo. El modelo asociado a este método es el siguiente:

$$\begin{aligned} Y_{\max} + K(f_1 - f_0)^2 &= Y_1 \\ Y_{\max} + K(f_2 - f_0)^2 &= Y_2 \\ Y_{\max} + K(f_3 - f_0)^2 &= Y_3 \end{aligned} \quad (2.8)$$

Se trata de un sistema de tres ecuaciones con tres incógnitas,  $Y_{\max}$ ,  $K$ ,  $f_0$ , de las cuales sólo  $f_0$  es relevante al problema. Podemos eliminar  $Y_{\max}$  restando sucesivamente la primera menos la segunda y la primera menos la tercera, y luego podemos remover  $K$  dividiendo las dos ecuaciones así resultantes. Resulta

$$\frac{(f_1 - f_0)^2 - (f_2 - f_0)^2}{(f_1 - f_0)^2 - (f_3 - f_0)^2} = \frac{Y_1 - Y_2}{Y_1 - Y_3} \quad (2.9)$$

donde sólo figura la incógnita  $f_0$ . Es ésta una formulación implícita del modelo de medición. En este caso es claro que se podría despejar  $f_0$  ya que se trata básicamente de una

<sup>24</sup> La curva de resonancia no tiene forma exactamente parabólica, pero en las proximidades del pico se la puede aproximar por una parábola. La aproximación será tanto mejor cuanto más cercanas sean las frecuencias ensayadas a la de resonancia.

ecuación de segundo grado (que se reduce a una de primer grado). En casos más complicados podría no existir una solución explícita, o ser prohibitivamente complicada.<sup>25</sup>

## 2.5 Incertidumbre

Al medir cualquier mensurando, no sólo el resultado obtenido discrepa del valor verdadero sino que al repetir la medición los resultados diferirán probablemente entre sí. Esta característica de cualquier proceso de medición se conoce cualitativamente como *incertidumbre*. Entre las causas de la incertidumbre se encuentran:

- 1) Causas atribuibles al mensurando: Una definición incompleta del mensurando (por ejemplo, no se especifica una variable que tiene un efecto significativo sobre el mensurando); imposibilidad de garantizar que la realización del mensurando satisface la definición (por ejemplo, la definición especifica la temperatura y ésta se conoce sólo aproximadamente).
- 2) Causas atribuibles al procedimiento de medición: Modelo subyacente en el método incompleto, insuficiente o inexacto (por ejemplo, corrección del efecto de la temperatura en el que el coeficiente de temperatura se conoce aproximadamente); cantidad finita de repeticiones; presencia de interferencias o perturbaciones (por ejemplo, ruido de fondo); presencia de factores que alteran el mensurando (por ejemplo, una superficie reflectora del sonido próxima al punto de medición; excesivo efecto de las condiciones ambientales).
- 3) Causas atribuibles al instrumental: Resolución finita; falta de precisión; tolerancia de fabricación; falta de calibración; respuesta en frecuencia inadecuada.
- 4) Causas atribuibles al operador: Alteración del mensurando a causa de su presencia (por ejemplo, alteración del campo acústico que se desea medir); errores de paralaje en la lectura de instrumentos analógicos; en instrumentos digitales, indecisión sobre qué valor consignar ante una lectura cuyos dígitos menos significativos fluctúan permanentemente en el tiempo.

No se deben confundir estas fuentes de incertidumbre en la medición con los errores groseros o errores de procedimiento (Bell, 2001; COGUANOR, s/f), por ejemplo confundir una cifra al registrar un dato, confundir dos mensurandos diferentes, o apartarse del protocolo sin causas fundadas (y sin tener en cuenta los efectos que esto puede ocasionar). En todo lo que sigue se supone que estos vicios se han eliminado o corregido y que las mediciones han sido realizadas por personal idóneo.

En términos generales, la incertidumbre es la dispersión que cabe esperar en los resultados de diversas mediciones de un mensurando en condiciones especificadas. Obsérvese que en general no es necesario hacer mención explícita al valor verdadero. Más específicamente, se define la *incertidumbre estándar*,  $u$ , como el desvío estándar de las mediciones, es decir,<sup>26</sup>

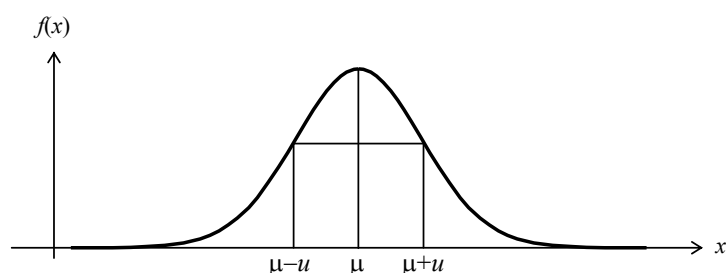
$$u = \sigma = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx}, \quad (2.10)$$

<sup>25</sup> Un ejemplo donde existen soluciones explícitas pero se prefieren métodos numéricos es en la resolución de sistemas de ecuaciones lineales de orden alto

<sup>26</sup> En el Apéndice 2 se introducen los conceptos básicos de estadística, y en el Glosario del Apéndice 1 se encuentra un resumen de las definiciones.



donde  $x$  es un resultado genérico de la medición realizada en condiciones especificadas,  $\mu$  es la media o valor esperado de la variable  $x$  y  $f(x)$  es su densidad de probabilidad. En la figura 25 se muestra un ejemplo.



**Figura 25.** Ejemplo de incertidumbre estándar para una función de densidad de probabilidad normal.

En muchos casos no se conoce el desvío estándar de las mediciones, entendido como el correspondiente a la población<sup>27</sup> de todas las mediciones posibles del mismo mensurando en condiciones especificadas, sino el desvío estándar de una *muestra*, es decir una cantidad finita  $n$  de mediciones. En ese caso, la incertidumbre estándar se obtiene como el desvío estándar muestral.

$$u = s = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2}. \quad (2.11)$$

donde  $x_k$  es el  $k$ -ésimo valor medido y  $\bar{x}$  es la media de dichos valores.

Ambas formas de determinar la incertidumbre estándar se consideran válidas. Más aún, como veremos luego, la incertidumbre puede tener varias componentes. Esto da origen a una clasificación de las componentes de la incertidumbre en dos categorías, denominadas incertidumbre tipo A e incertidumbre tipo B, descriptas en lo que sigue.

### 2.5.1 Incertidumbre tipo A

Es cualquier componente de la incertidumbre que se determina empíricamente por métodos estadísticos a partir de los datos obtenidos repitiendo la medición una cierta cantidad de veces. Por ejemplo, supongamos que se han medido los siguientes diez valores de nivel de presión sonora:

73,6	80,2	76,7	75,5	76,7	75,3	75,4	77,8	77,7	77,7
------	------	------	------	------	------	------	------	------	------

Entonces, aplicando (2.11) resulta una media 76,66 dB y un desvío estándar 1,83 dB. Una segunda muestra de mediciones correspondientes a la misma población es:

76,6	73,8	76,7	78,0	76,3	77,2	76,7	75,1	76,0	74,4
------	------	------	------	------	------	------	------	------	------

<sup>27</sup> El término “población” se utiliza en Estadística para hacer referencia a la totalidad de los elementos homogéneos sobre los que se especifican parámetros estadísticos

En este caso se obtiene una media 76,08 dB y un desvío estándar 1,29 dB. Se aprecia que los valores estimados pueden diferir bastante entre dos muestras.

### 2.5.2 Incertidumbre tipo B

Es cualquier componente de la incertidumbre que se obtiene por otros métodos que no se basan en el tratamiento estadístico de los propios datos, entre ellos:

- a) El conocimiento o estimación previa por consenso científico o técnico, por ejemplo incertidumbres en constantes físicas o en procedimientos
- b) Información provista en normas
- c) Especificaciones del instrumental empleado
- d) Información provista en certificados de calibración
- e) Simulaciones por software
- f) Aplicación de leyes físicas

NOTA: Probablemente el cuerpo de conocimientos involucrado en la mayor parte de las componentes de incertidumbre tipo B se haya obtenido originalmente por métodos estadísticos. La distinción importante aquí es que dichos métodos se aplicaron a datos generales o material de investigación, y no a los propios datos medidos en una ocasión particular, que es a lo que se refiere la incertidumbre tipo A.

#### EJEMPLOS:

a) La constante  $R$  de los gases interviene en el cálculo de la velocidad  $c$  del sonido. Según información de CODATA (Mohr et al., 2007),  $R = 8,314\,472\,15\text{ J}/(\text{mol}\cdot\text{K})$  con una incertidumbre estándar  $u = 0,000\,014\,13\text{ J}/(\text{mol}\cdot\text{K})$ . Esta incertidumbre debe ser tomada en cuenta y propagada a la velocidad del sonido.

b) La norma ISO 1996 establece que, en mediciones de ruido de tránsito y en ausencia de mejores datos, la incertidumbre estándar está dada por  $10/\sqrt{n}$  dB, donde  $n$  es la cantidad de vehículos que han pasado frente al micrófono durante la medición.

c) Un sonómetro especificado como clase 1 tendrá, a 1 kHz, una incertidumbre estándar de 0,7 dB (IEC 61672).

d) El certificado de calibración de un calibrador acústico indica una incertidumbre estándar de 0,15 dB. Por consiguiente, un sonómetro calibrado con este calibrador incorpora una componente de incertidumbre de 0,15 dB debida a la calibración.

e) Si se mide un ruido cuyo espectro tiene un comportamiento con la frecuencia conocido (por ejemplo ruido rosa), es posible obtener a partir de simulaciones por software la incertidumbre en el resultado de una medición con ponderación A conociendo la incertidumbre de la respuesta en cada frecuencia (dada por la norma IEC 61672), para lo cual se generará una gran cantidad de respuestas compatibles con lo especificado en la norma y se determinará el nivel ponderado A con cada una de esas respuestas. Finalmente se determinará por análisis estadístico, el desvío estándar de los valores obtenidos.

f) Si se conoce una expresión matemática de la cantidad a determinar en función de otras cantidades a medir, es posible propagar las incertidumbres de estas últimas para obtener la incertidumbre en el valor. La forma de hacerlo se verá oportunamente (sección 2.5.4).

### 2.5.3 Incertidumbre expandida

La incertidumbre estándar  $u$ , como desvío estándar de los valores medidos, por sí sola no da una idea unívoca de cuán dispersos pueden hallarse los posibles resultados de la medición. Podemos preguntarnos, por ejemplo, cuál es la probabilidad de que el resultado se encuentre en un rango  $\pm u$  con respecto a la media. La respuesta depende del tipo de distribución estadística del resultado de la medición. Si es una distribución normal, dicha probabilidad será de un 68,3 %. Si es una distribución uniforme, será de un 57,7 %.

Se introduce, en consecuencia, la *incertidumbre expandida*  $U$ , de modo tal que el intervalo  $\pm U$  respecto a la media abarque un porcentaje especificado de los posibles resultados de la medición. Expresado de otra forma, el resultado de la medición caerá en dicho intervalo con una probabilidad  $p$  especificada (que puede expresarse en %).

Dicha probabilidad, denominada *probabilidad de cobertura*, no está normalizada y, por lo tanto, existe cierta libertad para elegirla según las necesidades. En cualquier caso debería explicitarse el valor adoptado, y a veces se lo incluye como subíndice:  $U_p$ .

En el caso más frecuente en que la distribución es normal (o aproximadamente normal), se suele adoptar valores de 95,4 % y 99,7 % que corresponden, respectivamente, a  $2u$  y  $3u$ .

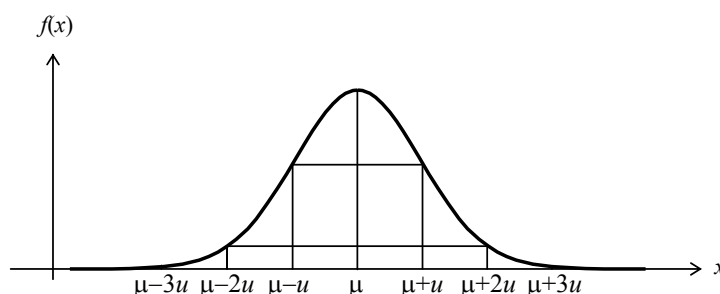
El factor  $k$  que permite pasar de  $u$  a  $U$  se denomina *factor de cobertura*:

$$U = k u. \quad (2.12)$$

Para el caso normal podemos escribir, por ejemplo

$$U_{95} = 2u. \quad (2.13)$$

donde, como es práctica habitual, para aliviar escritura se ha escrito  $U_{95}$  en lugar de  $U_{95,4}$ . En la figura 26 se muestran algunos ejemplos.



**Figura 26.** Ejemplos de intervalos correspondientes a diferentes probabilidades de cobertura para una función de densidad de probabilidad normal.

### 2.5.4 Incertidumbre estándar combinada

Vimos que en muchos casos la medición requiere a su vez medir varias cantidades subsidiarias o intermedias y combinarlas de acuerdo al modelo de medición pertinente. Nos preguntamos en qué medida contribuirá la incertidumbre asociada a cada cantidad intermedia a la incertidumbre general.

Para fijar ideas, analizaremos primero el caso más elemental en que el modelo sea una simple combinación lineal de varias variables independientes:

$$Y = a_1 X_1 + \dots + a_N X_N. \quad (2.14)$$

En primer lugar notemos que si  $\mu_k$  es la media de  $X_k$ , entonces la media de  $Y$ , es decir,  $\mu_Y$ , se obtiene con la misma combinación lineal:

$$\mu_Y = a_1 \mu_1 + \dots + a_N \mu_N. \quad (2.15)$$

Asimismo, se puede demostrar que el desvío estándar  $\sigma_Y$  satisface

$$\sigma_Y^2 = \sum_{i=1}^N \sum_{j=1}^N a_i a_j \text{cov}(X_i, X_j), \quad (2.16)$$

donde  $\text{cov}(X_i, X_j)$  es la covarianza entre  $X_i$  y  $X_j$  dada por

$$\text{cov}(X_i, X_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mu_i)(x_j - \mu_j) f(x_i, x_j) dx dy, \quad (2.17)$$

siendo  $f(x_i, x_j)$  la densidad de probabilidad conjunta (ver Apéndice 1).

Definimos la incertidumbre estándar combinada,  $u_c$ , como el desvío estándar total dado en (2.16).

$$u_c = \sigma_Y. \quad (2.18)$$

En el caso de una incertidumbre tipo A, donde la covarianza se estima a partir de una muestra finita, tendremos

$$u_c = s_Y = \sqrt{\sum_{i=1}^N \sum_{j=1}^N a_i a_j c_{ij}}, \quad (2.19)$$

donde  $c_{ij}$  es la covarianza muestral entre  $x_i$  y  $x_j$  calculada a partir de una muestra de tamaño  $n$  de ambas variables mediante la ecuación

$$c_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i)(y_{jk} - \bar{y}_j). \quad (2.20)$$

En muchos casos las variables intervinientes pueden considerarse estadísticamente independientes, en cuyo caso las covarianzas cruzadas son 0. Dado que la covarianza de una variable consigo misma es la varianza  $s^2$ , la ecuación (2.19) queda

$$s_Y = \sqrt{\sum_{i=1}^N a_i^2 s_i^2}. \quad (2.21)$$

Dado que la varianza de cada variable puede expresarse como la incertidumbre de esa variable al cuadrado, podemos reescribir la ecuación anterior en la forma

$$u_c = \sqrt{\sum_{i=1}^N a_i^2 u_i^2}. \quad (2.22)$$

El caso más general en el que el modelo de medición (es decir la relación entre el mensurando y las cantidades intermedias medidas en forma directa) no necesariamente es lineal, como en la ecuación (2.2),

$$Y = f(X_1, \dots, X_N),$$

puede analizarse en forma similar. Para ello se realiza un desarrollo de Taylor de primer grado alrededor de los valores medidos. Se demuestra en el Apéndice 8 que la varianza muestral del mensurando vale

$$s_y^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial x_i}(\bar{x}_i) \frac{\partial f}{\partial x_j}(\bar{x}_j) c_{ij}, \quad (2.23)$$

donde  $c_{ij}$  es la covarianza entre las variables  $x_i$  y  $x_j$  estimada mediante la fórmula muestral

$$c_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j). \quad (2.24)$$

En el caso en que las variables aleatorias  $X_1, \dots, X_N$  sean independientes, las covarianzas entre variables diferentes tienden a anularse y por lo tanto sólo aparecen las covarianzas entre una variable y sí misma, es decir, las varianzas:

$$s_y^2 = \sum_{i=1}^N \left( \frac{\partial f}{\partial x_i}(\bar{x}_i) \right)^2 s_i^2, \quad (2.25)$$

o bien

$$u_c = \sqrt{\sum_{i=1}^N \left( \frac{\partial f}{\partial x_i}(\bar{x}_i) \right)^2 u_i^2}, \quad (2.26)$$

Las derivadas parciales pueden calcularse analíticamente o numéricamente. Incluso pueden calcularse en caso de funciones implícitas (ver un caso en el apéndice 9).

## 2.6 Ejemplos

A continuación se analizan diversas situaciones que aparecen con frecuencia en mediciones acústicas, ya sean de campo o de laboratorio.

### 2.6.1 Relación entre las incertidumbres en el nivel y en la presión

Aunque el micrófono de un sonómetro produce una señal eléctrica proporcional a la presión sonora, los sonómetros calculan y muestran el nivel de presión sonora  $L_p$ . Conocido éste se puede calcular la presión sonora por medio de la fórmula

$$P_{\text{ef}} = P_{\text{ref}} 10^{L_p / 20}. \quad (2.27)$$

Para obtener la incertidumbre conviene pasar la exponencial a base  $e$ :

$$P_{\text{ef}} = P_{\text{ref}} e^{\frac{\ln 10}{20} L_p}. \quad (2.28)$$

Ahora es más fácil derivar. Hecho el cálculo y reemplazado en (2.26), resulta, finalmente, la siguiente expresión

$$u_{P_{\text{ef}}} = \frac{\ln 10}{20} P_{\text{ef}} u_{L_p}. \quad (2.29)$$

Vemos que si la incertidumbre de  $L_p$  es independiente del nivel, entonces la incertidumbre de  $u_{P_{\text{ef}}}$  será proporcional a la presión.

### 2.6.2 Incertidumbre en la corrección de condiciones ambientales

Consideremos el caso del ejemplo 4 de la sección 2.4, donde el nivel de presión sonora leído debe corregirse para llevarlo a unas condiciones ambientales de referencia diferentes de las que había en el momento de la medición. El modelo de medición era

$$L_p(T_1, h_{r1}, P_{a1}) = L_p(T_0, h_{r0}, P_{a0}) + K_T \Delta T + K_{h_r} \Delta h_r + K_{P_a} \Delta P_a. \quad (2.30)$$

El problema es que la determinación de las condiciones ambientales adolece de incertidumbre. Supongamos que la incertidumbre estándar en la medición de las respectivas variables es:

$$\begin{aligned} u_T &= 0,8 \text{ }^\circ\text{C} \\ u_{h_r} &= 3 \% \\ u_{P_a} &= 3 \text{ hPa} \end{aligned}$$

La incertidumbre combinada debida a estos factores es de tipo B, ya que surge de las especificaciones y/o normas cumplidas por los instrumentos (termómetro, higrómetro y barómetro). Su valor, calculado mediante la ecuación (2.22) bajo el supuesto razonable de independencia de dichas variables, es:

$$u_{c, \text{amb}} = \sqrt{(-0,001 \times 0,8)^2 + (0,001 \times 3)^2 + (-0,001 \times 3)^2} \text{ dB} = 0,0043 \text{ dB}.$$

Vemos que el aporte a la incertidumbre de los errores de medición de las condiciones ambientales es muy pequeño. Por esta razón, en general no se requieren instrumentos de gran jerarquía para medir dichas variables.

Para obtener la incertidumbre combinada total habría que agregar la incertidumbre en la lectura del instrumento, cuyo tratamiento veremos un poco más adelante.

### 2.6.3 Incertidumbre en el cálculo de nivel equivalente

Supongamos que se ha medido o estimado el nivel equivalente  $L_{Aeq,T_i}$  en  $N$  intervalos de tiempo  $T_i$  y queremos obtener el nivel equivalente total mediante la fórmula

$$L_{Aeq,T} = 10 \log \left( \frac{1}{T} \sum_{i=1}^N T_i 10^{\frac{L_{Aeq,T_i}}{10}} \right). \quad (2.31)$$

Este problema se presenta, por ejemplo cuando se desea determinar el nivel equivalente durante una jornada laboral a partir de la medición de cada proceso individual en que se descompone la secuencia diaria de tareas. Estas mediciones se hacen durante tiempos relativamente cortos y se extrapolan a la duración total de cada proceso.

Nos interesa determinar la incertidumbre estándar combinada, la cual tiene dos tipos de componentes: la debida a los niveles, y la debida a los tiempos. Nótese particularmente que la incertidumbre debida a los tiempos no es sólo la atribuible al proceso de medición (incluyendo el instrumento de medición, en este caso un cronómetro) sino que incluye una componente asociada a la indefinición del propio mensurando, ya que el tiempo de un proceso puede depender de factores de difícil control y que se ajustan dinámicamente en función del resultado.

Tenemos, entonces,

$$L_{Aeq,T} = g(L_1, \dots, L_N, T_1, \dots, T_N) = 10 \log \left( \frac{1}{T} \sum_{i=1}^N T_i 10^{\frac{L_{Aeq,T_i}}{10}} \right)$$

Para calcular las derivadas parciales conviene expresar el logaritmo como logaritmo natural, que es más fácil de derivar. Resulta, para las derivadas con respecto a  $L_{Aeq,T_i}$ ,

$$\begin{aligned} \frac{\partial g}{\partial L_{Aeq,T_i}} &= \frac{10}{\ln 10} \frac{\frac{\partial}{\partial L_{Aeq,T_i}} \left( \frac{1}{T} \sum_{i=1}^N T_i e^{\ln 10 \frac{L_{Aeq,T_i}}{10}} \right)}{\frac{1}{T} \sum_{i=1}^N T_i 10^{\frac{L_{Aeq,T_i}}{10}}} = \\ &= \frac{10}{\ln 10} \frac{T_i \frac{\ln 10}{10} 10^{\frac{L_{Aeq,T_i}}{10}}}{\frac{1}{T} \sum_{i=1}^N T_i 10^{\frac{L_{Aeq,T_i}}{10}}} \end{aligned}$$

es decir,

$$\frac{\partial g}{\partial L_{\text{Aeq}, T_i}} = \frac{T_i}{T} 10^{\frac{L_{\text{Aeq}, T_i} - L_{\text{Aeq}, T}}{10}}. \quad (2.32)$$

Y para las derivadas con respecto a  $T_i$ , teniendo en cuenta que  $T = \sum T_i$ ,

$$\frac{\partial g}{\partial T_i} = \frac{10}{\ln 10} \frac{\frac{T - T_i}{T^2} 10^{\frac{L_{\text{Aeq}, T_i}}{10}}}{\frac{1}{T} \sum_{i=1}^N T_i 10^{\frac{L_{\text{Aeq}, T_i}}{10}}} = \frac{10}{\ln 10} \frac{T - T_i}{T^2} 10^{\frac{L_{\text{Aeq}, T_i} - L_{\text{Aeq}, T}}{10}}. \quad (2.33)$$

Como caso numérico, supongamos tres intervalos de duraciones 2 h, 3 h y 3 h, determinados con incertidumbre estándar de 12 min, cuyos niveles equivalentes fueron 85,1 dBA, 83,0 dBA y 80,2 dBA respectivamente, con incertidumbre estándar de 1 dB. Entonces el nivel equivalente total resulta

$$L_{\text{Aeq}, T} = 10 \log \left( \frac{2}{8} 10^{\frac{85,1}{10}} + \frac{3}{8} 10^{\frac{83,0}{10}} + \frac{3}{8} 10^{\frac{80,2}{10}} \right) \text{dBA} = 82,9 \text{ dBA}$$

Las derivadas con respecto a los niveles son

$$\frac{\partial g}{\partial L_{\text{Aeq}, T_1}} = \frac{2}{8} 10^{\frac{85,1 - 82,9}{10}} = 0,415.$$

$$\frac{\partial g}{\partial L_{\text{Aeq}, T_2}} = \frac{3}{8} 10^{\frac{83,0 - 82,9}{10}} = 0,384.$$

$$\frac{\partial g}{\partial L_{\text{Aeq}, T_3}} = \frac{3}{8} 10^{\frac{80,2 - 82,9}{10}} = 0,201.$$

Las derivadas temporales se obtienen simplemente multiplicando por  $10/(T_k \ln 10)$ , es decir,

$$\frac{\partial g}{\partial T_1} = 0,676 \frac{\text{dBA}}{\text{h}}.$$

$$\frac{\partial g}{\partial T_2} = 0,347 \frac{\text{dBA}}{\text{h}}.$$

$$\frac{\partial g}{\partial T_3} = 0,182 \frac{\text{dBA}}{\text{h}}.$$



Entonces

$$u_c = \sqrt{(0,415^2 + 0,384^2 + 0,201^2) \times 1^2 + (0,676^2 + 0,347^2 + 0,182^2) \times 0,2^2} = 0,62 \text{ dBA}$$

Obsérvese que la incertidumbre es menor que la de cada medición (1 dB). Esto se debe al efecto de reducción de dispersión en un promedio de varias variables.

Si no hubiera errores temporales y los tres niveles fueran iguales y tuvieran la misma ponderación, entonces la incertidumbre sería  $1/\sqrt{3}$  dB = 0,58 dB.

#### 2.6.4 Incertidumbre en el nivel A a partir del espectro

Supongamos que se midieron los niveles de banda de octava  $\{L_1, \dots, L_{10}\}$  de un determinado ruido. Entonces el nivel ponderado A puede obtenerse como

$$L_A = 10 \log \left( \sum_{i=1}^{10} 10^{\frac{L_i + A_i}{10}} \right), \quad (2.34)$$

donde  $A_i$  es la ponderación A correspondiente a la  $i$ -ésima banda. Los niveles de banda medidos son variables aleatorias no necesariamente independientes. Para determinar la incertidumbre será preciso determinar las derivadas parciales de  $L_A$  con respecto a los niveles  $L_i$  y las covarianzas de dichos niveles.

Las derivadas parciales pueden calcularse convirtiendo el logaritmo decimal y las potencias de 10 a logaritmo natural y exponencial de base  $e$ :

$$L_A = \frac{10}{\ln 10} \ln \left( \sum_{i=1}^{10} e^{\ln 10 \frac{L_i + A_i}{10}} \right).$$

Entonces, aplicando regla de la cadena,

$$\frac{\partial L_A}{\partial L_i} = \frac{10}{\ln 10} \frac{1}{\left( \sum_{i=1}^{10} e^{\ln 10 \frac{L_i + A_i}{10}} \right)} e^{\ln 10 \frac{L_i + A_i}{10}} \frac{\ln 10}{10} = \frac{10^{\frac{L_i + A_i}{10}}}{10^{\frac{L_A}{10}}}$$

de donde

$$\frac{\partial L_A}{\partial L_i} = 10^{\frac{L_i + A_i - L_A}{10}}. \quad (2.35)$$

La obtención de la matriz de covarianza para calcular una incertidumbre tipo B sólo es posible si se conocen propiedades físicas del fenómeno que origina el ruido a medir. Caso contrario se debe realizar una serie de mediciones y estimar a partir de ellas la matriz de covarianza. Si  $L_{i,k}$  y  $L_{j,k}$  son los niveles de las bandas  $i$ -ésima y  $j$ -ésima correspondientes a la medición  $k$ , entonces las covarianzas se pueden estimar mediante

$$u(L_i, L_j) = \frac{1}{n-1} \sum_{k=1}^n (L_{ik} - \bar{L}_i) (L_{jk} - \bar{L}_j) \quad (2.36)$$

donde  $\bar{L}_i$  y  $\bar{L}_j$  son los niveles promedio de las bandas  $i$ -ésima y  $j$ -ésima respectivamente y  $n$ , la cantidad de espectros medidos. Finalmente podemos obtener la incertidumbre combinada estándar como

$$u_c = \sqrt{\sum_{i=1}^{10} \sum_{j=1}^{10} \frac{\partial L_A}{\partial L_i} \frac{\partial L_A}{\partial L_j} u(L_i, L_j)}. \quad (2.37)$$

Esta incertidumbre corresponde a la estimación del nivel A a partir de una única medición. Sin embargo, habiendo realizado  $n$  mediciones, es natural estimar el nivel A promediando los resultados de las mediciones. En tal caso las covarianzas quedan divididas por  $n$ , lo cual resulta en una mayor confiabilidad de la estimación de  $L_A$ .

Nótese que cuando se está midiendo un tipo de ruido específico (por ejemplo ruido de tránsito) las diferentes bandas del espectro no son independientes y por lo tanto las covarianzas cruzadas ( $i \neq j$ ) no son nulas.

### 2.6.5 Incertidumbre en la medición del índice de reducción acústica

El índice de reducción acústica de un material se mide en el laboratorio mediante la medición de los niveles a ambos lados de una cámara de transmisión entre los que se intercala una muestra del material a evaluar. También se mide el tiempo de reverberación  $T_k$  de la cámara receptora, el área  $S_{12}$  de la muestra a colocar en el vano entre ambos recintos y el volumen  $V_2$  del recinto receptor. La medición se realiza en bandas de tercio de octava. Para la  $k$ -ésima banda se tiene

$$R_k = L_{1k} - L_{2k} + 10 \log_{10} \frac{T_k S_{12}}{0,161 V_2}, \quad (2.38)$$

Las derivadas con respecto a  $L_{1k}$  y  $L_{2k}$  son, respectivamente, 1 y  $-1$ . Las derivadas con respecto a  $T_k$ ,  $S_{12}$  y  $V_2$  son, respectivamente

$$\frac{\partial R_k}{\partial T_k} = \frac{10}{\ln 10} \frac{1}{T_k} \quad (2.39)$$

$$\frac{\partial R_k}{\partial S_{12}} = \frac{10}{\ln 10} \frac{1}{S_{12}} \quad (2.40)$$

$$\frac{\partial R_k}{\partial V_2} = -\frac{10}{\ln 10} \frac{1}{V_2} \quad (2.41)$$

EJEMPLO: Supongamos que se ha medido la  $k$ -ésima banda, habiéndose obtenido los valores siguientes:  $T_k = 2,3$  s,  $L_{1k} = 87,3$  dB,  $L_{2k} = 45,6$  dB,  $S_{12} = 10,4$  m<sup>2</sup>,

$V_2 = 51,84 \text{ m}^3$ . En cuanto a las incertidumbres, se sabe que  $u_L = 1 \text{ dB}$ ,  $u_T = 0,2 \text{ s}$ ,  $u_S = 0,099 \text{ m}^2$ ,  $u_V = 0,38 \text{ m}^3$ . Tenemos

$$\begin{aligned}\frac{\partial R_k}{\partial T_k} &= 1,888 \frac{\text{dB}}{\text{s}} \\ \frac{\partial R_k}{\partial S_{12}} &= 0,418 \frac{\text{dB}}{\text{m}^2} \\ \frac{\partial R_k}{\partial V_2} &= -0,0838 \frac{\text{dB}}{\text{m}^3}\end{aligned}$$

Entonces, la incertidumbre combinada para  $R_k$  será

$$\begin{aligned}u_{c,R_k} &= \sqrt{\left(\frac{\partial R_k}{\partial L_{1k}}\right)^2 u_{L_{1k}}^2 + \left(\frac{\partial R_k}{\partial L_{2k}}\right)^2 u_{L_{2k}}^2 + \left(\frac{\partial R_k}{\partial T_k}\right)^2 u_{T_k}^2 + \left(\frac{\partial R_k}{\partial S_{12}}\right)^2 u_{S_{12}}^2 + \left(\frac{\partial R_k}{\partial V_2}\right)^2 u_{V_2}^2} = \\ &= \sqrt{1^2 \times 1^2 + 1^2 \times 1^2 + 1,888^2 \times 0,2^2 + 0,418^2 \times 0,099^2 + 0,0838^2 \times 0,38^2} = \\ &= \sqrt{1 + 1 + 0,143 + 0,0017 + 0,00022} = 1,46 \text{ dB}\end{aligned}$$

Ex-profeso se detallan los resultados intermedios para hacer notar que la mayor contribución a la incertidumbre total corresponde a las componentes acústicas, es decir, las mediciones de los niveles a uno y otro lado del tabique bajo ensayo.

## 2.7 Incertidumbre y resolución

La resolución finita de un instrumento de medición implica una incertidumbre ya que cualquier valor que caiga entre dos divisiones consecutivas de la escala o entre dos dígitos menos significativos. Podemos suponer que la distribución estadística del error es uniforme, siendo su amplitud igual a la resolución. Según se detalla en el apéndice 2, si llamamos  $\Delta x_{\min}$  a la resolución, la incertidumbre estándar será

$$u_{\text{resol}} = \frac{\Delta x_{\min}}{\sqrt{12}}. \quad (2.42)$$

Por ser la distribución estadística uniforme, la incertidumbre expandida es menor que para una distribución gaussiana. Por ejemplo, la  $U_{95}$  se logra con  $k$  tal que

$$k \frac{\Delta x_{\min}}{\sqrt{12}} \frac{1}{\Delta x_{\min}} = \frac{0,95}{2},$$

es decir, un factor de cobertura

$$k = \frac{0,95 \sqrt{12}}{2} = 1,65.$$

Sin embargo esta componente de la incertidumbre no aparece sola sino en conjunto con otras. El teorema central del límite dice que al sumar muchas variables aleatorias la distribución resultante tiende a ser normal. Por otra parte, en los instrumentos de medición acústica en general se cumple que la componente debida a la resolución es considerablemente menor que las otras componentes de la incertidumbre.

EJEMPLO: Si un sonómetro tiene una resolución de 0,1 dB, entonces la incertidumbre debida a la resolución será

$$u_{\text{resol}} = \frac{0,1 \text{ dB}}{\sqrt{12}} = 0,029 \text{ dB}.$$

Resulta despreciable frente a la incertidumbre propia del instrumento, del orden de 1 dB.

## 2.8 Incertidumbre y error sistemático

Una situación muy frecuente en metrología es la presencia de un *error sistemático*, es decir, un sesgo o desplazamiento constante o predecible de todas las mediciones realizadas bajo determinadas condiciones. Puede deberse al operador, al procedimiento o al instrumento de medición.

En el caso del operador, la causa suele ser un error de paralaje al realizar la lectura de un instrumento de deflexión (de aguja). En el caso del procedimiento, podría ser la falta de corrección de un efecto (conocido o no), por ejemplo, una dependencia de la temperatura. En el caso del instrumento en sí, la causa típica es una descalibración.

El error sistemático puede ser aditivo (error de cero), multiplicativo (error de escala) o no lineal (error de linealidad).

### 2.8.1 Error sistemático aditivo

Podemos definirlo como la diferencia entre la media de los valores observados,  $\mu_o$ , y el valor verdadero  $x$  del mensurando,<sup>28</sup> es decir,

$$\varepsilon_{\text{sist}} = \mu_o - x. \quad (2.43)$$

Para poder aplicar (2.43) necesitaríamos conocer la media poblacional de todas las posibles mediciones,  $\mu_o$ , y el valor verdadero de  $x$ , lo cual no es posible. Podemos, en cambio, particularizar (2.43) para el caso de un patrón de referencia, por ejemplo la longitud de una varilla patrón, la masa de una pesa patrón o el nivel de un tono de calibración. En cualquier caso se trata de una magnitud con un valor convencional conocido y con una incertidumbre estándar también conocida (en general menor que la del instrumento de medición). Llamando  $x_{\text{cal}}$  a dicho valor, podemos estimar la media de los valores observados  $\mu_{x_{\text{cal}},o}$  mediante la media  $\bar{x}_{\text{cal},o}$  de una muestra de  $m$  mediciones, y adoptar el valor convencional como valor verdadero. Nuestra estimación del error sistemático será, entonces,

$$\hat{\varepsilon}_{\text{sist}} = \bar{x}_{\text{cal},o} - x_{\text{cal}}. \quad (2.44)$$

<sup>28</sup> En forma equivalente, puede definirse como la media de la diferencia entre el valor medido y el valor verdadero

donde  $\bar{x}_{\text{cal},0}$  es la media de  $m$  observaciones o mediciones de la referencia. La incertidumbre de esta estimación será

$$u_{\hat{\varepsilon}_{\text{sist}}} = \sqrt{u_{\bar{x}_{\text{cal},0}}^2 + u_{x_{\text{cal}}}^2}, \quad (2.45)$$

o bien, llamando  $u_o$  a la incertidumbre de una medición individual,

$$u_{\hat{\varepsilon}_{\text{sist}}} = \sqrt{\frac{u_o^2}{m} + u_{x_{\text{cal}}}^2}, \quad (2.46)$$

Una vez estimado el error sistemático, podemos estimar el valor del mensurando corregido para eliminar efectos sistemáticos como

$$x_c = \bar{x}_o - \hat{\varepsilon}_{\text{sist}}. \quad (2.47)$$

donde  $\bar{x}_o$  es la media de  $n$  observaciones (no necesariamente igual al  $m$  anterior) del mensurando a medir.

La incertidumbre en esta determinación será

$$u_{x_c} = \sqrt{u_{x_c}^2 + u_{\hat{\varepsilon}_{\text{sist}}}^2}, \quad (2.48)$$

o bien, en términos de la incertidumbre general para una medición individual:

$$u_{x_c} = \sqrt{\frac{u_o^2}{n} + u_{\hat{\varepsilon}_{\text{sist}}}^2}. \quad (2.49)$$

Es interesante observar que si bien se puede (y se debe) corregir el error sistemático, no se puede eliminar su incertidumbre.

**EJEMPLO:** Supongamos que realizamos cuatro mediciones de un ruido en idénticas condiciones, obteniendo 83,3 dB, 83,6 dB, 84,1 dB y 83,0 dB con un instrumento cuya incertidumbre estándar es 1 dB. Antes de realizar la medición se constata la calibración con un calibrador acústico cuyo valor nominal es 93,85 dB y cuyo certificado de calibración reciente indica una incertidumbre estándar de 0,15 dB y un error de -0,1 dB. Se constata tres veces la calibración obteniéndose en los tres casos un valor de 94,1 dB.

Calculemos primero la corrección y su incertidumbre. El valor nominal del patrón es 93,85 dB, pero como el error sistemático del calibrador es de -0,1 dB, entonces un instrumento perfecto mediría

$$L_{\text{cal}} + \varepsilon_{L_{\text{cal}},\text{sist}} = 93,75 \text{ dB}$$

Nuestro instrumento mide

$$\bar{x}_{\text{cal,o}} = \frac{94,1 + 94,1 + 94,1}{3} \text{ dB} = 94,1 \text{ dB}$$

El error sistemático estimado debido a la descalibración será entonces

$$\hat{\varepsilon}_{\text{sist}} = 94,1 \text{ dB} - 93,75 \text{ dB} = 0,35 \text{ dB}.$$

La incertidumbre de este valor será

$$\begin{aligned} u_{\hat{\varepsilon}_{\text{sist}}} &= \sqrt{\frac{1^2 + (0,1/\sqrt{12})^2}{3} + 0,15^2} = \\ &= \sqrt{0,3333 + 0,0003 + 0,0225} = 0,597 \text{ dB} \end{aligned}$$

El segundo término dentro de la raíz corresponde a la incertidumbre por resolución finita, que es en este caso prácticamente despreciable.

Vayamos ahora al valor medido y corregido del mensurando:

$$x_c = \frac{83,3 + 83,6 + 84,1 + 83,0}{4} - 0,35 = 83,15 \text{ dB}$$

La incertidumbre estándar será

$$\begin{aligned} \varepsilon_{x_c} &= \sqrt{\frac{1^2 + (0,1/\sqrt{12})^2}{4} + 0,597^2} = \\ &= \sqrt{0,25 + 0,0002 + 0,3564} = 0,78 \text{ dB} \end{aligned}$$

Vemos, nuevamente, que la incertidumbre por resolución finita tiene poca importancia frente a la incertidumbre del instrumento y de la calibración. Vemos también que esta última finalmente prima cuando se promedian muchas mediciones. En este caso, ya con promediar 4 mediciones la incertidumbre del instrumento es menor que la incertidumbre en la calibración.

### 2.8.2 Error sistemático multiplicativo

También llamado *error de escala* o *error sistemático relativo*, se presenta cuando el principio físico de la medición es del tipo

$$y = Kx, \quad (2.50)$$

donde  $x$  es un mensurando que no se puede medir directamente (por ejemplo una presión sonora),  $y$ , una variable asociada que sí se puede medir directamente (por ejemplo una tensión eléctrica) y  $K$ , la sensibilidad del transductor encargado de convertir el mensurando en una cantidad medible.

Lo que realmente se mide, por lo tanto, es  $y$ , luego de lo cual se podría obtener el valor deseado como

$$x = \frac{1}{K} y, \quad (2.51)$$

El error surge cuando la constante  $K$  de (2.50) difiere de la de (2.51). En general en (2.50) hay una constante  $K_{\text{real}}$  desconocida, pero al determinar  $x$  en (2.51) se supone una constante nominal o ideal  $K_{\text{ideal}}$ . En consecuencia la relación entre el valor observado  $x_o$  y el verdadero será

$$x_o = \frac{K_{\text{real}}}{K_{\text{ideal}}} x. \quad (2.52)$$

Surge aquí otra complicación: el valor de  $y$  medido posee a su vez cierta incertidumbre debido a que el instrumento utilizado no es ideal, de modo que (2.50) debe reescribirse como

$$y = K_{\text{real}} x + \varepsilon_y, \quad (2.53)$$

donde  $\varepsilon_y$  es una variable aleatoria centrada en 0 que representa la dispersión en la medición de  $y$ . Entonces el valor observado de  $x$  será

$$x_o = \frac{1}{K_{\text{ideal}}} y = \frac{K_{\text{real}}}{K_{\text{ideal}}} x - \frac{1}{K_{\text{ideal}}} \varepsilon_y. \quad (2.54)$$

Dado que  $\varepsilon_y$  tiene media 0, podríamos eliminar el error aleatorio de medición tomando la media  $\mu_o$  de los valores observados de  $x$ :

$$\mu_o = \frac{K_{\text{real}}}{K_{\text{ideal}}} x. \quad (2.55)$$

Definimos, finalmente, el error sistemático relativo  $\varepsilon_{\text{rel}}$  como

$$\varepsilon_{\text{rel}} = \frac{\mu_o}{x} - 1. \quad (2.56)$$

Como se aprecia, representa el error relativo entre las constantes real e ideal. El error relativo es adimensional y es 0 en el caso ideal en que fuera  $K_{\text{real}} = K_{\text{ideal}}$ . Para mediciones aceptables debería ser mucho menor que 1.

Dado que  $\mu_o$  es la media extendida a toda la población de posibles observaciones de  $x$ , no es posible conocerla. A efectos de estimar  $\varepsilon_{\text{rel}}$  elegimos un valor de referencia de calibración  $x_{\text{cal}}$  y promediamos  $m$  observaciones de  $x$ . Se supone que  $x_{\text{cal}}$  y su incertidumbre son conocidas. Se obtiene

$$\hat{\varepsilon}_{\text{rel}} = \frac{\bar{x}_{\text{cal}, o}}{x_{\text{cal}}} - 1. \quad (2.57)$$

La incertidumbre estándar combinada se calcula mediante la ecuación (2.26) y resulta

$$\begin{aligned}
u_{\hat{\epsilon}_{\text{rel}}} &= \sqrt{u_{\bar{x}_{\text{cal},o}}^2 \left( \frac{1}{x_{\text{cal}}} \right)^2 + u_{x_{\text{cal}}}^2 \left( \frac{-\bar{x}_{\text{cal},o}}{x_{\text{cal}}^2} \right)^2} = \\
&= \frac{1}{x_{\text{cal}}} \sqrt{u_{\bar{x}_{\text{cal},o}}^2 + u_{x_{\text{cal}}}^2 (1 + \hat{\epsilon}_{\text{rel}})^2}
\end{aligned} \tag{2.58}$$

Ahora que disponemos del valor estimado del error sistemático relativo, se puede calcular el valor corregido de un mensurando  $x$  genérico:

$$x_c = \bar{x}_o (1 + \hat{\epsilon}_{\text{rel}}). \tag{2.59}$$

La incertidumbre estándar combinada para esta medición resulta

$$u_{x_c} = \sqrt{u_{\bar{x}_o}^2 (1 + \hat{\epsilon}_{\text{rel}})^2 + \bar{x}_o^2 u_{x_{\text{cal}}}^2} \tag{2.60}$$

O, poniendo en evidencia la reducción de incertidumbre de una media respecto a una medición individual,

$$u_{x_c} = \sqrt{\frac{u_{x_o}^2}{n} (1 + \hat{\epsilon}_{\text{rel}})^2 + \bar{x}_o^2 u_{x_{\text{cal}}}^2}. \tag{2.61}$$

Nótese que, como

$$\bar{x}_o = \frac{x_c}{1 + \hat{\epsilon}_{\text{rel}}}, \tag{2.62}$$

la incertidumbre en  $x_c$  crece con  $x_c$ , no así el error sistemático relativo.

**EJEMPLO:** Consideremos la medición de una presión sonora mediante un micrófono cuya sensibilidad  $S$  difiere de la nominal. En ese caso, si  $v$  es la tensión generada por el micrófono ante una presión  $p$ , tendremos<sup>29</sup>

$$p = \frac{1}{S} v. \tag{2.63}$$

Supongamos un micrófono Rion UC-53A, cuya sensibilidad nominal (ideal) es  $-28,0$  dB referidos a  $1$  V/Pa, y cuyo certificado de calibración indica que la sensibilidad real (a  $1$  kHz) es  $-26,2$  dB con una incertidumbre estándar de  $0,2$  dB. Nos interesa obtener el error sistemático relativo. Para ello necesitamos obtener las sensibilidades ideal y real, y su incertidumbre, en V/Pa

<sup>29</sup> Es interesante notar que en este caso estamos interesados en mediciones de presión sonora en lugar de nivel de presión sonora.



$$S_{\text{ideal}} = 10^{\frac{S_{\text{ideal,dB}}}{20}} = 10^{\frac{-28,0}{20}} \text{ V/Pa} = 0,03981 \text{ V/Pa}$$

La sensibilidad real, a su vez, es

$$S_{\text{real}} = 10^{\frac{S_{\text{real,dB}}}{20}} = 10^{\frac{-26,2}{20}} \text{ V/Pa} = 0,04898 \text{ V/Pa}$$

La incertidumbre estándar se obtiene aplicando la ecuación (2.26), en este caso con una sola variable:

$$u_{S_{\text{real}}} = \frac{\partial S_{\text{real}}}{\partial S_{\text{real,dB}}} u_{S_{\text{real,dB}}} = \frac{\ln 10}{20} S_{\text{real}} u_{S_{\text{real,dB}}} = 0,0011 \text{ V/Pa}$$

Si suponemos que el micrófono genera una tensión  $v$ , el valor, la presión sonora observada será

$$p_o = \frac{v}{S_{\text{ideal}}},$$

mientras que la presión real es

$$p = \frac{v}{S_{\text{real}}}.$$

El error relativo estimado será

$$\hat{\varepsilon}_{\text{rel}} = \frac{p_o}{p} - 1 = \frac{S_{\text{real}}}{S_{\text{ideal}}} - 1 = 0,2303.$$

La incertidumbre estándar de este error es

$$u_{\hat{\varepsilon}_{\text{rel}}} = \frac{u_{S_{\text{real}}}}{S_{\text{ideal}}} = 0,0276.$$

Si, por ejemplo, midiéramos una presión  $p_o = 1 \text{ Pa}$ , el valor corregido sería

$$p_c = 1 \text{ Pa} \times (1 + 0,2303) = 1,2303 \text{ Pa}$$

La componente de incertidumbre debida al error relativo será

$$u_{p_c, \text{rel}} = 1 \text{ Pa} \times 0,0276 = 0,0276 \text{ Pa}$$

La incertidumbre total deberá incluir, además, la incertidumbre debida a efectos aleatorios (no sistemáticos), heredada entre otras cosas de la incertidumbre en la medición de  $v$ .

NOTA: El error sistemático multiplicativo se convierte en aditivo cuando se usa una escala logarítmica, ya que el error en el factor de escala se convierte en un término aditivo.

### 2.8.3 Error sistemático no lineal

El error sistemático no lineal o error de linealidad se presenta en situaciones donde el modelo supuesto para la medición es lineal:

$$y = Kx \quad (2.64)$$

pero en realidad es no lineal y posee dispersión:

$$y = g(x) + \varepsilon_y. \quad (2.65)$$

En estas ecuaciones  $y$  es una cantidad que se puede medir directamente y  $x$  el mensurando a obtener a partir de  $y$ . Cuando se determina  $x$  en general se supone que vale (2.64) con constante nominal o ideal  $K_0$ , por lo cual el valor observado,  $x_0$  será

$$x_0 = \frac{y}{K_0} = \frac{g(x)}{K_0} + \frac{\varepsilon_y}{K_0}. \quad (2.66)$$

Supongamos, por ejemplo, que la función  $g(x)$  puede aproximarse por un polinomio de tercer grado sin término independiente

$$y = Kx + K_1x^2 + K_2x^3. \quad (2.67)$$

Para una ecuación cúbica como ésta es posible despejar  $x$  en función de  $y$ , pero para el caso de mayor interés donde  $K_1$  y  $K_2$  son pequeños (una relación prácticamente lineal), podemos aproximar

$$x \cong Ay + By^2 + Cy^3. \quad (2.68)$$

Si bien la cantidad realmente medida es  $y$ , el instrumento en general proporciona un valor observado  $x_0$  del mensurando, relacionado con  $y$  según la relación proporcional de la ecuación (2.66). El valor verdadero de  $x$  puede, entonces, aproximarse como:

$$x \cong ax_0 + bx_0^2 + cx_0^3. \quad (2.69)$$

Si conociéramos  $K$ ,  $K_1$  y  $K_2$ , podríamos obtener  $a$ ,  $b$  y  $c$  del siguiente modo. Primero se reemplaza  $x_0$  por  $y/K_0$  según (2.66) e  $y$  en función de  $x$  mediante (2.67). Se obtiene en el segundo miembro de (2.69) un polinomio de grado 9. Reteniendo sólo los términos hasta el cúbico e igualando coeficientes se llega a la siguiente corrección sobre el valor leído  $x_0$ :

$$x_c = \frac{K_0}{K}x_0 - \frac{K_1K_0^2}{K^3}x_0^2 + \frac{2K_1^2 - K_2K}{K^5}K_0^3x_0^3. \quad (2.70)$$

Nótese que se reemplazó el signo  $\cong$  por una igualdad. Ello se debe a que no afirmamos que el valor  $x_c$  sea igual a  $x$ ; sólo es un valor corregido sobre el valor leído  $x_o$  que, al tener en cuenta el efecto sistemático no lineal se aproxima mejor al valor real.

En este caso no se introduce una medida específica de error sistemático ya que se requerirían varios parámetros, aunque todavía es posible utilizar el concepto de error sistemático relativo en relación con el término lineal.

En general, los valores de  $K$ ,  $K_1$  y  $K_2$  no son conocidos, por lo que el modelo deberá obtenerse empíricamente. En lugar del modelo (2.67) usaremos el (2.69), ya que nos da el valor corregido directamente a partir del valor leído. Para determinar  $a$ ,  $b$  y  $c$  hace falta medir tres valores de referencia distintos de valores convencionales  $x_1, x_2, x_3$  e incertidumbres  $u_{x_1}, u_{x_2}, u_{x_3}$  conocidos.<sup>30</sup> Llamando  $x_{1o}, x_{2o}, x_{3o}$  a los valores medidos u observados (o las respectivas medias de  $m$  mediciones), se obtiene el sistema

$$\begin{cases} ax_{1o} + bx_{1o}^2 + cx_{1o}^3 = x_1 \\ ax_{2o} + bx_{2o}^2 + cx_{2o}^3 = x_2 \\ ax_{3o} + bx_{3o}^2 + cx_{3o}^3 = x_3 \end{cases} \quad (2.71)$$

La solución se obtiene numéricamente mediante software matemático matricial.

Este sistema de 3 ecuaciones con 3 incógnitas se resuelve numéricamente, para lo cual es preferible presentarlo matricialmente

$$\begin{bmatrix} x_{1o} & x_{1o}^2 & x_{1o}^3 \\ x_{2o} & x_{2o}^2 & x_{2o}^3 \\ x_{3o} & x_{3o}^2 & x_{3o}^3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad (2.72)$$

de donde

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_{1o} & x_{1o}^2 & x_{1o}^3 \\ x_{2o} & x_{2o}^2 & x_{2o}^3 \\ x_{3o} & x_{3o}^2 & x_{3o}^3 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \quad (2.73)$$

Esto completa los datos para aplicar la corrección de los efectos sistemáticos no lineales:

$$x_c \cong ax_o + bx_o^2 + cx_o^3. \quad (2.74)$$

Calcularemos ahora la incertidumbre. Para ello tengamos en cuenta que si bien una vez determinadas las constantes queda definido un modelo que puede aplicarse a muchos datos observados  $x_o$ , dicho modelo no es el único que podríamos haber obtenido siguiendo el mismo procedimiento, ya que los valores medidos de las referencias están sujetos a incertidumbre. Por lo tanto, en realidad

<sup>30</sup> Nótese que se conocen los valores convencionales, no los valores reales. Sin embargo dichos valores convencionales tienen menor dispersión y no poseen efectos sistemáticos conocidos

$$x_c = g(x_0, x_{10}, x_{20}, x_{30}). \quad (2.75)$$

La dependencia funcional de  $x_{10}$ ,  $x_{20}$  y  $x_{30}$  es a través de los coeficientes  $a$ ,  $b$  y  $c$ :

$$x_c \cong a(x_{10}, x_{20}, x_{30})x_0 + b(x_{10}, x_{20}, x_{30})x_0^2 + c(x_{10}, x_{20}, x_{30})x_0^3. \quad (2.76)$$

La incertidumbre es, entonces

$$u_{x_c}^2 = \left( \frac{\partial g}{\partial x_0} \right)^2 u_{x_0}^2 + \left( \frac{\partial g}{\partial x_{10}} \right)^2 u_{x_{10}}^2 + \left( \frac{\partial g}{\partial x_{20}} \right)^2 u_{x_{20}}^2 + \left( \frac{\partial g}{\partial x_{30}} \right)^2 u_{x_{30}}^2 \quad (2.77)$$

Las incertidumbres básicas son las correspondientes a la medición y podrían depender del valor del mensurando en cada caso. Las derivadas son:

$$\frac{\partial g}{\partial x_0} = a + 2bx_0 + 3cx_0^2 \quad (2.78)$$

$$\frac{\partial g}{\partial x_{i0}} = \frac{\partial a}{\partial x_{i0}}x_0 + \frac{\partial b}{\partial x_{i0}}x_0^2 + \frac{\partial c}{\partial x_{i0}}x_0^3 \quad (2.79)$$

donde  $i = 1, 2, 3$ . Las funciones  $a$ ,  $b$ ,  $c$  son funciones definidas implícitamente a través del sistema (2.71). Si bien es posible obtener una expresión analítica, es preferible la derivación implícita. Para ello reescribimos (2.71) así:

$$\begin{cases} ax_{10} + bx_{10}^2 + cx_{10}^3 - x_1 = 0 \\ ax_{20} + bx_{20}^2 + cx_{20}^3 - x_2 = 0 \\ ax_{30} + bx_{30}^2 + cx_{30}^3 - x_3 = 0 \end{cases} \quad (2.80)$$

Si derivamos cada una de estas expresiones con respecto a  $x_{10}$  obtenemos

$$\begin{cases} \frac{\partial a}{\partial x_{10}}x_{10} + \frac{\partial b}{\partial x_{10}}x_{10}^2 + \frac{\partial c}{\partial x_{10}}x_{10}^3 = -(a + 2bx_{10} + 3cx_{10}^2) \\ \frac{\partial a}{\partial x_{10}}x_{20} + \frac{\partial b}{\partial x_{10}}x_{20}^2 + \frac{\partial c}{\partial x_{10}}x_{20}^3 = 0 \\ \frac{\partial a}{\partial x_{10}}x_{30} + \frac{\partial b}{\partial x_{10}}x_{30}^2 + \frac{\partial c}{\partial x_{10}}x_{30}^3 = 0 \end{cases} \quad (2.81)$$

Este es un sistema de tres ecuaciones con tres incógnitas que permite despejar las derivadas de  $a$ ,  $b$  y  $c$  respecto a  $x_{10}$ . En versión matricial:

$$\begin{bmatrix} x_{10} & x_{10}^2 & x_{10}^3 \\ x_{20} & x_{20}^2 & x_{20}^3 \\ x_{30} & x_{30}^2 & x_{30}^3 \end{bmatrix} \begin{bmatrix} \partial a / \partial x_{10} \\ \partial b / \partial x_{10} \\ \partial c / \partial x_{10} \end{bmatrix} = - \begin{bmatrix} a + 2bx_{10} + 3cx_{10}^2 \\ 0 \\ 0 \end{bmatrix} \quad (2.82)$$

Vemos que la matriz del sistema de ecuaciones es la misma del sistema (2.72) planteado para despejar  $a$ ,  $b$  y  $c$ , lo cual es ventajoso porque basta calcular una vez la inversa.

Procediendo análogamente obtenemos y resolvemos sistemas para las derivadas respecto a  $x_{20}$  y  $x_{30}$ . Con todos los valores obtenidos de (2.82) y sus análogos reemplazados en (2.79) y similares, sustituimos (2.78) y (2.79) en (2.77), obteniendo finalmente la incertidumbre buscada. El cálculo es bastante largo y por lo tanto se realiza por software. En el apéndice 9 se desarrolla un ejemplo.

#### 2.8.4 Errores sistemáticos no corregidos

Hay casos en donde existe un error sistemático que no es posible corregir. Un caso típico es cuando un instrumento se calibra con un calibrador cuyo valor real no coincide con el nominal. Esta situación se maneja teniendo en cuenta la incertidumbre de la calibración.

Otra situación se da cuando el calibrador experimenta derivas a largo plazo. Por ejemplo, el calibrador Brüel & Kjaer Type 4231 indica en sus especificaciones que la estabilidad a un plazo de un año es mejor que 0,05 dB con una confiabilidad del 96 %. Podemos suponer que ello implica un factor de cobertura  $k = 2$ . Por consiguiente, dentro del plazo de un año la incertidumbre estándar será de 0,025 dB adicional a la incertidumbre estándar propia del nivel de referencia del calibrador, por lo cual la incertidumbre debida al error sistemático no corregido será

$$u_{\text{cal}} = \sqrt{u_{\text{inicial}}^2 + u_{1 \text{ año}}^2} \quad (2.83)$$

Con los valores del ejemplo de 2.8.1 resulta  $u_{\text{cal}} = 0,152$  dB, es decir, un incremento despreciable frente a la incertidumbre de 0,15 dB indicada en el certificado de calibración más reciente.

#### 2.9 Cálculos de incertidumbre en cadena

En cálculos complejos de incertidumbre suele suceder que el problema se puede descomponer en una serie de cálculos encadenados a través de composición de funciones anidadas. Supongamos, por ejemplo, que tenemos un mensurando que depende de varias variables que a su vez dependen de otras variables:

$$w = f(z_1, \dots, z_L),$$

$$z_k = z_k(y_1, \dots, y_M),$$

$$y_j = y_j(x_1, \dots, x_N).$$

Entonces

$$\begin{aligned} u_w^2 &= \sum_{i=1}^L \left( \frac{\partial w}{\partial x_i} \right)^2 u_{x_i}^2 = \sum_{i=1}^L \left( \sum_{j=1}^M \frac{\partial w}{\partial y_j} \frac{\partial y_j}{\partial x_i} \right)^2 u_{x_i}^2 = \\ &= \sum_{i=1}^L \left( \sum_{j=1}^M \left( \sum_{k=1}^N \frac{\partial f}{\partial z_k} \frac{\partial z_k}{\partial y_j} \right) \frac{\partial y_j}{\partial x_i} \right)^2 u_{x_i}^2 \end{aligned}$$

Es decir,

$$u_w^2 = \sum_{i=1}^L \left( \sum_{j=1}^M \sum_{k=1}^N \frac{\partial f}{\partial z_k} \frac{\partial z_k}{\partial y_j} \frac{\partial y_j}{\partial x_i} \right)^2 u_{x_i}^2$$

A pesar de que las  $x_i$  sean independientes las  $y_j$  y las  $z_k$  no necesariamente lo serán, ya que dependen funcionalmente de unas mismas variables. Es importante, en casos como éste, procurar que las variables intermedias dependan de variables independientes diferentes, lo cual a veces se logra agrupando en una sola función todas aquellas variables que dependan de una misma variable

## Capítulo 3

# REGISTRO DIGITAL

### 3.1 Introducción

Las mediciones acústicas basadas en software requieren que, previamente, una señal acústica haya sido convertida en eléctrica mediante un transductor calibrado y luego transformada en audio digital mediante un conversor analógico/digital. Sólo entonces es posible aplicar algoritmos que completen el proceso de medición. En lo que sigue se introducen los conceptos fundamentales del audio digital.

### 3.2 Audio digital

Una señal digital es, en primer lugar, una señal representada con números en lugar de valores continuos de una magnitud física como la presión sonora o la tensión eléctrica. Mientras una señal continua puede adoptar infinitos valores posibles, una señal digital sólo puede adoptar una cantidad finita de valores, según la cantidad de dígitos disponibles. Pero además, la señal original toma valores en una cantidad infinita de instantes de tiempo, por lo tanto si pretendemos que la señal digital pueda ser procesada o almacenada necesitamos hacer un *muestreo*, es decir, obtener el valor de la señal en algunos instantes de tiempo, pues de lo contrario deberíamos manipular infinitos números. La señal digital es, entonces, una señal de *valores discretos en tiempo discreto*. Hay, por lo tanto, dos discretizaciones: una en el tiempo y otra en la magnitud.

Las señales digitales tienen varias ventajas con respecto a las señales analógicas que las hacen atractivas en muchas aplicaciones, particularmente las metrológicas:

- 1) Si se utiliza un medio de almacenamiento apropiado, la señal es prácticamente inalterable. Esto contrasta con los medios de almacenamiento magnético como los registros fonográficos o magnéticos.
- 2) Es posible hacer tantas copias de respaldo como se desee, incluso sucesivas copias de copias, todas idénticas al original. En el caso de señales analógicas cada copia supone una degradación por ruido y distorsión.
- 3) La señal digital puede ser transmitida tanto por cable como inalámbricamente sin sufrir alteraciones por la existencia de ruido o distorsiones moderadas en el canal de transmisión, a diferencia del caso analógico.
- 4) En caso de canales ruidosos o medios delicados se pueden implementar algoritmos de corrección de errores basados en la redundancia. En medios analógicos la redundancia sólo permite reducir, pero no anular, las alteraciones de la señal.
- 5) Es posible implementar una gran variedad de procesos digitales tales como filtros, reducción de ruido, análisis de espectro, análisis estadístico, integraciones, convolución, postprocesamientos, marcado, presentación gráfica, aplicación de criterios, etc.

Como contrapartida, el hardware de base requerido para cualquiera de estos procesos es generalmente más sofisticado, requiriendo filtros, conversores, procesadores digitales de señal, microprocesadores o microcontroladores. La discretización en tiempo impo-

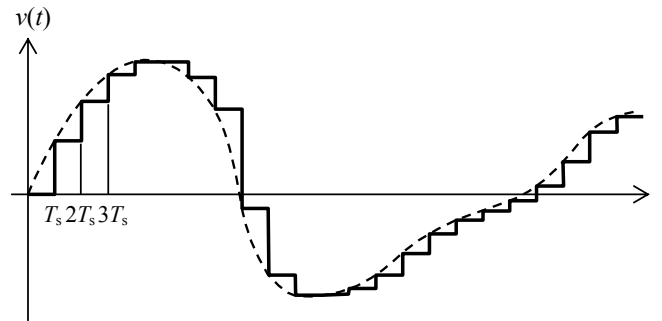
ne, asimismo, una severa limitación del espectro en frecuencias elevadas. Con todo, las ventajas superan ampliamente a las desventajas.

Comenzaremos analizando el proceso de digitalización en sus de dos etapas básicas: el muestreo y la conversión analógico/digital.

### 3.3 Muestreo

El proceso de muestreo consiste en retener el valor de la señal en el instante en que se desea muestrear. Ello se logra mediante circuitos de *muestreo y retención* (en inglés, *sample and hold*). Este paso es necesario porque los conversores analógico/digitales requieren cierto tiempo para completar la conversión, y durante ese tiempo la señal a convertir debe permanecer constante.

La retención se logra conectando durante un breve lapso un condensador a la fuente de señal (en nuestro caso, la salida de un micrófono previamente amplificada) y luego desconectándolo. El condensador retiene su carga por un tiempo, permitiendo la conversión. Este proceso se repite cada un tiempo  $T_s$  llamado *periodo de muestreo*. En la figura 27 se muestra el resultado del muestreo.



**Figura 27.** Ejemplo de muestreo y retención de una señal. En cada instante  $kT$  se actualiza el valor de la muestra y el nuevo valor se retiene hasta  $(k + 1)T$ .

Una cuestión muy importante en el muestreo es cuántas muestras se deben tomar por unidad de tiempo, es decir, la *tasa de muestreo*,  $F_s$  (también llamada *frecuencia de muestreo*), dada por

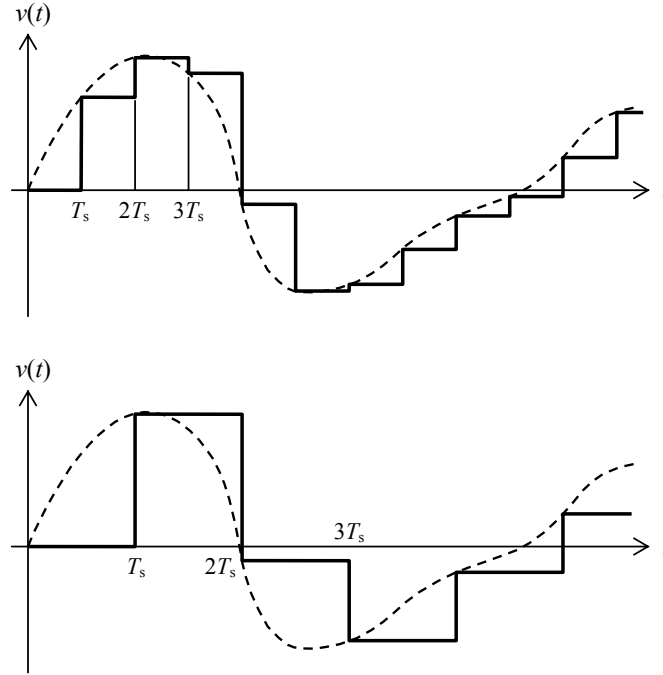
$$F_s = \frac{1}{T_s}. \quad (3.1)$$

Tanto desde el punto de vista del espacio de memoria requerido para almacenar la señal como del de la capacidad de cómputo del sistema que luego procesará la señal digital (cantidad de operaciones por unidad de tiempo), convendría que  $F_s$  fuera pequeña. Pero al mismo tiempo un valor demasiado pequeño haría que se pierdan detalles importantes de la señal. En la figura 28 se muestra la misma señal de la figura 27 pero muestreada a menores tasas de muestreo. Se aprecia que cuanto menor sea la tasa de muestreo menor será la cantidad de detalles que las muestras pueden representar.

Se demuestra que si se adopta una tasa de muestreo mayor o igual que el doble de la máxima frecuencia contenida en la señal continua, es decir

$$F_s \geq 2f_{\text{máx}} \quad (3.2)$$





**Figura 28.** La misma señal de la figura 27 muestreada a tasas de muestreo menores. El resultado muestra cada vez menor fidelidad con respecto al original.

entonces es posible reconstruir perfectamente la señal a partir de las muestras obtenidas, lo cual implica que se pueden representar todos los detalles contenidos en la señal, particularmente todos aquellos de interés metrológico. Este importante resultado se conoce como *teorema del muestreo* (ver apéndice 10) y la desigualdad (3.2) como *condición de Nyquist*. En la figura 29 se muestra una señal correctamente muestreada ( $F_s = 14,7f_{\text{máx}}$ ) y la reconstrucción a partir de las muestras. Como se aprecia, la señal reconstruida coincide con la original.

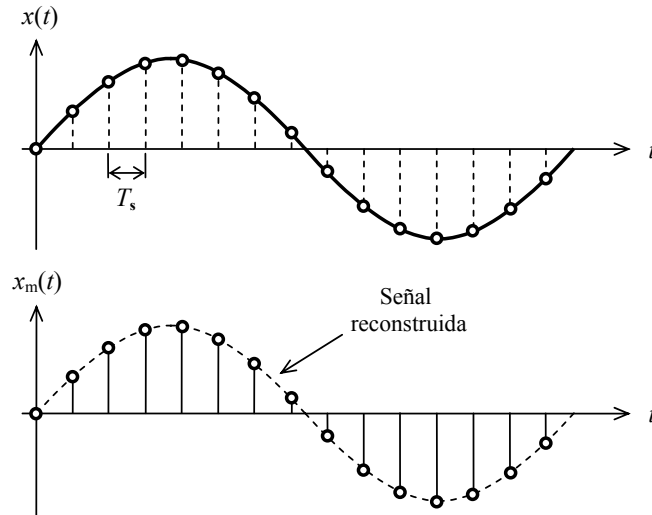
La relación (3.2) puede interpretarse también como que el muestreo será correcto si la frecuencia máxima satisface

$$f_{\text{máx}} \leq \frac{F_s}{2}. \quad (3.3)$$

La frecuencia  $F_s/2$  se denomina *frecuencia de Nyquist*.

Si la señal no satisface la desigualdad (3.3) se produce un fenómeno denominado *aliasing*, consistente en la aparición de frecuencias inexistentes al intentar reconstruir la señal, denominadas *frecuencias alias*. Estas frecuencias equivalen a reflexiones alrededor de la frecuencia  $F_s/2$  de todas las frecuencias que excedan dicha frecuencia. Así, dada una frecuencia  $f_0$  tal que

$$\frac{F_s}{2} < f_0 < F_s, \quad (3.4)$$

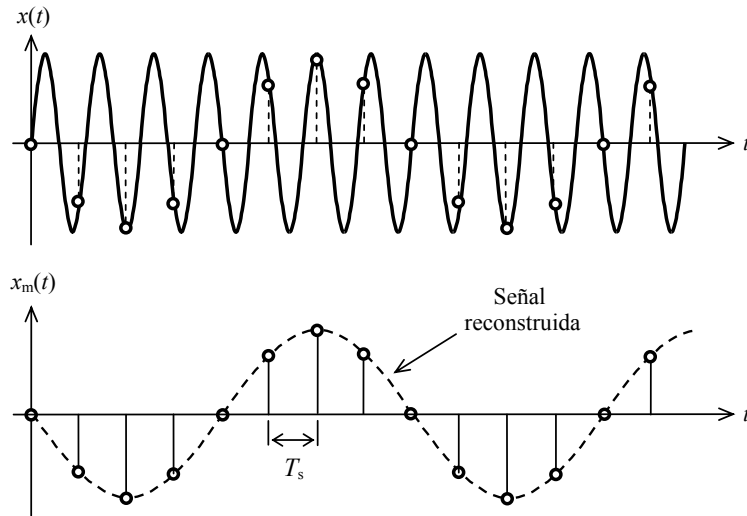


**Figura 29.** Reconstrucción de una señal correctamente muestreada. La señal reconstruida es igual a la original.

entonces aparecerá una frecuencia alias

$$f_{\text{alias}} = \frac{F_s}{2} - \left( f_0 - \frac{F_s}{2} \right) = F_s - f_0 \quad (3.5)$$

El fenómeno se ilustra en la figura 30, donde una señal de  $f_0 = 35$  kHz se ha muestreado con una tasa de muestreo  $F_s = 40$  kHz. La reconstrucción resulta de 5 kHz.



**Figura 30.** Efecto del muestreo con una frecuencia menor que el doble de la máxima frecuencia contenida en la señal. Una señal de 35 kHz se muestrea con una frecuencia de 40 kHz; al intentar reconstruirla, aparece una frecuencia *alias* de 5 kHz.

Las frecuencias superiores a  $F_s$  se reflejan en  $F_s/2$ , luego en 0, luego en  $F_s/2$ , y así sucesivamente, de modo que la frecuencia alias se encuentre siempre en el intervalo entre 0 y  $F_s/2$ . Si

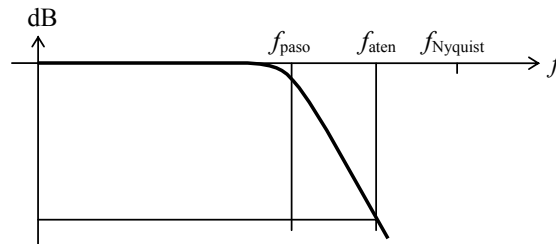
$$\begin{aligned} F_o &= f_o \bmod (F_s/2) \\ k &= \frac{f_o}{F_s/2} \end{aligned} \quad (3.6)$$

donde mod es la operación de módulo, consistente en el resto de la división entera. Entonces podemos escribir,

$$f_{\text{alias}} = \begin{cases} \frac{F_s}{2} - F_o & k \text{ impar} \\ F_o & k \text{ par} \end{cases} \quad (3.7)$$

Debido a esto, es importante comprender que la frecuencia máxima que aparece en la desigualdad (3.3) no se refiere sólo a la máxima frecuencia *de interés*, sino en realidad a la máxima frecuencia que *efectivamente* aparece en la señal a muestrear, aunque dicha frecuencia provenga de un ruido inaudible de alta frecuencia que contamina la señal. En caso de no cumplirse el criterio, al intentar recuperar la señal aparecerán componentes de frecuencia espurias en la banda útil. Un caso típico en que esto podría suceder es cuando la señal de audio está contaminada con ruido eléctrico de banda ancha por, ejemplo ruido térmico proveniente del micrófono, del amplificador o de cualquier componente que se intercale antes del muestreo.

Por esa razón, como parte del acondicionamiento de la señal previo al muestreo se debe utilizar un filtro que bloquee virtualmente todo contenido espectral más allá de la frecuencia de Nyquist,  $F_s/2$ . Dicho filtro, denominado *filtro antialias* (ver figura 31), debe atenuar el ruido por encima de  $F_s/2$  lo suficiente como para hacerlo irrelevante comparado con otras fuentes de incertidumbre.



**Figura 31.** Respuesta en frecuencia de un filtro pasabajos antialias. La frecuencia de atenuación debe ser menor que la frecuencia de Nyquist.

Aunque el filtro antialias puede parecer un componente relativamente simple del sistema, en la práctica suele ser el eslabón que mayor distorsión introduce en un proceso

de medición basado en software. En efecto, para ahorrar espacio de almacenamiento de los datos digitales asociados a una señal la tasa de muestreo se suele tomar apenas mayor que lo estrictamente necesario para cumplir la condición de Nyquist. Por ejemplo, para señales de audio, cuyo límite relevante está en 20 kHz, se utiliza normalmente una frecuencia de 44,1 kHz.<sup>31</sup> Esto implica una frecuencia de Nyquist de 22,05 kHz, lo cual deja un exiguo margen de 2,05 kHz donde el filtro debe efectuar una rápida transición entre ninguna atenuación y una atenuación elevada (típicamente más de 90 dB). Para ello se requieren filtros de elevado orden, por ejemplo un filtro elíptico de orden 13 ó uno de Chebychev de orden 29. Estos filtros poseen una respuesta transitoria muy marcada, lo cual puede ocasionar distorsiones en señales fuertemente aleatorias. A ello se agrega que algunas celdas constitutivas de estos filtros son de elevado  $Q$ , lo cual se traduce en elevados picos de resonancia, imponiendo limitaciones del rango dinámico antes de llegar a la saturación.

El problema anterior se puede encarar de dos maneras. La primera, elevando la tasa de muestreo a 48 kHz o incluso 96 kHz, a costa de un mayor requisito de memoria. La segunda, trabajar con técnicas de *sobremuestreo* (oversampling),<sup>32</sup> que al utilizar tasas de muestreo mucho más altas reducen enormemente la exigencia de los filtros.

### 3.4 Digitalización

Una vez muestreada, la señal discreta se convierte al dominio digital. Para ello se utiliza un conversor analógico/digital cuya función es asignar a cada valor analógico muestreado un número proporcional a dicho valor a través de una constante de calibración. Si bien podría utilizarse cualquier base de numeración, todos los conversores conocidos emplean numeración binaria, es decir, en base 2, debido a que permite utilizar circuitos de lógica binaria con dos niveles de tensión, 0 y  $V_{cc}$ , o aplicar el concepto de llave abierta y cerrada, que corresponden a los dígitos binarios 0 y 1. Esto permite la mayor inmunidad posible frente al ruido, ya que es necesario un ruido superior a  $V_{cc}/2$  para que un nivel pueda ser confundido con otro.<sup>33</sup> Se denomina *bit* (contracción de *binary digit* en inglés) a cada dígito binario, y *palabra digital* a un grupo cualquiera de bits (aunque más específicamente suele contener un múltiplo de 8 bits).

La conversión se realiza fijando primero un valor de referencia analógico  $V_{ref}$  y estableciendo una subdivisión de dicha referencia en  $2^n$  escalones de amplitud  $V_{ref}/2^n$ , donde  $n$  es la cantidad de bits o *resolución* del dato digital. Para señales unidireccionales se hace corresponder el 0 digital (es decir, una palabra formada por  $n$  ceros) al 0 analógico, y el máximo digital,  $2^n - 1$  (es decir  $n$  ceros) a  $V_{ref}(1 - 1/2^n)$ . Para señales bidi-

<sup>31</sup> El valor preciso de 44,1 kHz en realidad obedece a que las primeras grabadoras digitales en cinta magnética aprovecharon la tecnología de cabezal rotativo que se venía utilizando en el video, ubicaban tres muestras de 12 bit en cada línea de barrido horizontal mediante un adaptador PCM (pulse code modulation). Había 294 líneas activas a 50 Hz (barrido entrelazado equivalente a 588 líneas a 25 Hz).

<sup>32</sup> Consiste en realizar el muestreo a una tasa mucho más elevada (por ejemplo 128 veces mayor) pero convertir cada muestra con baja resolución (típicamente 1 bit). En lugar de representar el valor completo de la señal, se representa la diferencia entre muestras sucesivas. Por la altísima tasa de muestreo el cambio entre dos muestras sucesivas será muy pequeño, de allí que se lo pueda representar con 1 bit. Luego se calcula por integración numérica el valor completo y se procede a un submuestreo.

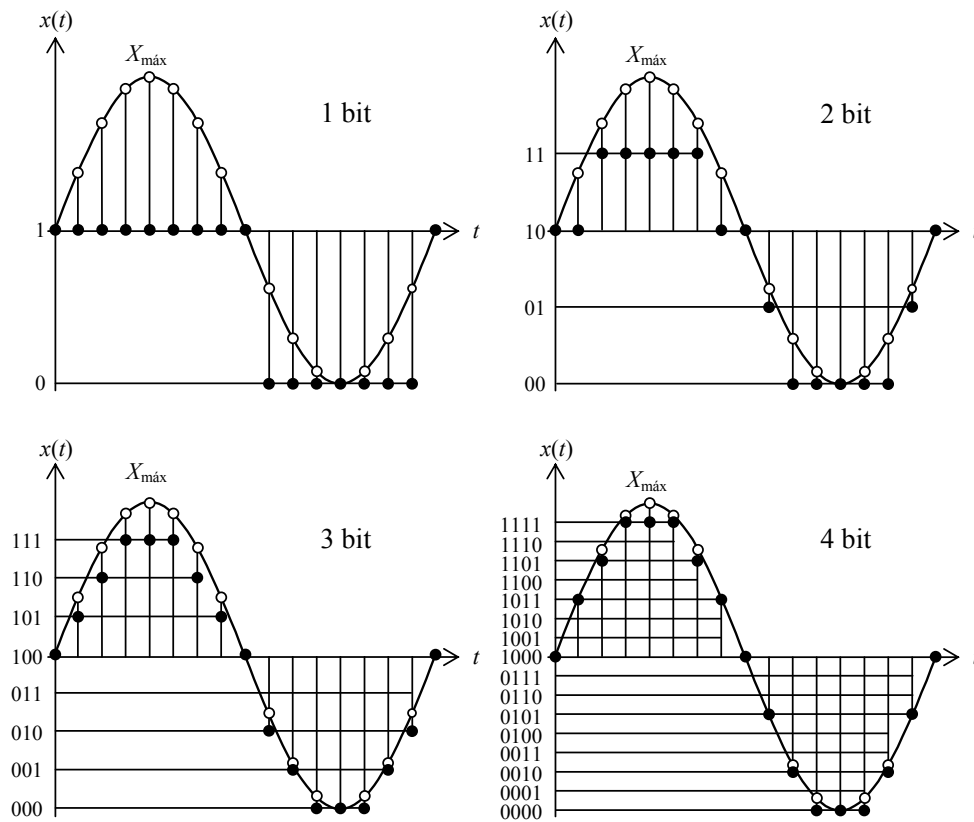
<sup>33</sup> Desde el punto de vista de la Teoría de la Información, se puede demostrar que el código más eficiente, es decir el que maximiza la cantidad de información por dígito, es un código ideal de base  $e = 2,71828...$ , por lo que el código de base entera más eficiente es el ternario (base 3). ¿Debería sorprendernos que la Naturaleza haya seleccionado evolutivamente un código ternario para almacenar y transmitir la información genética...?

reccionales, el 0 digital corresponde a  $-V_{\text{ref}}/2$  y el máximo digital a  $V_{\text{ref}}(1/2 - 1/2^n)$ . En este caso el 0 analógico corresponde a  $2^{n-1}$ , es decir, un 1 seguido por  $n - 1$  ceros.

En la figura 32 se muestra la conversión analógico/digital de una senoide de máxima amplitud ( $V_{\text{ref}}/2$ ) para resoluciones de 1 bit, 2 bit, 3 bit y 4 bit. En este caso se adopta para cada valor analógico el máximo escalón que no lo supera, es decir:

$$D = \begin{cases} 0 & x < -V_{\text{ref}}/2 \\ \left[ \frac{x - (-V_{\text{ref}}/2)}{V_{\text{ref}}} 2^n \right] & -V_{\text{ref}}/2 \leq x < V_{\text{ref}}/2 \\ 2^n - 1 & V_{\text{ref}}/2 \leq x \end{cases} \quad (3.8)$$

donde  $D$  es el dato digital y  $[z]$  representa la parte entera de  $z$ . Las definiciones para  $x < V_{\text{ref}}/2$  y para  $V_{\text{ref}}/2 \leq x$  corresponden a efectos de saturación. Vemos que a mayor cantidad de bits más fielmente queda representada cada muestra analógica por su versión digital



**Figura 32.** Ejemplos de conversión analógico/digital con resoluciones de 1 bit, 2 bit, 3 bit y 4 bit. Los círculos vacíos representan el valor exacto de cada muestra y los círculos llenos el valor digital correspondiente. En este ejemplo se tomó como criterio el mayor escalón que no supera al valor de la muestra.

Se logra otra posible codificación que respeta mejor la simetría de la onda eligiendo el 0 analógico a mitad de camino entre los valores digitales  $2^{n-1} - 1$  y  $2^{n-1}$ . Esta solución, no obstante, no se suele utilizar pues es de implementación más compleja y además no permite representar correctamente una señal nula.

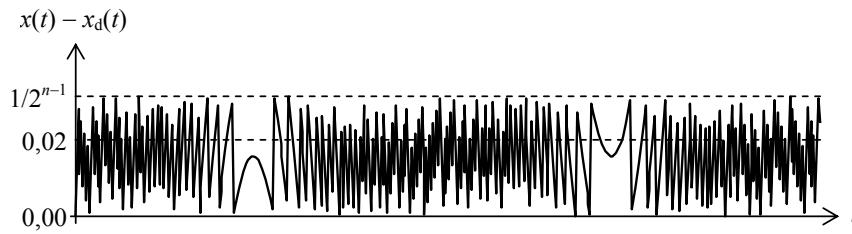
La falta de simetría, que es muy notoria con representaciones de pocos bits, se vuelve irrelevante cuando se utilizan resoluciones más altas, por ejemplo, 16 bit.

### 3.5 Relación señal a ruido

Cuando la amplitud de la señal es relativamente grande comparada con el salto correspondiente al bit menos significativo, es decir, 1 LSB (por sus siglas en inglés, *Least Significant Bit*), la diferencia entre la señal original  $x(t)$  y la señal digitalizada  $x_d(t)$  es una variable cuasi aleatoria que se comporta como un ruido:

$$\varepsilon(t) = x(t) - x_d(t). \quad (3.9)$$

En la figura 33 se muestra un ejemplo en el que se ha muestreado una señal senoidal de máxima amplitud con una resolución de 6 bit.



**Figura 33.** Diferencia o error entre una señal senoidal de amplitud 1 y su versión digitalizada con una resolución  $n = 6$  bit. NOTA: La tasa de muestreo era mucho más alta que la frecuencia de la señal.

Las porciones cercanas al pico forman un patrón distintivo que, al aumentar la resolución  $n$ , va ocupando cada vez menor proporción de la onda. Por consiguiente, a medida que aumenta la resolución la señal residual se va pareciendo progresivamente a un ruido con distribución uniforme entre 0 y  $V_{\text{máx}}/2^{n-1}$ . Su valor eficaz coincide con su desvío estándar y vale (ver apéndice 2)

$$E_{\text{ef}} = \frac{V_{\text{máx}}/2^{n-1}}{\sqrt{12}}. \quad (3.10)$$

El valor eficaz de una señal senoidal de máxima amplitud será

$$V_{\text{ef}} = \frac{V_{\text{máx}}}{\sqrt{2}}, \quad (3.11)$$

de donde la relación señal a ruido resulta

$$S/R|_{\text{dB}} = 20 \log \frac{V_{\text{ef}}}{E_{\text{ef}}} = 20 \log \sqrt{\frac{3}{2}} + 20 \log 2^n$$

es decir,

$$S/R|_{\text{dB}} = 1,76 + 6,02 n. \quad (3.12)$$

Una regla nemotécnica simple para aproximar la relación señal a ruido de digitalización utilizada frecuentemente es

$$S/R|_{\text{dB}} \cong 6 n. \quad (3.13)$$

Así, por ejemplo, para una resolución estándar de 16 bit la relación señal/ruido es de unos 96 dB, suficientemente alta como para representar correctamente todos los aspectos cuantitativos de los fenómenos sonoros relevantes para la percepción humana y, en nuestro caso, para fines metrológicos.

### 3.6 Distorsión a bajos niveles

Cuando la amplitud de la señal es pequeña, del orden de unos pocos LSB, la señal digitalizada presenta una marcada distorsión, como se puede apreciar en los dos espectros superiores de la figura 34.

Esta distorsión es claramente perceptible, pudiendo, en caso de haber sonidos de frecuencias diferentes, aparecer productos de intermodulación. Es un tipo de distorsión antinatural porque en vez de darse en niveles altos se presenta en niveles muy bajos. En general, a partir de una resolución de 8 bit el efecto de distorsión para un tono de máxima amplitud desaparece, presentándose en cambio como un ruido aproximadamente blanco. Pero con la misma resolución, los tonos de pequeña amplitud sí exhibirán distorsiones. En la actualidad todos los sistemas de digitalización soportan como mínimo 16 bit y muchos incluso 24 bit, con lo cual los efectos de distorsión sólo se manifiestan recién a niveles extremadamente bajos.

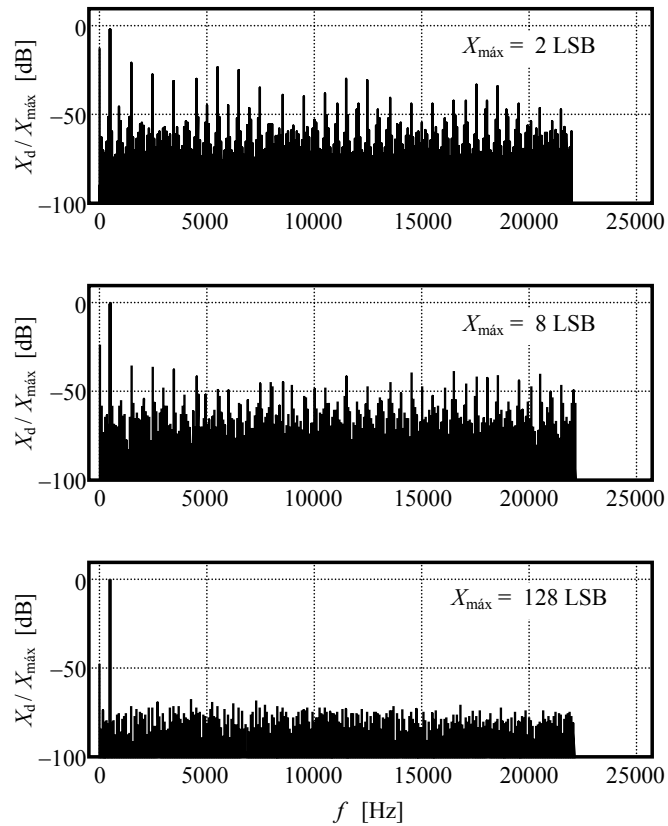
### 3.7 Dither

Una forma de reducir el problema de la distorsión a bajos niveles es agregar una pequeña cantidad de ruido aleatorio  $x(t)$  previo a digitalizar. Este ruido se denomina *dither*, y el proceso de agregarlo, *dithering*.<sup>34</sup> El valor eficaz de dicho ruido es en general una fracción del LSB, típicamente entre 0,2 LSB y 0,7 LSB, y su distribución estadística  $f(x)$  puede ser normal, triangular o uniforme. Si llamamos  $x_d$  a la versión digital del ruido, podemos calcular el valor eficaz del ruido agregado como

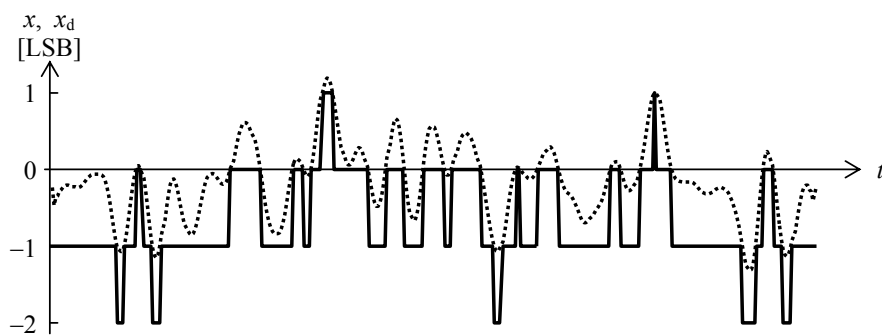
$$\begin{aligned} X_{d,\text{ef}} &= \int_{-\infty}^{\infty} x_d^2(x) f(x) dx = \sum_{k=-\infty}^{\infty} \int_k^{k+1} x_d^2(x) f(x) dx = \\ &= \sum_{k=-\infty}^{\infty} \int_k^{k+1} k^2 f(x) dx = \sum_{k=-\infty}^{\infty} k^2 \int_k^{k+1} f(x) dx = \\ &= \sum_{k=-\infty}^{\infty} k^2 (F(k+1) - F(k)) \end{aligned} \quad (3.14)$$

<sup>34</sup> La palabra inglesa *dither* equivale a indecisión, vacilación o temblequeo y alude al hecho de que se hace oscilar irregularmente a la señal alrededor del valor ideal.

Esta serie converge muy rápidamente. En la figura 35 se muestra un ejemplo de dither, en este caso aplicando un ruido filtrado para poder visualizar la conversión. Debido a que la digitalización trunca hacia abajo, el valor medio del dither es  $-0,5$  LSB.



**Figura 34.** Espectro de una señal senoidal de 500 Hz digitalizada para diversas amplitudes expresadas en LSB. Los valores indicados equivalen a digitalizar una senoide de máxima amplitud con 2 bit, 4 bit y 8 bit. En el caso de 2 bit y de 4 bit se aprecian claramente los armónicos de la fundamental. En el caso de 8 bit, sólo queda ruido residual.

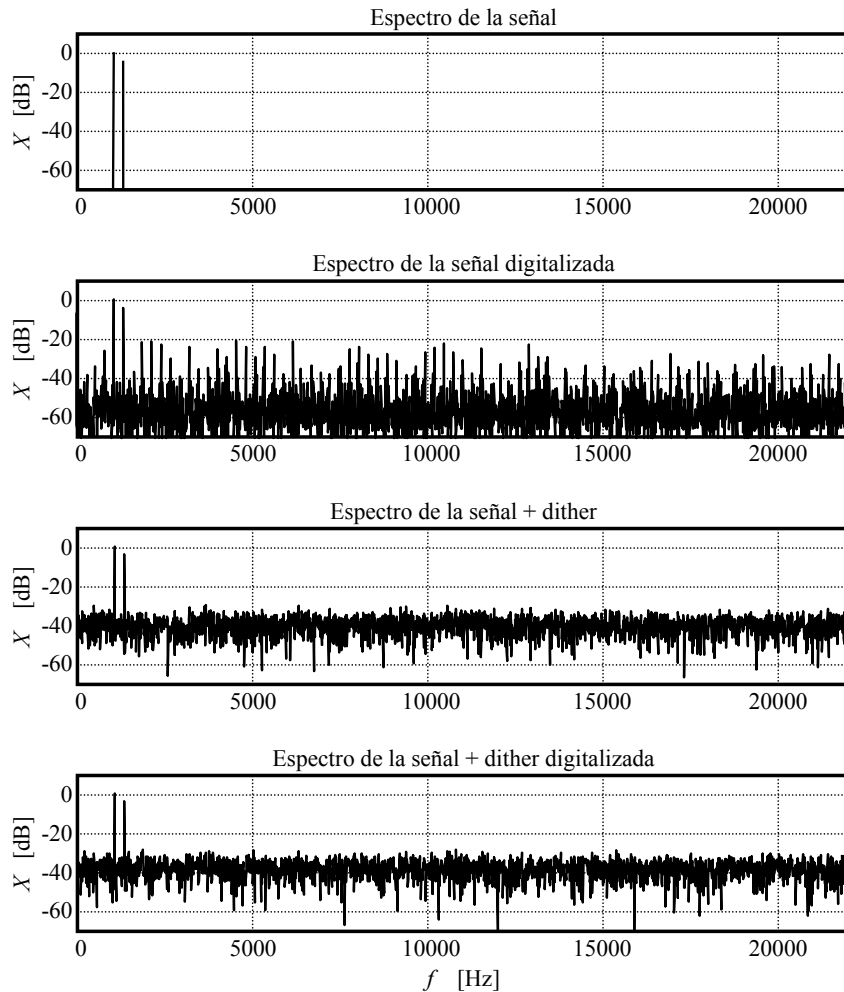


**Figura 35.** Ejemplo de señal de dithering (un ruido gaussiano filtrado con valor eficaz equivalente a  $0,6$  LSB) y el mismo ruido digitalizado.



El agregado de dither tiende a descorrelacionar el fuerte efecto no lineal que tiene lugar cuando la amplitud de la señal es de unos pocos LSB. El resultado neto es que deja de comportarse como distorsión y pasa a comportarse como ruido, con la ventaja de que la distorsión concentra su energía espectral en armónicos aislados, en tanto el ruido la distribuye en todo el espectro, haciéndose mucho más tolerable.

En la figura 36 se muestra este efecto para una señal formada por dos tonos de frecuencia  $f_0$  y  $1,25 f_0$  y amplitudes 1,1 LSB y 0,7 LSB. En la segunda gráfica se aprecia que luego de la digitalización aparece una gran cantidad de productos de intermodulación superpuestos a un ruido aproximadamente blanco. Al agregar un dither gaussiano de 0,4 LSB, antes de digitalizar aparece un ruido blanco que después de digitalizar prácticamente no cambia. Han desaparecido los productos de intermodulación.



**Figura 36.** Efecto linealizador del dithering aplicado a una señal que contiene dos tonos de 1077 Hz y 1346 Hz y amplitudes 1,1 LSB y 0,7 LSB respectivamente. De arriba hacia abajo se muestran los espectros de líneas de la señal antes y después de la digitalización, y luego la señal con agregado de dither gaussiano de 0,4 LSB, antes y después de la digitalización.

En la tabla 3.1 se muestran varios valores de valor eficaz de un dither gaussiano y de su versión digitalizada, así como la relación señal/ruido para un tono de máxima amplitud. Se aprecia la disminución de dicha relación en mayor proporción que la que se tiene antes de digitalizar. Es el precio que se paga por sustituir distorsión por ruido de banda ancha.

**Tabla 3.1.** Valor eficaz en LSB de un dither gaussiano y de su versión digitalizada. También se indica la relación señal/ruido resultante para señales de máxima amplitud

$X_{\text{ef}}$	$X_{\text{d,ef}}$	S/R <sub>dB</sub> (16 bit)
0,20	0,70711	96,33
0,25	0,7072	96,328
0,30	0,70832	96,315
0,35	0,71313	96,256
0,40	0,72446	96,119
0,45	0,74335	95,895
0,50	0,76893	95,602
0,55	0,79949	95,263
0,60	0,83343	94,902
0,65	0,86963	94,533
0,70	0,90745	94,163
0,75	0,9465	93,797
0,80	0,98658	93,437
0,85	1,0275	93,083
0,90	1,0693	92,738
0,95	1,1117	92,4
1,00	1,1547	92,07

Desde el punto de vista metrológico esto en general no tiene importancia, siempre que se ajusten los niveles de ganancia para aprovechar lo más posible el rango dinámico disponible. En el caso de señales con rango dinámico amplio (por ejemplo, la música sinfónica) puede suceder que, para evitar la saturación en momentos de gran intensidad, se ajuste la ganancia a un valor relativamente bajo, de modo que la mayor parte del tiempo la relación señal a ruido sea mucho menor que la máxima que permite la resolución. En este caso los niveles más bajos pueden quedar sumergidos en el ruido.

Una solución a este problema puede ser trabajar con resoluciones más altas. Los sistemas actuales de registro digital (grabadores o tarjetas de sonido) ofrecen resoluciones de 24 bit. Es necesario tener en cuenta que aunque tal resolución permite idealmente una relación señal a ruido de  $24 \times 6 \text{ dB} = 144 \text{ dB}$ , la electrónica analógica no permite en la actualidad un desempeño tan destacado en materia de ruido.<sup>35</sup> Por lo tanto los bits menos significativos están representando, en el mejor de los casos, el ruido circuital y el del propio transductor, que pasan, así, a constituir una suerte de dither natural.

<sup>35</sup> En efecto, si suponemos una señal máxima de 20 V, una relación señal a ruido de 144 dB implica que la tensión de ruido a la salida debe ser de 1,26  $\mu\text{V}$ . Con una ganancia 100, ello equivale a una tensión de ruido a la entrada de 12,6 nV. Para un ruido principalmente térmico, según la fórmula del ruido de Johnson el valor eficaz del ruido de una resistencia  $R$  es  $V_{\text{ef}} = \sqrt{4kTRB}$ , donde  $k = 1,38 \times 10^{-23} \text{ J/K}$  es la constante de Boltzmann;  $T$ , la temperatura absoluta; y  $B$ , el ancho de banda, resulta que la resistencia de entrada debe ser de 0,5  $\Omega$  por lo que la resistencia del micrófono debería ser bastante menor que esto.

### 3.8 Jitter

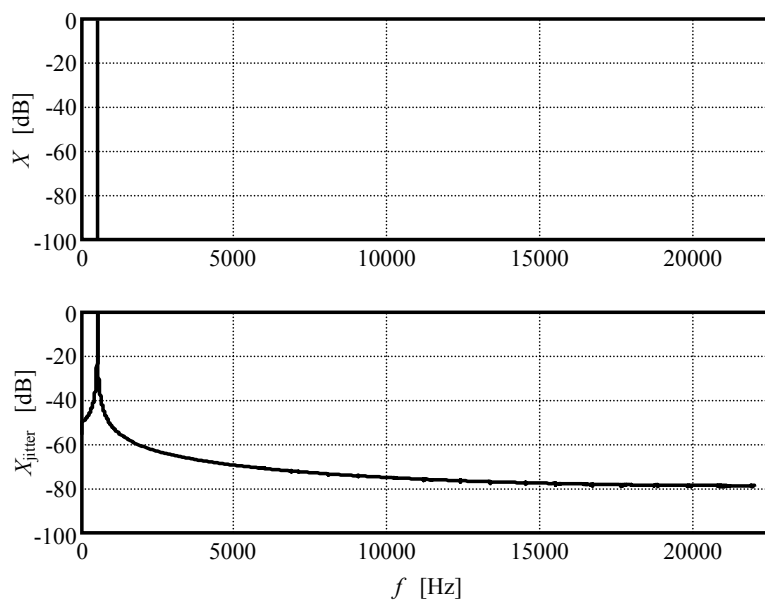
Idealmente, un sistema de muestreo toma las muestras a intervalos perfectamente regulares. En un sistema real ello no es posible debido a varias causas: inestabilidad de la frecuencia del reloj electrónico que comanda los instantes de muestreo, ruido superpuesto a la señal analógica, ruido superpuesto a la señal digital o ruido en la línea que lleva la señal de reloj entre dos puntos distantes. Estos fenómenos agregan incertidumbre al instante de muestreo que se manifiesta como una modulación de frecuencia aleatoria denominada jitter. El efecto equivale a la presencia de un ruido superpuesto a la señal.

En la figura 37 se muestra el espectro de un tono puro

$$x(t) = \sin 2\pi f_0 t.$$

y del mismo tono con una modulación de frecuencia mediante un ruido gaussiano:

$$f(t) = f_0 (1 + n(t)).$$



**Figura 37.** Ruido causado por la presencia de un jitter consistente en la modulación de frecuencia de un tono senoidal de frecuencia  $f_0 = 44100 \times 100 / 8192 \text{ Hz} = 538,33 \text{ Hz}$  con un ruido gaussiano  $n(t)$  de modo que la frecuencia instantánea es  $f(t) = f_0(1 + n(t))$ . El valor eficaz de  $n(t)$  es 0,005. Se indican los espectros de  $N = 8192$  líneas de la señal pura y contaminada por jitter.

Se observa que el espectro se ha alterado por un ruido de banda ancha pero de distribución no uniforme, concentrándose principalmente en las proximidades de la frecuencia original  $f_0$ .

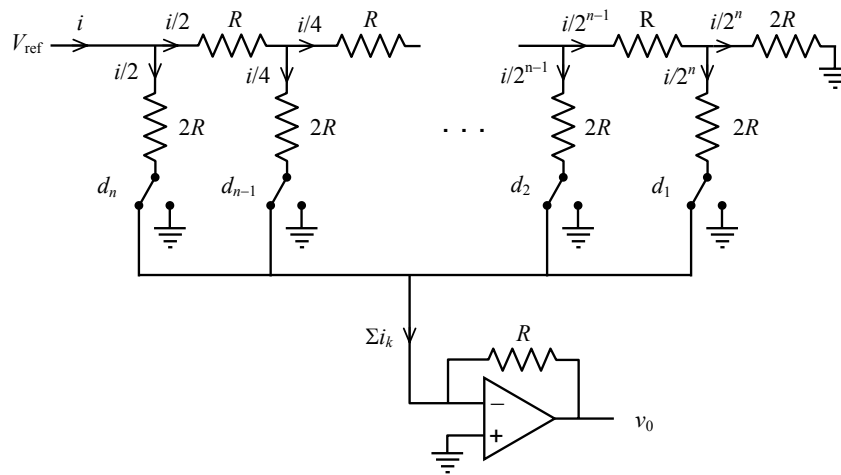
El jitter debido a ruido en las líneas de reloj o inestabilidades en la frecuencia del reloj de reconstrucción de la señal puede corregirse en forma casi perfecta, mediante una resincronización con un reloj estable, no así el generado durante el muestreo de la señal, ya que no hay manera de saber cuáles eran los valores de la señal en los instantes ideales.

### 3.9 Conversión D/A y A/D

#### 3.9.1 Conversión digital / analógica

Aunque podría parecer que desde el punto de vista de la metrología basada en software sólo interesa la conversión analógico / digital, en muchas ocasiones es necesario generar digitalmente señales de prueba que luego deben ser llevadas al dominio analógico para su utilización en un ensayo.

La conversión digital / analógica se logra en general sumando corrientes ponderadas por factores  $1/2^k$  que atraviesan llaves que permiten o no el paso de la corriente según que los bits homólogos sean 1 ó 0 respectivamente. En la figura 38 se muestra uno de tales circuitos, donde  $n$  es la resolución de los datos digitales a convertir. En este caso se utiliza una ingeniosa disposición que, alternando resistencias de valor  $R$  y  $2R$  en una red escalera, consigue sucesivos divisores de corriente por 2 que van reduciendo en factores 2, 4, 8, etc. la corriente de entrada. Todas las corrientes son proporcionales a una tensión de referencia  $V_{\text{ref}}$ .



**Figura 38.** Conversor analógico-digital R-2R en modo de corriente. La masa virtual en la entrada inversora del amplificador operacional garantiza que la propiedad de la red R-2R se cumpla.

Se obtiene, idealmente,

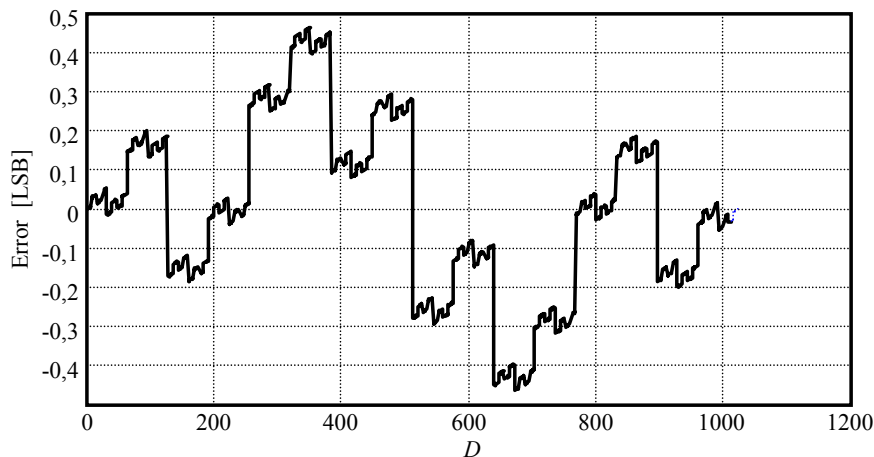
$$v_o = -\frac{V_{\text{ref}}}{R} \sum_{k=1}^n \frac{d_k}{2^{n-k+1}}. \quad (3.15)$$

donde  $d_k$  es el  $k$ -ésimo bit, contando a partir del menos significativo (LSB)

Un problema de este tipo de circuito es que requiere tolerancias muy ajustadas en las resistencias. Por ejemplo, si la tolerancia fuera de un 1 %, resultaría que un error máximo en la primera  $2R$  y mínimo en todas las demás resistencias implicaría que al pasar de 011...11 a 100...00 (un salto de tan sólo 1 LSB) la salida se redujera en 0,5 %, que para un conversor de 16 bit, por ejemplo, corresponde a un inadmisibles error de 328 LSB.

Para conseguir tolerancias mucho más estrictas se ajustan las resistencias más críticas, durante la fabricación, mediante la técnica de *laser-trimming*. Ésta consiste en descomponer cada resistencia en un conjunto de tramos en serie que luego se van cortocircuitando (por fundido mediante un haz láser) mientras se monitorea el valor de la resistencia, hasta alcanzar un valor que garantice las especificaciones.

En la figura 39 se muestra un ejemplo del error que puede causar la tolerancia de las resistencias obtenido mediante una simulación de Montecarlo.<sup>36</sup> Se grafica el error en función del dato digital  $D$  de entrada, haciendo variar éste en todo el rango desde 0 hasta  $2^n - 1$ . El curioso patrón que se observa cambia completamente de una simulación a otra, pero en general se logra una cota para el error máximo.



**Figura 39.** Ejemplo del error en LSB en un convertor  $R$ - $2R$  de 10 bit cuando la tolerancia de las resistencias es del 0,2 % (corresponde a una incertidumbre relativa con  $\sigma = 0,001$  para una confiabilidad del 95 %).

El error ilustrado es un *error de linealidad*, ya que equivale a que el factor de escala varíe entre muestras sucesivas. Se produce, por lo tanto, una pequeña distorsión que se agrega a la propia de la discretización, aunque si se logra mantener por debajo de 1 LSB su efecto en general es despreciable. En este ejemplo, para una onda senoidal de máxima amplitud, la distorsión debido a la digitalización más un dither de 0,3 LSB es de un 0,116 %, mientras que al incluir el error de no linealidad se eleva a 0,127 %. La diferencia no es significativa.

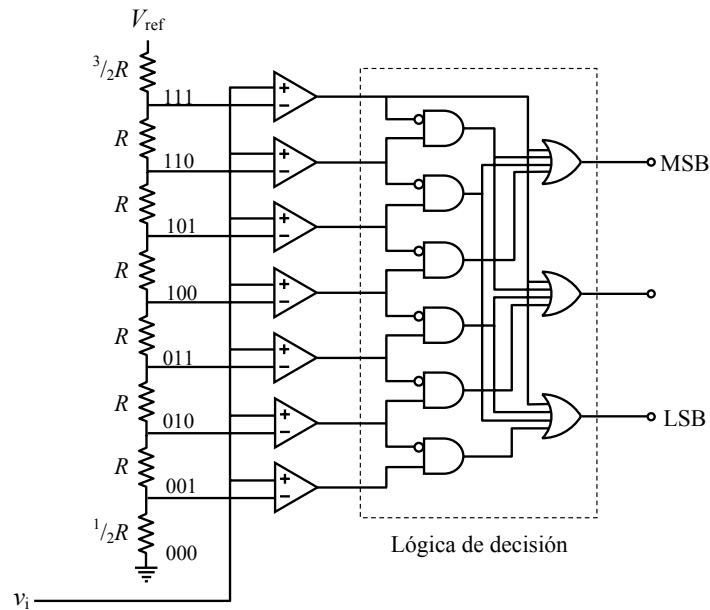
### 3.9.2 Conversión analógica / digital

Hay muchas formas de realizar una conversión analógica / digital, aunque en sistemas que requieren una alta tasa de conversión en el tiempo sólo se utilizan dos: la conversión flash (en paralelo) y la conversión por aproximaciones sucesivas.

Los conversores flash consisten en una serie de comparadores que comparan la tensión de entrada con cada uno de los  $2^n$  niveles posibles con  $n$  bit, logrados mediante

<sup>36</sup> El método de Montecarlo consiste en seleccionar los valores de los parámetros (en este caso las resistencias) mediante un generador de números aleatorios con la misma distribución estadística con que se distribuyen dichos parámetros

un divisor resistivo múltiple. Las salidas de los comparadores son 1 para los niveles menores que la entrada y 0 para los mayores. Luego un circuito lógico combinatorio transforma los  $2^n$  valores intermedios en una representación binaria de  $n$  bit. En la figura 40 se ilustra un ejemplo de 3 bit.



**Figura 40.** Estructura de un convertor analógico-digital flash (en paralelo) de 3 bit.

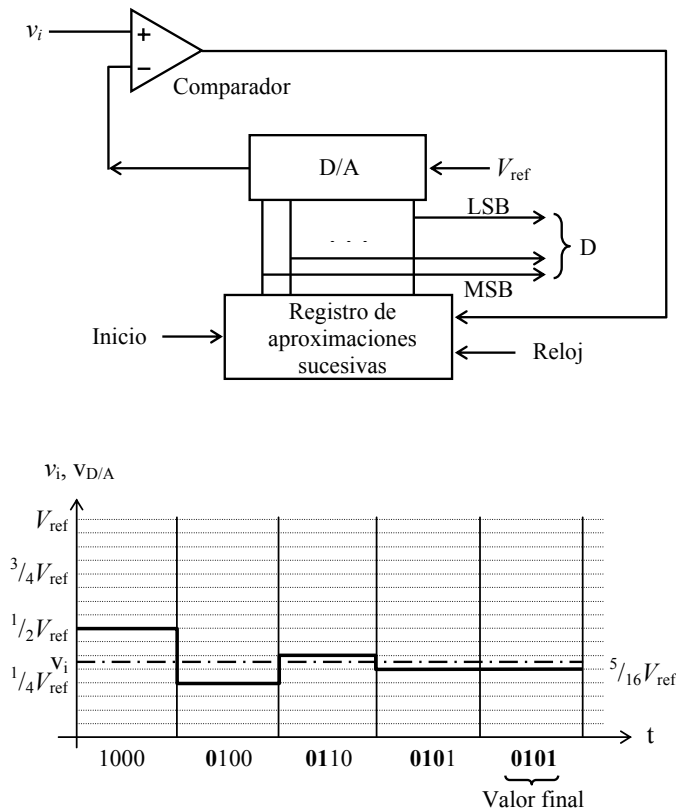
La principal ventaja de los convertidores flash es su gran rapidez, ya que la única latencia (demora) en completar la conversión está en el tiempo de conmutación de los comparadores y en el tiempo de operación del circuito lógico.

El mayor inconveniente es que para altas resoluciones la cantidad de comparadores de precisión que se requieren crece exponencialmente, al igual que la complejidad de la lógica asociada. Se logran implementaciones de hasta 10 bit. Estas implementaciones suelen dividir el problema mediante una preselección de rangos y posterior asignación de los valores extremos de tensión del divisor resistivo.

La conversión por aproximaciones sucesivas (figura 41) compara la señal de entrada con la salida de un convertor digital / analógico que se alimenta con valores que se van aproximando dicotómicamente al valor final. Para ello, en un primer paso se aplica un 1 al bit más significativo (MSB) del convertor D/A y 0 en los restantes. Si la entrada es mayor que la salida del convertor D/A se confirma dicho bit; si es menor se vuelve a 0. En el siguiente paso se aplica un 1 en el bit que sigue. Nuevamente, si la entrada es mayor que la salida del convertor D/A, se confirma el segundo bit; de lo contrario, se vuelve a 0. El proceso se repite hasta llegar al bit menos significativo, de modo que en  $n$  pasos se completa la conversión. El elemento que administra este algoritmo se denomina registro de aproximaciones sucesivas.

Es importante observar que, a diferencia del convertor flash, en este caso se requiere que la entrada se mantenga rigurosamente constante, de lo contrario podrían producirse errores muy groseros. En efecto, una vez que los bits más significativos han

quedado fijados, ya no es posible cambiarlos hasta la próxima conversión, por lo cual el proceso continúa buscando la mejor aproximación que sea posible con los restantes bits. Por esta razón se requiere un *sample and hold* (muestreo y retención) a la entrada.



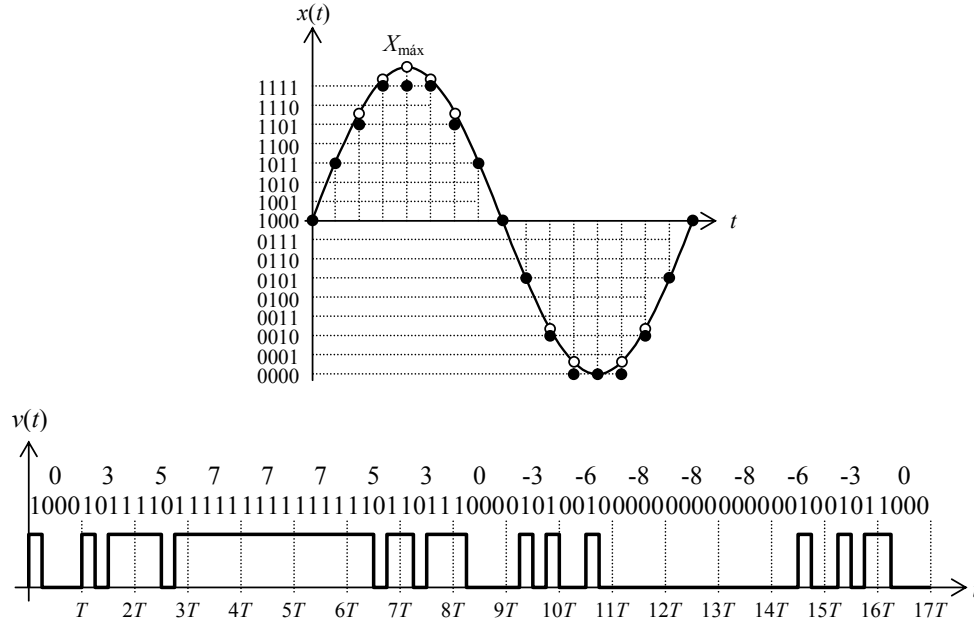
**Figura 41.** Arriba, esquema de un convertor analógico / digital de aproximaciones sucesivas. Abajo, ejemplo del proceso de acercamiento al valor final en un convertor de 4 bits. Los dígitos en **negrita** representan los que en cada etapa han quedado estabilizados.

### 3.10 Modulación de código de pulsos (PCM)

Cuando la conversión analógica / digital se realiza reiteradamente a lo largo del tiempo, se obtiene una secuencia sincronizada de datos digitales. Dicha secuencia se puede representar en paralelo, utilizando una línea por cada bit (con un total de  $n$  líneas), donde la tensión eléctrica de cada línea puede ser 0 ó  $V$ , según que el bit sea 0 o 1, o bien puede representarse en serie, utilizando una única línea que transmite los  $n$  bits en forma sucesiva. Este proceso, denominado *modulación de código de pulsos*, *PCM* (del inglés, *pulse code modulation*) se ilustra en la figura 42

En este código es preciso contar con un adecuado sincronismo, dado que el desplazamiento de un solo bit es suficiente para que la señal completa se desvirtúe, tornándose inutilizable. En muchos casos se agregan bits adicionales, por ejemplo de detección de errores aislados por paridad, o destinados a mantener la componente de continua a mitad de camino entre ambos niveles. En otros casos se realiza una recodificación completa de la señal siguiendo una tabla de equivalencias donde los bits dejan de inter-

pretarse como un sistema de numeración posicional a cambio de ganar propiedades interesantes como la capacidad de corrección de errores o la posibilidad de albergar un código específico de sincronismo. En todos estos casos un bloque lógico se encargará de decodificar la información a efectos de su almacenamiento o procesamiento ulterior.



**Figura 42.** Ejemplo de modulación por código de pulsos (PCM). Una onda senoidal de máxima amplitud ha sido codificada con 4 bit con un período de muestreo  $T$ . Los bits se representan como valores bajos (0) y altos (1) de tensión. Cada bit ocupa  $1/4$  del periodo de muestreo. Sobre los datos binarios se indica el valor decimal correspondiente.

### 3.11 Modulación de código de pulsos diferencial (DPCM)

En lugar de digitalizar el valor muestreado de la señal, a veces es conveniente digitalizar la diferencia entre el valor de la muestra actual y el de la muestra anterior. El resultado se conoce como *modulación de código de pulsos diferencial*, *DPCM*. Para señales de tasa de variación acotada, esto puede suponer una conveniente reducción en la resolución necesaria. Supongamos, por ejemplo, una señal  $v(t)$  de amplitud máxima  $V_{\text{máx}}$  que ha sido filtrada mediante un filtro pasabajos con constante de tiempo  $\tau$ . Entonces su máxima pendiente (correspondiente a una conmutación entre  $-V_{\text{máx}}$  y  $V_{\text{máx}}$ ) será

$$\left. \frac{dv}{dt} \right|_{\text{máx}} = \frac{2V_{\text{máx}}}{\tau}$$

Si el período de muestreo es  $T$ , el máximo salto posible entre dos muestras sucesivas será

$$\Delta v|_{\text{máx}} = 2V_{\text{máx}} \frac{T}{\tau},$$



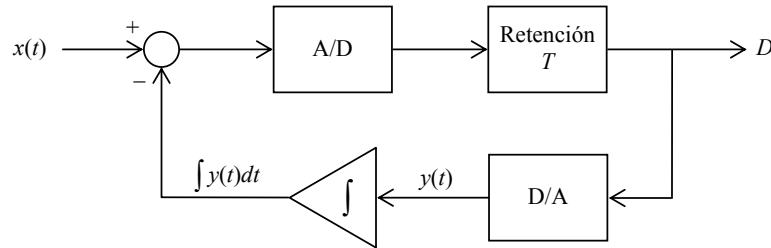
que, si  $T \ll \tau$ , resulta mucho menor que  $2V_{\text{máx}}$ . Puede, por lo tanto, codificarse con menos bits. De hecho, por cada reducción de  $T$  a la mitad se ahorra un bit.

Obsérvese que a partir de las diferencias entre muestras sucesivas es posible recuperar la señal original simplemente haciendo la suma acumulada de todas las diferencias anteriores desplazada en el valor inicial de la señal:

$$v(t_n) = v(t_1) + \sum_{k=1}^{n-1} (v(t_{k+1}) - v(t_k)). \quad (3.16)$$

La conversión de la diferencia entre los valores actual y anterior implica disponer al mismo tiempo de ambos valores analógicos. Una forma, aunque impráctica, de hacerlo, sería mediante una línea de retardo analógica con dos *sample and hold* en cascada. Luego, un restador analógico obtendría la diferencia a digitalizar.

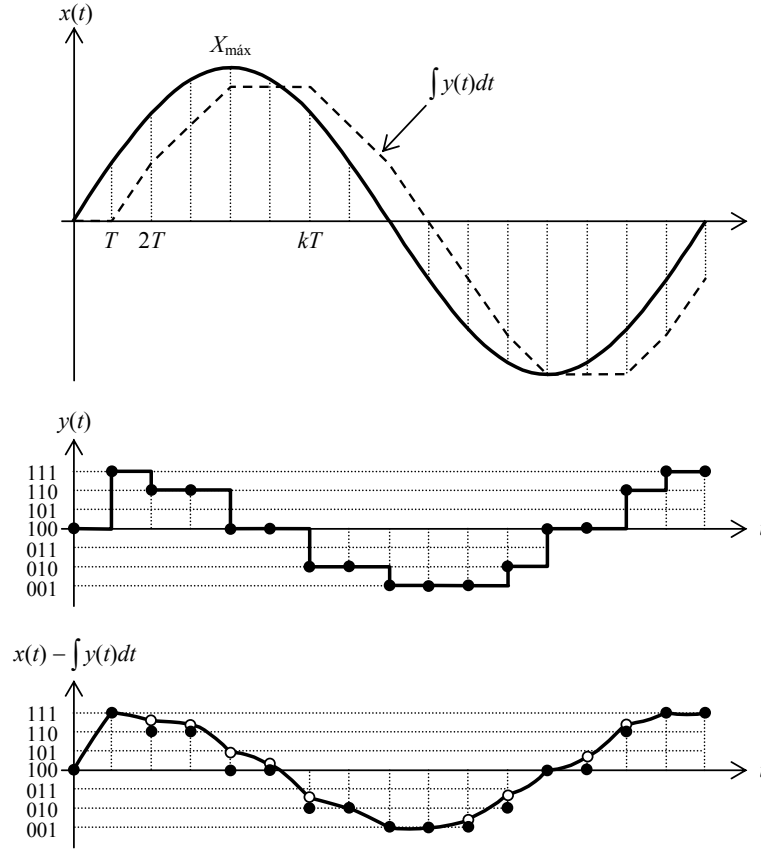
Una forma más apropiada es como se muestra en la figura 43, donde el valor analógico anterior se reconstruye a partir de la diferencia digitalizada. Para ello, primero se la convierte en analógica y luego se la integra. El bloque de retención simplemente mantiene el valor durante un periodo de muestreo  $T$  (hasta la próxima muestra).



**Figura 43.** Diagrama de bloques de un conversor DPCM. El conversor A/D convierte la diferencia entre la muestra actual y la reconstrucción de la muestra anterior.

En la figura 44 se ilustran las formas de onda en diversos puntos del diagrama anterior durante la conversión de una senoide. En el instante inicial se muestrea y convierte a digital el valor  $x(0) - \int y(t)dt$ , que en este primer ciclo es 0 por ser  $x(0) = \int y(t)dt = 0$  ya que el integrador aún no ha recibido señal distinta de 0 en su entrada. Este valor es retenido durante un tiempo  $T$ , convertido a analógico e integrado (el resultado es nulo por ser nulo el integrando). Al cumplirse el primer periodo de muestreo vuelve a muestrearse y convertirse la diferencia entre la señal y la integral que aproxima el valor de la señal en  $t - T$ . El nuevo valor es retenido, convertido a analógico e integrado. Por ser constante la integral da una rampa de pendiente proporcional a su entrada. El restador analógico produce la curva indicada en la grafica inferior de la figura 44.

Al término de este segundo período de muestreo se toma la tercera muestra de  $x(t) - \int y(t)dt$ , se convierte a digital y se retiene hasta  $3T$ . Se integra, obteniéndose una nueva rampa que es restada de la señal y esta diferencia es convertida en el instante  $3T$ . El proceso se repite generando los valores del código DPCM.



**Figura 44.** Formas de onda en el conversor DPCM para una senoide de máxima amplitud ( $X_{\text{máx}} = X_{\text{ref}}$ ). La onda en línea de trazos es la salida del integrador

A partir de la señal digital DPCM se puede obtener la señal PCM mediante una integral discreta o suma acumulativa:

$$D_{\text{PCM}}(k) = \sum_{h=1}^k D_{\text{DPCM}}(h), \quad (3.17)$$

o bien, mediante la expresión recursiva

$$D_{\text{PCM}}(k) = D_{\text{PCM}}(k-1) + D_{\text{DPCM}}(k). \quad (3.18)$$

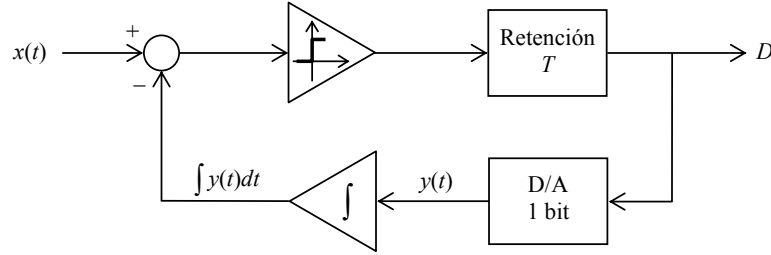
que es computacionalmente más eficiente.

La modulación DPCM requiere en general menos bits que la modulación PCM, aunque para poder explotar esta ventaja se requiere incrementar la tasa de muestreo o limitar la velocidad de excursión de la señal, lo cual implica limitar el contenido espectral muy por debajo de la frecuencia de Nyquist. La cantidad de bits por unidad de tiem-

po se mantendrá, por lo tanto, constante. Tal como se mencionó anteriormente, esto es ventajoso, ya que permite simplificar los filtros necesarios.

### 3.12 Modulación delta (DM)

Dado que la modulación DPCM requiere menos bits que la modulación PCM, podemos llevar esta idea hasta el extremo en que se requiera un solo bit. El conversor analógico/digital será simplemente un detector de signo que asignará un 1 cuando la señal sea positiva y un 0 cuando sea negativa. La modulación así obtenida se denomina *modulación delta*, *DM*. En la figura 45 se muestra un diagrama de bloques y en la figura 46 se ven las formas de onda.



**Figura 45.** Diagrama de bloques de un conversor DM. El conversor D/A se limita a desplazar hacia abajo  $\frac{1}{2}$  LSB la salida binaria de modo que un 0 corresponda a  $-\frac{1}{2}$  LSB y un 1 a  $\frac{1}{2}$  LSB.

Para que la idea sea viable es necesario que la pendiente de la señal sea lo suficientemente baja como para que el incremento máximo en un periodo de muestreo no supere 1 LSB, es decir

$$\left| \frac{dx}{dt} \right| < \frac{1 \text{ LSB}}{T}. \quad (3.19)$$

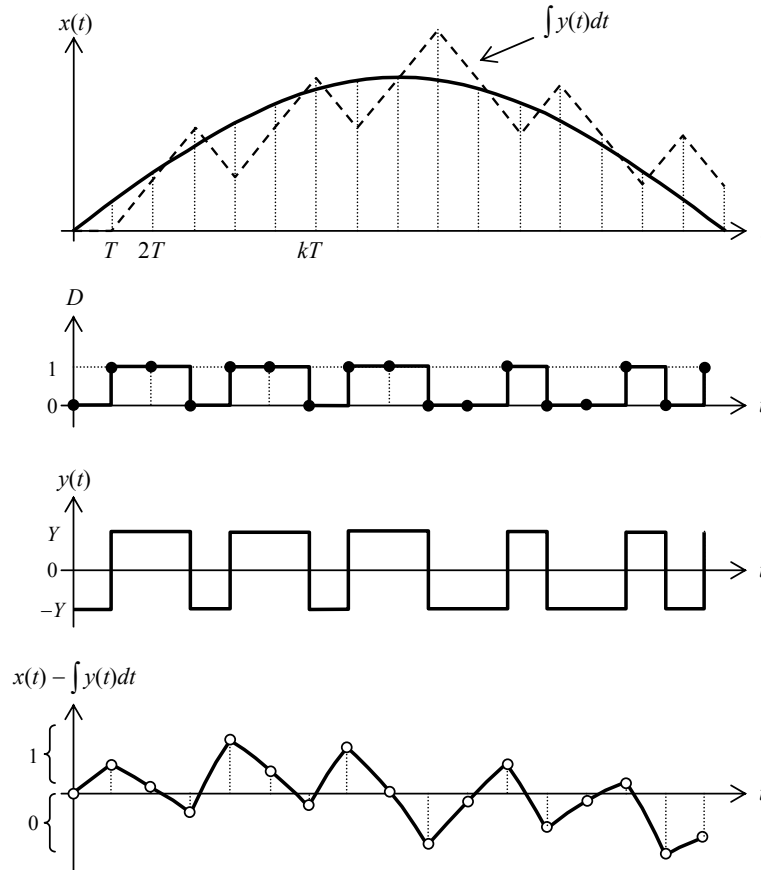
Para ello es preciso, en general, aumentar enormemente la tasa de muestreo.

Es interesante observar que en modulación delta no hay necesariamente un límite de amplitud, ya que una señal podría crecer indefinidamente con una pendiente acotada y aun así podría representarse con saltos de 1 LSB. Sin embargo, para una dada frecuencia, por ejemplo la máxima de interés (20 kHz en el caso de audio), la pendiente máxima es directamente proporcional a la amplitud e inversamente proporcional a la frecuencia:

$$\left| \frac{d}{dt} V \sin \omega t \right| \leq V \omega. \quad (3.20)$$

De las ecuaciones (3.19) y (3.20) se obtiene la siguiente cota para el periodo de muestreo:

$$T < \frac{1 \text{ LSB}}{V_{\text{máx}} \omega_{\text{máx}}} \quad (3.21)$$



**Figura 46.** Formas de onda en el conversor DM para una senoide. La onda en línea de trazos es la salida del integrador

A modo de ejemplo, supongamos una señal de 20 kHz y 10 V de amplitud que se quiere representar con 16 bit. En ese caso

$$1 \text{ LSB} = \frac{10 \text{ V}}{2^{15}} = 0,00030518509 \text{ V}$$

Deberá ser

$$T < \frac{0,00030518509}{10 \text{ V} \times 2\pi \times 20000 \text{ Hz}} = 0,24286 \text{ ns} \quad (3.22)$$

que corresponde a 4,1176 GHz, frecuencia demasiado alta para ser práctica. En caso de utilizarse menor tasa de muestreo sobrevendrá una distorsión en la pendiente máxima del mismo tipo que la distorsión por slew-rate (limitación de la máxima pendiente de salida) de algunos amplificadores.

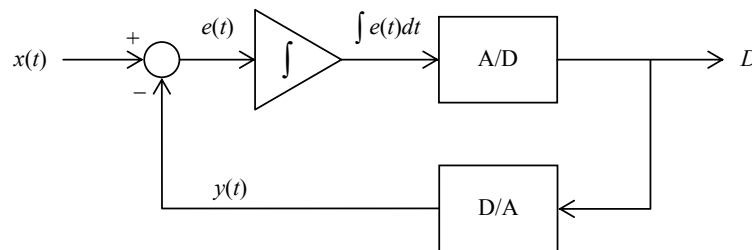
La modulación delta presenta el problema de *granularidad*, esto es, dado que sólo detecta valores positivos o negativos de pendiente, no permite representar una señal constante, sino que la aproxima oscilando alrededor del valor real. En particular, no es posible una señal nula. La granularidad se manifiesta como un ruido superpuesto a la señal.

La decodificación de este tipo de modulación digital se logra mediante un simple filtro pasabajos.

En lo que respecta al ruido, permite repartir el ruido de digitalización en una banda mucho más amplia, con lo cual se reduce el ruido dentro de la banda audible.

### 3.13 Modulación sigma-delta (SDM)

La *modulación sigma-delta*, SDM, se basa, nuevamente, en el concepto de realimentar la señal convertida, pero en este caso ubica la integración después de la comparación, en lugar de antes, tal como se muestra en la figura 47.



**Figura 47.** Diagrama de bloques de un conversor sigma-delta.

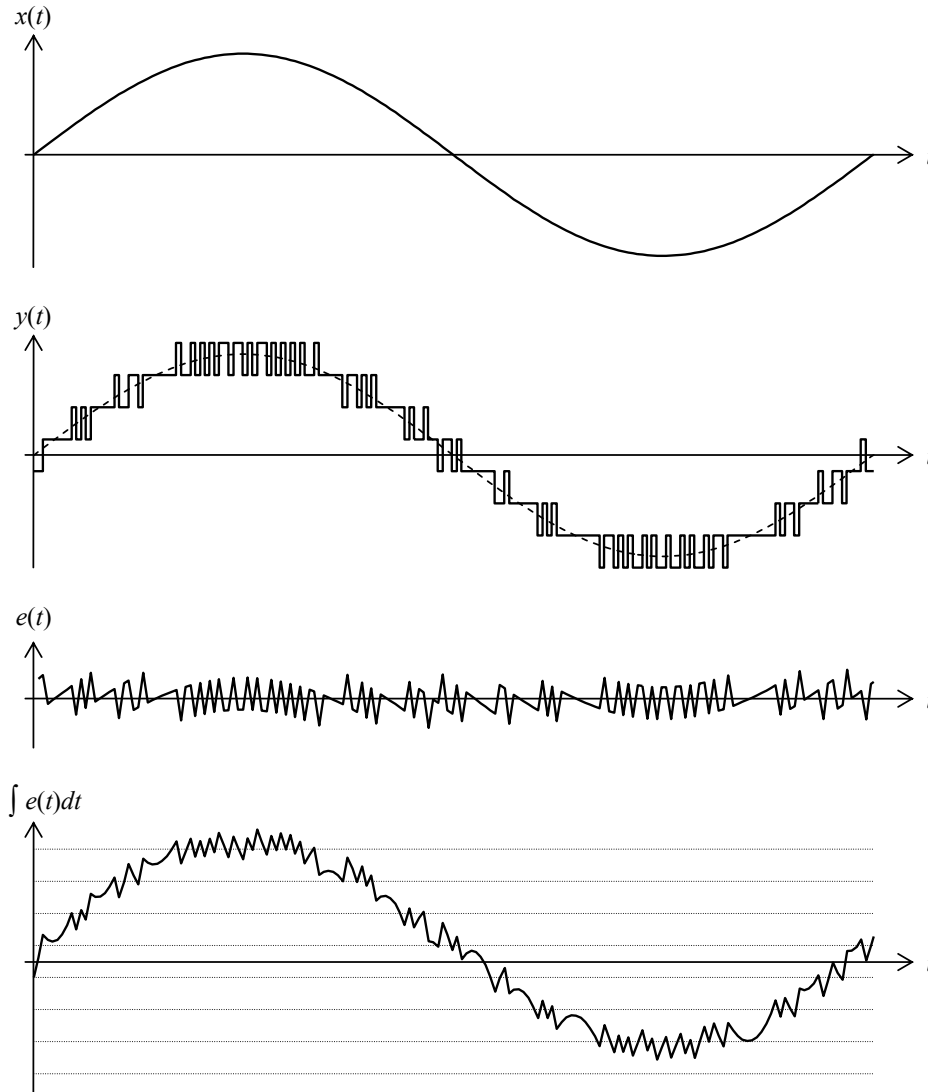
En este caso se compara la señal original con la versión analógica de la señal convertida, obteniéndose una señal de error que es integrada antes de la conversión a digital. El conversor sigma-delta utiliza el concepto de *sobremuestreo*, es decir, un muestreo a una tasa  $F_{os}$  muy superior a la tasa de muestreo  $F_s$  deseada. El cociente entre ambas es el *factor de sobremuestreo*,  $N_{os}$ , de modo que

$$F_{os} = N_{os} F_s. \quad (3.23)$$

El factor de sobremuestreo suele ser una potencia de 2 relativamente alta, por ejemplo 64 ó 128, y la resolución de los conversores A/D y D/A es baja, típicamente entre 1 bit y 5 bit. Luego de la modulación será necesario reducir la tasa de muestreo hasta  $F_s$  (proceso conocido como *diezmado* o *submuestreo*) y aumentar la resolución hasta el valor requerido por el formato final. Para lograrlo se requieren tres etapas. Primero se aplica un filtro antialias digital que elimina las frecuencias superiores a  $F_s/2$ , luego se realiza un submuestreo digital, tomando una de cada  $N_{os}$  muestras, y finalmente se procede a recuantizar digitalmente la señal para tener mayor resolución. Es interesante notar que la recuantización digital es un proceso equivalente a la conversión analógica/digital. Si bien en principio podrían truncarse los valores de las muestras provenientes del filtro antialias,<sup>37</sup> en la práctica es conveniente aplicar una señal de *dither* digital para reducir la distorsión a bajos niveles de señal.

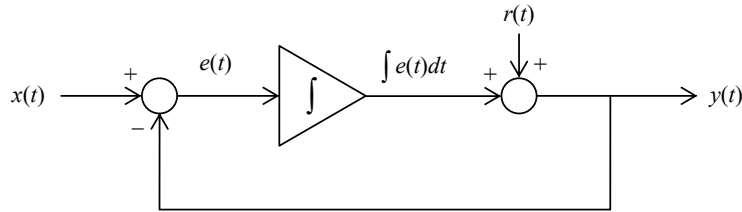
<sup>37</sup> Tengamos en cuenta que el filtro digital trabaja con una resolución muy superior a la de la señal  $D$ . Dicha resolución depende de la aritmética del procesador utilizado.

En la figura 48 se observa una simulación numérica de la respuesta del conversor sigma-delta ante un tono senoidal de 1 kHz, siendo la conversión de 3 bit, el factor de sobremuestreo  $N_{os} = 4$  y la tasa de muestreo deseada  $F_s = 44\,100$  Hz. La señal  $y(t)$  de retorno de la realimentación, que equivale a la señal de salida, oscila alrededor de la señal de entrada entre los dos escalones sucesivos más próximos. Como dichas oscilaciones son de alta frecuencia, serán eliminadas por el filtro pasabajos antialias que se coloca a la salida.



**Figura 48.** Formas de onda de un conversor sigma-delta de 3 bit con  $F_s = 44\,100$  Hz y factor de sobremuestreo 4. De arriba hacia abajo: Entrada senoidal  $x(t)$  de 1 kHz; retorno de la realimentación  $y(t)$ , equivalente a la salida; señal de error  $e(t)$ ; integral del error y niveles de cuantización del conversor A/D.

Esta estructura tiene la ventaja de que reduce el error de baja frecuencia. En efecto, para una señal lentamente variable, de existir un error importante, el mismo sería rápidamente integrado tendiendo a modificar la señal realimentada de modo de reducir el error.<sup>38</sup> Esto puede verificarse analíticamente mediante el modelo linealizado de la figura 49. Allí se ha reemplazado la acción conjunta de los conversores A/D y D/A por el agregado de un ruido  $r(t)$  de cuantización.



**Figura 49.** Modelo lineal del convertidor sigma-delta.

Teniendo en cuenta que la función de transferencia en transformada de Laplace del integrador es  $1/\tau s$  y que la realimentación es unitaria, se verifica que

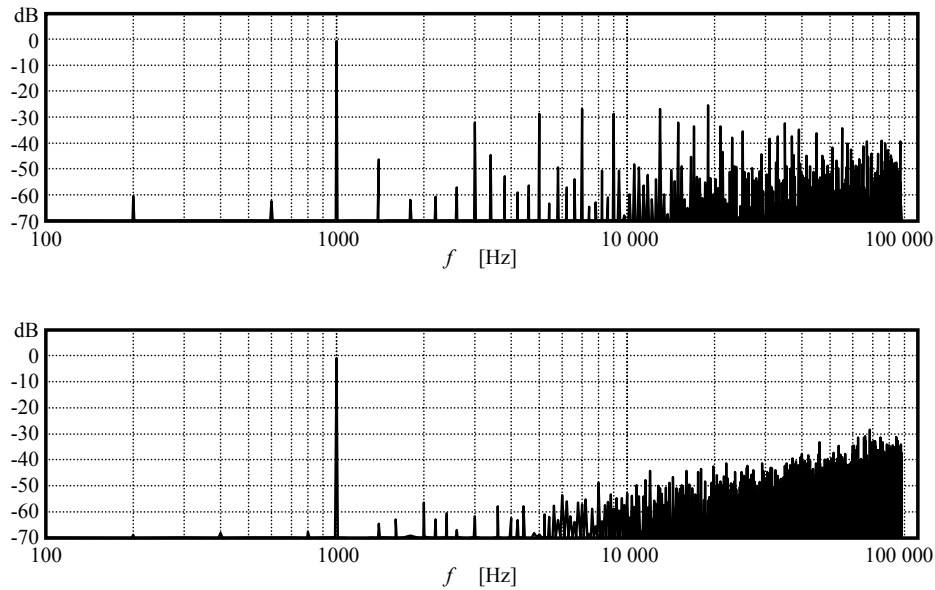
$$Y(s) = \frac{1}{1 + \tau s} X(s) + \frac{\tau s}{1 + \tau s} R(s). \quad (3.24)$$

Observemos que para la señal  $x(t)$  la función de transferencia es un pasabajos con frecuencia angular de corte  $\omega_c = 1/\tau$ , mientras que para el ruido  $r(t)$  es un pasaaltos con igual frecuencia de corte. Esto implica que el ruido que subsiste en  $y(t)$  es el de alta frecuencia, por lo cual el filtro pasabajos que se utilice como filtro antialias previo al submuestreo lo reducirá en gran medida.

En la figura 50 se muestra el espectro de la señal de entrada cuantizada en forma tradicional (por clasificación de niveles) y cuantizada mediante modulación sigma-delta, en ambos casos luego de la conversión digital/analógica. Según se puede apreciar, la conversión tradicional produce una serie de componentes tonales de niveles considerables dentro de la banda base ( $f < F_s/2$ ), que incluye la serie de armónicos impares. La modulación sigma-delta, en cambio, reduce las componentes en la banda base a costa de incrementar la amplitud del espectro en alta frecuencia. Esto último, sin embargo, no supone un problema, dado que las componentes espectrales de alta frecuencia se eliminan mediante el filtro pasabajos antialias.

La idea puede llevarse aún más lejos sustituyendo el integrador, que es, en definitiva, un tipo de filtro, por un filtro de mayor orden, para lograr lo que se denomina *conformación de ruido* (*noise shaping*). Consiste en que el filtro desplace el espectro del error de cuantización desde las frecuencias correspondientes a la banda base ( $f < F_s$ ) hacia las frecuencias por encima de ésta de una manera controlada. De esa forma, la atenuación del ruido de cuantización en la banda base será mucho mayor.

<sup>38</sup> La idea se inspira en la teoría del Control Automático, donde se utilizan integradores para anular el error estático y reducirlo considerablemente para las variaciones más lentas.



**Figura 50.** Arriba, espectro de un tono de 1 kHz cuantizado con 3 bit por clasificación de niveles. Abajo, espectro de un tono de 1 kHz cuantizado con 3 bit por modulación sigma-delta. En los dos casos la tasa de muestreo deseada es  $F_s = 44100$  Hz y el factor de sobremuestreo es  $N_{os} = 4$ .

Un caso particular frecuente en conversores sigma-delta es que el número de bits sea 1. En este caso el conversor A/D es un simple comparador que entrega una tensión  $V_{\text{máx}}$  (correspondiente a un 1) si su entrada es positiva y  $-V_{\text{máx}}$  (correspondiente a un 0) si es negativa. En la figura 51 se muestran las formas de onda obtenidas ante una señal de entrada que crece en forma de rampa. Según se aprecia, la señal de salida modulada en 1 bit acumula más 1's seguidos cuanto mayor sea la entrada.

### 3.14 Dispositivos de registro digital

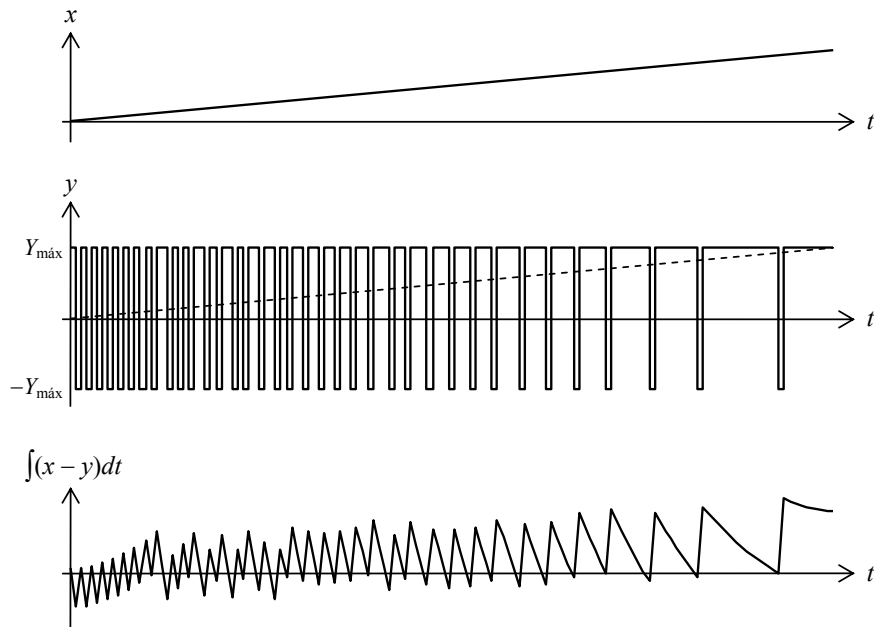
Tradicionalmente, la adquisición de señales de audio para su ulterior procesamiento y análisis ha sido posible a través del uso de grabadores de instrumentación, por ejemplo, el grabador portátil a cinta magnética Nagra IV-SJ, equipo analógico de la década de 1970. Con la introducción de la tecnología de registro digital en la década de 1990, otra opción fue el grabador de cinta de audio digital (DAT).

Varios equipos portátiles existentes en el mercado, tales como el DAT, el grabador de Minidisc (MD) y el grabador de disco duro (HD), han sido estudiados con respecto a su idoneidad para el mismo uso (Miyara, 2001, 2001a). El DAT y el HD se han encontrado adecuados para prácticamente todas las aplicaciones, mientras que el MD ha mostrado una respuesta en frecuencia algo irregular a altas y bajas frecuencias, como así también un rango dinámico efectivo reducido para la recuperación de tonos puros en casos comprometidos, por ejemplo, un pozo espectral angosto. Otro problema del MD es que, salvo que se lea en un equipo de escritorio, la salida es analógica y debe volverse a digitalizar en una computadora, lo cual supone una nueva degradación de la señal.



Por otra parte, muchos equipos tienen incorporados recursos anticopia que impiden efectuar una segunda copia digital del material.<sup>39</sup>

En años recientes, se ha popularizado una nueva tecnología que usa una memoria flash de estado sólido removible como medio de grabación. Son ejemplos las tarjetas de memoria SD (*Secure Digital*), HCSD (*High Capacity Secure Digital*) o CF (*Compact Flash*). La baja potencia y capacidades de almacenamiento que exceden hoy los 32 Gb permiten efectuar grabaciones prácticamente continuas<sup>40</sup> durante más de 48 horas, constituyéndose en una tecnología muy adecuada para sesiones de grabaciones de campo de larga duración. Como estos grabadores no emplean partes mecánicas, su operación es completamente silenciosa, siendo ideales para grabar en ambientes de muy bajo ruido.



**Figura 51.** Formas de onda de un conversor sigma-delta de 1 bit. Arriba, entrada en forma de rampa. Centro, retorno de la realimentación, equivalente a la salida. Abajo, integral del error. El conversor A/D de 1 bit produce un 0 si su entrada es negativa y un 1 si es positiva. El conversor D/A transforma el 0 y en 1 en dos niveles opuestos,  $-Y_{\text{máx}}$  e  $Y_{\text{máx}}$ .

A pesar de que existen sonómetros que permiten directamente la grabación digital en memoria flash, esta característica incrementa considerablemente el costo del instrumento, aun cuando la opción sea sólo una actualización de firmware o incluso una habilitación de un software preinstalado. Por ejemplo, la licencia para habilitar el módulo de grabación de sonido del equipo Brüel & Kjær 2250 tiene un costo de más de U\$S 2500,

<sup>39</sup> La finalidad de esto es prevenir el copiado de material sujeto a derechos de propiedad intelectual (copyright)

<sup>40</sup> El sistema de archivos FAT 16 / 32 de las memorias flash impone un límite máximo de 4 Gb para el tamaño de cualquier archivo, por lo cual es necesario interrumpir la grabación. Algunos grabadores sortean esta dificultad iniciando un nuevo archivo una vez superado el tamaño máximo de modo tal que los dos archivos puedan luego empalmarse para formar uno más grande.

mientras que grabadores digitales como el Zoom H4n de dos canales están valuados en alrededor de U\$S 700 (precios en el mercado argentino en 2012).

También se encuentran en estado experimental equipos de muy bajo costo desarrollados enteramente mediante software libre y hardware libre<sup>41</sup> que graban en formatos comprimidos sin pérdida, como el formato FLAC, con capacidad de transferencia inalámbrica a equipos con conexión WiFi, incluyendo *routers* WiFi que permiten subir los archivos a un repositorio ubicado en un sitio específico de Internet (Raynaudo, 2013).

Dado que los grabadores digitales de bajo costo están orientados al mercado de consumo, particularmente para la grabación personal de conciertos o demos de interpretaciones musicales, existe una legítima preocupación en cuanto a si la calidad de la señal digital así obtenida es suficiente para cumplir con la normativa. Sobre todo porque las especificaciones omiten parámetros que son de interés fundamental para nuestra aplicación, como la relación señal/ruido, la distorsión o la respuesta en frecuencia.<sup>42</sup>

Normas internacionales tales como las ISO 1996-2:2007 especifican que para que una pieza del equipo sea idónea para propósitos de medición, el sistema completo de instrumentación que la incluye debería cumplir plenamente con la norma IEC 61672 sobre sonómetros. Esto significa, en particular, que los límites de tolerancia para la respuesta en frecuencia son muy estrictos. Además, se incluyen otras condiciones sobre ruido auto-generado, linealidad y respuesta transitoria.

Es importante tener en cuenta que el proceso de digitalización involucra algunos procesos que podrían dejar secuelas en la señal. Uno de ellos es el filtro antialias (ya descrito), necesario para evitar severas distorsiones en la señal. Este filtro podría alterar la respuesta en frecuencia, sobre todo en el extremo superior del espectro audible. Otro proceso es la propia conversión analógica / digital, ya que impone un límite al rango dinámico de la señal resultante, como consecuencia del ruido de cuantización. A su vez, este ruido tiende a correlacionarse con la señal cuando ésta es pequeña, transformándose en una distorsión. Para minimizar este efecto se suelen agregar *dither* (dosis controladas de ruido), como ya se ha visto, al precio de incrementar ligeramente el ruido.

Con el fin de verificar que estas limitaciones no afectan significativamente la calidad de la señal adquirida, se realizó un ensayo completo de tres unidades diferentes del grabador digital Zoom H4 (Miyara et al., 2010, 2010a). El estudio reveló que dicho modelo es apto para su uso en una cadena de medición. Así, por ejemplo, la respuesta en frecuencia es plana en al menos un factor de 10 menor que la tolerancia admisible (figura 3.6). La diferencia entre las respuestas a 24 °C y 40 °C resultó menor que 0,05 dB entre 20 Hz y 10 kHz, y sólo por encima de 15 kHz alcanzó los 0,2 dB, valor mucho menor que la tolerancia aceptable. La relación señal / ruido sin ponderar resultó mayor de 90 dB, y ponderada A, mayor de 95 dB. La linealidad, expresada en términos de la respuesta real menos la ideal, según lo estipula la Norma IEC 61672, resultó menor de 0,2 dB para niveles de señal referidos al máximo digital entre -10 dB y -70 dB. Expresada como distorsión total armónica, resultó del orden de 0,02 % entre -10 dB y -50 dB, llegando a un 5 % para señales de -73 dB respecto al máximo. Este incremento de la distorsión para

<sup>41</sup> El hardware libre es una reciente modalidad de hardware cooperativo que provee soluciones de procesamiento y conectividad sin restricciones de uso ni obligación de pago de regalías, con acceso a una amplia y creciente variedad de documentación. Son ejemplos las placas Arduino, BeagleBone y Raspberry Pi. Muchas de ellas admiten la instalación de sistemas operativos libres como GNU Linux y por lo tanto toda clase de software de procesamiento matemático como Scilab, Octave o Euler, que permiten a su vez la realización de numerosos análisis como los que se describen en el presente libro.

<sup>42</sup> Probablemente la razón sea que los dispositivos de audio de alta fidelidad se han abaratado tanto que están al alcance de muchas personas, aun cuando no posean los conocimientos técnicos para interpretar una hoja de especificaciones.

niveles bajos es un resultado esperable debido a que la señal continua es reemplazada por una señal escalonada. El único problema encontrado es una distorsión en el extremo de las frecuencias altas (arriba de 10 kHz) que produce un efecto de aliasing. Este efecto desaparece si se trabaja con un máximo nivel de grabación de  $-20$  dB, al precio de reducir el rango dinámico a unos 70 dB. Este rango es, no obstante, suficiente en la gran mayoría de los casos, sobre todo en la medición de ruido urbano, donde el ruido de fondo es bastante considerable comparado con el ruido eléctrico del grabador.

### 3.15 Formatos de archivos de sonido

Cualquiera sea el medio de almacenamiento, la información digital correspondiente a las sucesivas muestras de audio debe atenerse a cierto formato comprensible para el usuario o el software que se utilice para su análisis. En términos generales los formatos pueden clasificarse en formatos con y sin compresión.

Los formatos sin compresión presentan los sucesivos datos PCM tal cual son generados por el conversor analógico digital, en algunos casos sin más información (extensiones .pcm, .raw); en otros, con un encabezamiento que proporciona datos como la resolución, número de canales, tasa de muestreo, cantidad de datos de audio (extensión .wav). El primer caso no es recomendable para archivos a compartir con terceros o destinados a su almacenamiento a largo plazo pues sin información suplementaria se tornan inservibles. Pero sí podrían emplearse como almacenamiento transitorio de una variable interna dentro de un software de procesamiento con el fin de liberar memoria. En el segundo caso el archivo contiene la información suplementaria y es por lo tanto fácilmente interpretable por cualquier otro usuario o software.

La ventaja de los formatos sin compresión es que no hace falta prácticamente ninguna decodificación para recuperar la información original. La desventaja es que su tamaño es muy grande. Por ejemplo, si se utiliza una tasa de muestreo de 44 100 Hz y una resolución de 16 bit (2 B) por muestra, resulta que se requieren 5,292 MB por cada minuto de audio o, lo que es equivalente, 315,52 MB por hora. Tal tasa de información puede suponer un problema a la hora de almacenar grandes cantidades de audio, sobre todo a largo plazo, o a la hora de transmitirlo a través de redes.

Los formatos comprimidos se utilizan para reducir la cantidad de datos necesarios para representar una determinada cantidad de audio y, por lo tanto, ahorrar espacio de almacenamiento o ancho de banda de transmisión. Todos los métodos de compresión se apoyan en que las señales de audio poseen un grado considerable de redundancia. Típicamente usan alguna estrategia de modelización de la señal que permite predecirla a partir de algunos datos anteriores. Dicho modelo usa pocos parámetros por lo que, si en lugar de almacenar la señal se almacenan los parámetros, será posible ahorrar espacio de almacenamiento. Desde luego, no se puede esperar que la predicción sea perfecta, por lo que el método de compresión se completa con una representación del error, que normalmente es pequeño y requiere menos bits.

La compresión puede ser, a su vez, con pérdida (lossy) o sin pérdida (lossless). Los formatos con pérdida, como .mp3 (MPEG-1 Layer 3), .ogg (OGG Vorbis) o ATRAC (propio del Minidisc), utilizan criterios psicoacústicos para descartar información que en condiciones normales resultaría inaudible para la mayoría de las personas. Son, por consiguiente, ideales para el almacenamiento en reproductores digitales portátiles y para la descarga desde repositorios en Internet de archivos de audio destinados a la escucha casual. Se consiguen reducciones en un factor de 5 con diferencias prácticamente imperceptibles y 10 o más con una calidad bastante respetable. Su utilización con fines metrológicos, sin embargo, no es recomendable en general, sobre todo en mediciones

especializadas como la del espectro de líneas, pues pueden suprimir regiones completas del espectro que, si bien son en principio inaudibles, podrían en su versión original contener tonos reveladores de mecanismos de interés en una fuente de ruido.

Los formatos de compresión sin pérdida se basan en consideraciones estadísticas y teoría de la información, y permiten una reconstrucción perfecta de la señal. Consiguen reducciones típicas en un factor cercano a 2. Entre los más conocidos se encuentra .flac (Free Lossless Audio Coding), que, a diferencia de otros formatos propietarios, es un formato abierto y de uso libre. Es, además, de codificación y decodificación muy rápida. Los formatos sin pérdida son muy utilizados para almacenamiento a largo plazo, repositorios en Internet de música y bases de sonidos y transmisión por redes, ya que en muchos casos permiten reducir prácticamente a un 50 % el uso de recursos. La única desventaja es que se requiere un algoritmo específico para su codificación y decodificación, aunque la carga computacional es suficientemente baja como para permitir su ejecución varias veces más rápido que el tiempo real (Marengo et al, 2011).

### 3.15.1 Formato WAV

El formato WAV (apócope del inglés *wave*, onda) es un tipo particular de formato de archivo para intercambio de recursos<sup>43</sup> (Resource Interchange File Format, RIFF). Si bien es ampliamente reconocido por la mayoría de los programas de edición de audio digital, así como por los lenguajes de programación, en algunos casos no es el formato nativo y se accede a él por importación o exportación.

El formato de un archivo .wav está dividido en *fragmentos* (en inglés, *chunks*),<sup>44</sup> cada uno de los cuales contiene un tipo específico de información, por ejemplo: formato, datos de audio, marcas (*cues*), etiquetas (*labels*), anotaciones (*notes*). Cada fragmento contiene un identificador, constituido por una cadena de caracteres,<sup>45</sup> y una serie de datos numéricos. Éstos están expresados en sistema de numeración hexadecimal mediante 1 a 4 bytes, donde cada byte (8 bit) representa dos cifras hexadecimales, es decir, tomadas del conjunto {0, 1, 2, ..., 9, A, ..., F}. Así, por ejemplo el byte 01101101 representa al hexadecimal 0x6D (la notación 0x se utiliza para indicar que lo que sigue es una representación hexadecimal), es decir,

$$\begin{array}{cc} \underbrace{0110}_{0x6} & \underbrace{1101}_{0xD} \\ & = 6 \times 16 + 13 = 109. \end{array}$$

La disposición de los bytes en el archivo utiliza el estilo numérico *little-endian*<sup>46</sup> de los procesadores Intel 80x86, según el cual los bytes se ordenan desde el menos significativo al más significativo (orden inverso). Así, el número hexadecimal 0xA34E se almacenará y/o transmitirá como 4EA3. Es importante tener en cuenta esto cuando se requiera interpretar los valores numéricos contenidos en un archivo .wav.

Los archivos .wav comienzan siempre con un encabezamiento que contiene el identificador “RIFF” (representado en código ASCII por 0x52494646), seguido por 4 bytes que representan la cantidad de bytes que siguen hasta el final (es decir, la longitud total

<sup>43</sup> Los archivos RIFF se utilizan para audio, video, imágenes, MIDI y otra información multimedial.

<sup>44</sup> La traducción literal de *chunk* es *trozo*. En relación con los archivos WAV se lo suele traducir como segmento o fragmento.

<sup>45</sup> Cada carácter se representa mediante un byte utilizando el código ASCII.

<sup>46</sup> La terminología *little-endian* y *big-endian* está inspirada en la novela “Los viajes de Gulliver” de Jonathan Swift. Los *Big-endians* eran los que se oponían al decreto del emperador según el cual los huevos debían romperse por el extremo más pequeño y seguían rompiéndolos, en la clandestinidad, por su extremo grande.

del archivo menos 8 bytes) y el identificador de tipo de archivo, es decir, “WAVE” (en ASCII, 0x57415645). Ver un ejemplo en la tabla 3.2.

A continuación aparece el fragmento correspondiente a formato (*format chunk*), que comienza con el identificador “fmt” (en ASCII, 0x666D7420) y la cantidad de bytes que siguen dentro del fragmento de formato. Los dos bytes que siguen codifican el tipo de formato. Para PCM sin compresión, el valor correspondiente es 1 (0x0100). Es el único caso de interés en metrología. Luego se tiene el número de canales en 2 bytes, la tasa de muestreo, en 4 bytes, la cantidad de bytes por segundo (igual al número de canales por la tasa de muestreo), en 4 bytes, la cantidad de bytes por muestra, en 2 bytes, y la cantidad de bits por muestra, en 2 bytes. Para formatos comprimidos podría haber campos adicionales. Nótese que algunos campos<sup>47</sup> son redundantes. Así, conociendo los bits por muestra basta dividir por 8 para obtener los bytes por muestra. A partir de la tasa de muestreo se deduce la cantidad de bytes por segundo simplemente multiplicando por el número de bytes por muestra y por la cantidad de canales.

El fragmento correspondiente a datos de audio (*data chunk*) comienza con el identificador “data” seguido por la longitud total en bytes de dichos datos (excluyendo el identificador y los 4 bytes de longitud). Luego vienen los datos de audio específicamente. En el caso de grabaciones monofónicas (un canal), aparecen tantos bytes por muestra como se indica en el fragmento de formato. En el caso de grabaciones estereofónicas, aparecerá el doble de bytes, correspondiendo los primeros al canal izquierdo y los segundos al derecho, es decir los canales están entrelazados. Si hubiera más canales se procede en forma similar.

En todos los casos, las longitudes de los datos en cada fragmento se expresan en bytes y excluyen el identificador previo. En el ejemplo de la tabla 3.2, la longitud del fragmento de formato es 16 bytes, que corresponden a los campos de tipo de formato, canales, tasa de muestreo, bytes por segundo, bytes por muestra y bits por muestra. Esto implica que la longitud real del fragmento es 8 bytes mayor.

La longitud de los datos de audio, en este caso 1 728 000, corresponde a la totalidad de los bytes, no al número de muestras. Si se desea obtener el número de muestras, deberá dividirse por la cantidad de bytes por muestra, en este caso, 3, y por la cantidad de canales, en este caso, 1. Por lo tanto tenemos  $1\,728\,000 / 3 = 576\,000$  muestras que, a una tasa de muestreo de 48 000 Hz corresponden a 12 s de audio.

La interpretación de los datos de audio depende del número de bits por muestra. En el caso de 8 bit por muestra, son enteros sin signo, lo cual significa que es necesario restar 128 para obtener el verdadero valor. Este caso es, en general, irrelevante para propósitos metrológicos, debido a la reducida relación señal / ruido. En el caso de más bits por muestra, por ejemplo, 16 bit, son enteros con signo en el formato denominado *complemento a 2*, según el cual los valores entre 0 y  $2^{n-1} - 1$  (que son los que tienen el bit más significativo igual a 0) se interpretan textualmente, y a los valores entre  $2^{n-1}$  y  $2^n - 1$  (bit más significativo igual a 1) se les resta  $2^n$ , lo cual corresponde a números negativos. Así, por ejemplo, en el caso de  $n = 16$ , tenemos, en escritura normal (*big-endian*):

$$\begin{aligned} 0x0000 &\equiv 0 \\ 0x0001 &\equiv 1 \\ 0x7FFF &= 2^{15} - 1 \equiv 32\,767 \\ 0x8000 &\equiv 2^{15} - 2^{16} = -32\,768 \\ 0xFFFF &\equiv 2^{16} - 1 - 2^{16} = -1 \end{aligned}$$

<sup>47</sup> Se denomina *campo* al espacio dentro de una estructura de datos destinado a alojar una variable

**Tabla 3.2.** Ejemplo de archivo .wav de 1 canal, tasa de muestreo 48 000 Hz y resolución 24 bit.. Sólo se muestran los dos primeros y el último dato de audio.

Fragmento	Tipo	Contenido	Valor	Decimal
Encabezamiento	Identificador	“RIFF”	0x52494646	----
	Longitud	4 bytes	0XC2601A00	1728706
	Identificador	“WAVE”	0x57415645	----
Formato	Identificador	“fmt ”	0x666D7420	----
	Longitud	4 bytes	0x10000000	16
	Tipo de formato	2 bytes	0x0100	1
	Canales	2 bytes	0x0100	1
	Tasa de muestreo	4 bytes	0x80BB0000	48000
	Bytes por segundo	4 bytes	0x80320200	144000
	Bytes por muestra	2 bytes	0x0300	3
	Bits por muestra	2 bytes	0x1800	24
Datos de audio	Identificador	“data”	0x64617461	----
	Longitud	4 bytes	0x005E1A00	1728000
	Valor de muestra	3 bytes	0x5B8E00	36443
	Valor de muestra	3 bytes	0xC97701	96201
	...	...	...	...
	Valor de muestra	3 bytes	0x0685f6	16155910

Los valores así obtenidos son números enteros con el 0 ubicado aproximadamente a mitad de camino entre el mínimo y el máximo. Si bien son proporcionales al valor instantáneo de la muestra correspondiente, difieren del valor real en un factor de escala que luego deberá ser aplicado para obtener resultados físicamente significativos. Dicho factor de escala involucra la sensibilidad del transductor,  $S$ , en V/Pa, la ganancia del sistema de acondicionamiento de señal (preamplificador),  $G$ , y la referencia de tensión del conversor analógico / digital,  $V_{\text{ref}}$ . Si el valor de presión sonora correspondiente a la muestra  $k$  es  $p_k$  y su representación digital con  $n$  bits es  $p_{Dk}$ ,

$$p_k = \frac{V_{\text{ref}}}{2^{n-1} S G} p_{Dk}, \quad (3.25)$$

es decir,

$$p_k = K p_{Dk}, \quad (3.26)$$

La ecuación (3.25) puede ser útil al diseñar el hardware de un sistema, pero en la práctica, cuando las grabaciones se realizan mediante equipos comerciales de arquitectura cerrada, las constantes  $S$ ,  $G$  y  $V_{\text{ref}}$  no se conocen. Incluso si están especificadas o pueden deducirse a partir de las especificaciones, están sujetas a tolerancias de fabricación y a derivas con el tiempo, la temperatura, etc., lo cual incrementa la incertidumbre. Es más apropiado registrar una señal de calibración de nivel especificado y, a partir de allí, obtener la constante  $K$ . Si llamamos  $p_{\text{cal}}$  al valor de calibración y  $p_{\text{cal}, D}$  al valor digital obtenido, entonces

$$K = \frac{P_{\text{cal}}}{P_{\text{cal}, D}}. \quad (3.27)$$

En lugar de la presión instantánea de una muestra particular, conviene utilizar su valor eficaz (ya que en general se trata de un tono puro del cual se conoce su nivel de presión sonora), y en lugar del valor digital de una muestra se toma el valor eficaz correspondiente a una duración de algunos segundos, obtenido como:

$$P_{\text{cal}, D, \text{ef}} = \sqrt{\frac{1}{N} \sum_{k=1}^N P_{\text{cal}, D}^2(k)}. \quad (3.28)$$

Resulta, entonces,

$$K = \frac{P_{\text{cal}, \text{ef}}}{P_{\text{cal}, D, \text{ef}}}. \quad (3.29)$$

Es conveniente elegir  $N$  muy próximo a un número entero de ciclos, ya que de esa manera se minimiza el error residual oscilante que se mostraba en la figura 1.3. Si además  $N$  es grande, se logra reducir la incertidumbre hasta un valor muy próximo a la del tono de calibración, ya que en la práctica equivale a promediar una gran cantidad de mediciones.

La información anterior (incluyendo la totalidad de los datos de audio) es suficiente para constituir un archivo de audio digital válido. Sin embargo, resulta útil poder almacenar en el archivo cierta información adicional que no pertenece estrictamente al audio. Se trata de *marcas (cues)*, *etiquetas (labels)* y *anotaciones (notes)* relativas a eventos acústicos de interés, por ejemplo identificación de fuentes, datos sobre horario y lugar de medición, tonos de calibración, señales de sincronismo entre varias mediciones, etc. Esta información se conoce genéricamente como *metadatos*, es decir, datos que se refieren a otros datos, describiéndolos, clasificándolos o calificándolos.<sup>48</sup>

Mientras el archivo original obtenido por el proceso de grabación se genera en forma automática, esta nueva información será agregada interactivamente por el operador a posteriori, mediante componentes del software de edición provistos a tal fin. Estos componentes ofrecen ventanas de diálogo donde se pueden agregar marcas y editar sus correspondientes etiquetas y anotaciones.

La forma de generar la información a agregar varía según el protocolo de medición adoptado. Una manera es grabar la expresión oral de los eventos, ya sea en el mismo canal durante intervalos irrelevantes (por ejemplo antes o después de una calibración) o bien en un canal independiente sincronizado. Otra forma es llevar un meticuloso registro escrito mediante planillas donde se anotará el horario y la descripción de los eventos. En ambos casos, luego se pasará en limpio dicha información al archivo mediante las facilidades ya citadas que posea cada software de edición. Algunos equipos de grabación permiten el agregado en tiempo real de marcas, aunque no siempre informan el número correlativo de marca para vincularla a posteriori con una anotación manual específica en la planilla, por lo cual el uso de esta facilidad presenta riesgo de errores y no se recomienda.

Un buen protocolo de medición permitirá el aprovechamiento automático de la información almacenada en el propio archivo en forma de metadatos, especialmente si se

<sup>48</sup> El formato WAV permite otros metadatos, tales como autor, intérprete, copyright, etc., pensados para la industria de producción fonográfica, que en este caso son irrelevantes.

adopta para los mismos un formato normalizado que incluya descriptores y campos de datos bien definidos y uniformes. Por ejemplo, si los tonos de calibración se han delimitado con marcas apropiadas a las que se asocia el nivel de presión sonora correspondiente, será posible desarrollar un algoritmo que aplique automáticamente los datos de calibración a fin de obtener niveles equivalentes, espectros, etc.

Dado que la gran mayoría del software existente está orientado más bien al proceso de edición con fines de producción de música, bandas de sonido o locuciones, las marcas corresponden típicamente a puntos de edición, inserción o sustitución y son siempre destinadas a procesos interactivos. Para nuestro propósito metrológico, en cambio es necesario poder acceder o incluso generar o modificar automáticamente los metadatos, de allí la necesidad de conocer su organización dentro del archivo. En las tablas 3.3 a 3.5 se muestran ejemplos de la parte de un archivo WAV correspondiente a los metadatos.

El primer elemento (ver tabla 3.3) son las marcas (cues). Las marcas aparecen en el archivo en un nuevo fragmento (*cue chunk*) ubicado en general inmediatamente después de los datos de audio. Es posible conocer dónde comienza localizando el identificador “cue” luego de finalizados los datos de audio, lo cual sucede una vez completada una cantidad de bytes igual a la especificada en el fragmento de datos de audio (en el ejemplo de la tabla 3.2, luego de 1 728 000 bytes a partir del byte siguiente al campo de longitud del fragmento de datos). A continuación del identificador del fragmento aparecen 4 bytes que contienen la longitud del fragmento de ahí en más.

**Tabla 3.3.** Ejemplo de distribución de marcas (cue). Esta información aparece inmediatamente después del último dato de audio que se muestra en la tabla 3.2.

Fragmento	Tipo	Contenido	Valor	Decimal
Cue	Identificador	“cue ”	0x63756520	----
	Longitud	4 bytes	0xC4000000	196
	Cantidad de cues	4 bytes	0x41000000	8
	Nº de cue	4 bytes	0x01000000	1
	Muestra asociada	4 bytes	0x61C00000	52321
	Identificador	“data”	0x64617461	----
	Inicio de fragmento	4 bytes	0x00000000	0
	Inicio de bloque	4 bytes	0x00000000	0
	Offset	4 bytes	0x61C00000	52321
	Nº de cue	4 bytes	0x02000000	2
	Muestra asociada	4bytes	0x43C10100	115011
	Identificador	“data”	0x64617461	----
	Inicio de fragmento	4 bytes	0x00000000	0
	Inicio de bloque	4 bytes	0x00000000	0
	Offset	4 bytes	0x43C10100	115011
	...	...	...	...
	Nº de cue	4 bytes	0x08000000	8
	Muestra asociada	4bytes	0xA73C0500	343207
	Identificador	“data”	0x64617461	----
	Inicio de fragmento	4 bytes	0x00000000	0
	Inicio de bloque	4 bytes	0x00000000	0
	Offset	4 bytes	0xA73C0500	343207



Luego aparece la información de las marcas: 4 bytes para el número de marca (en orden correlativo de creación, que no tiene por qué coincidir con el orden cronológico, ya que se podría introducir una marca en un instante anterior a la última marca existente) y 4 bytes para el número de muestra asociada, que indica su posición en el tiempo. Luego, precedidos por el identificador “data” (que no debe confundirse con el correspondiente al fragmento de datos de audio), aparecen tres datos de 4 bytes cada uno que sólo tienen relevancia en el caso de formatos comprimidos o en el caso de archivos que contienen en realidad varios bloques sonoros independientes.

A continuación aparece una *lista de datos asociados* (en inglés, *associated data list*, *ADTL*) que contendrá las etiquetas, anotaciones e intervalos asociados a cada marca (ver tabla 3.4). Dicha lista está encabezada por el identificador “LIST”, seguido por la longitud de los datos que siguen y el identificador “adtl”.

Luego aparecen tres tipos de fragmentos. El primer tipo, que comienza con el identificador “ltxt” (*labeled text chunk*) seguido por la cantidad de bytes del segmento (sin incluir el identificador ni los 4 bytes de longitud, contiene el número de marca y la *cantidad de muestras del intervalo asociado a la marca*. Este valor puede ser 0 o cualquier valor compatible con la duración restante del archivo. Es útil para delimitar intervalos de interés, por ejemplo, un tono de calibración, un intervalo de integración o un evento como el paso de un vehículo, el canto de un pájaro o un ladrido. Luego de esta información aparecen unos datos irrelevantes para nuestra aplicación metrológica, ya que en realidad se proveen para compatibilidad con otros formatos.<sup>49</sup> Existen tantos fragmentos ltxt como marcas.

El segundo tipo de fragmento que aparece en la lista de datos asociados es el correspondiente a *etiquetas* (ver tabla 3.5) Comienza con el identificador “labl” (*label chunk*) seguido por la cantidad de bytes del segmento (sin incluir el identificador ni los 4 bytes de longitud), y sigue con el número de marca y el contenido, es decir una cadena alfanumérica que contiene el texto de la etiqueta en codificación ASCII (admite acentos). La etiqueta es arbitraria y no tiene por qué guardar ninguna relación con el número de marca a la que se asocia, como se puede apreciar en el ejemplo de la tabla 3.5. En todos los casos se requiere que la terminación sea un byte nulo (0x00), pero como además la cantidad total de bytes debe ser par, en caso de resultar impar se agrega otro byte nulo de emparejamiento (ver el fragmento correspondiente a la marca 2 en la tabla 3.5). El byte de emparejamiento no es tenido en cuenta en la longitud declarada. Por ejemplo, la longitud declarada del segundo fragmento de la tabla 3.5 (correspondiente a la marca 2) es 13 bytes, pero incluyendo el byte de emparejamiento resultan 14 bytes.

Cabe aclarar que sólo aparecen fragmentos de etiqueta asociados a etiquetas no vacías, de modo que si alguna marca no tiene etiqueta el correspondiente fragmento no aparece.

El tercer tipo de fragmento en la lista de datos asociados es el correspondiente a *notas* o *anotaciones* (ver tabla 3.6), que es muy similar a los fragmentos de etiquetas. Comienza con el identificador “note” (*note chunk*) seguido por la cantidad de bytes del segmento (sin incluir el propio identificador ni los 4 bytes de longitud), y sigue con el número de marca y el contenido, es decir una cadena alfanumérica que contiene el texto de la anotación en código ASCII. En general, las anotaciones son de mayor extensión que las etiquetas, brindando información detallada sobre la marca a la que se asocia (o sobre el intervalo asociado), por ejemplo una fecha, un horario, una posición GPS. Nuevamente, en todos los casos se requiere que la terminación sea un byte nulo (0x00), y

<sup>49</sup> Estos datos son el identificador de propósito del texto, en este caso “rgn ” (región), y códigos de país, idioma y dialecto, código de página y el texto, que en el caso de audio no aparece.

como además la cantidad total de bytes debe ser par, en caso de resultar impar se agrega otro byte nulo de emparejamiento. Al igual que en el caso de las etiquetas, el byte de emparejamiento, si está, no entra en la longitud declarada. Sólo aparecen fragmentos de etiqueta asociados a etiquetas no vacías.

**Tabla 3.4.** Ejemplo de fragmentos ltxt de cantidad de muestras asociada a cada marca. Esta información está inmediatamente después del último dígito en la tabla 3.3. Sólo la marca 3 tiene asociado un intervalo no nulo.

Fragmento	Tipo	Contenido	Valor	Decimal
Lista de datos asociados	Identificador	“LIST”	0x4C495354	----
	Longitud	4 bytes	0xCA010000	458
	Identificador	“adtl”	0x6164746C	----
Longitudes de las marcas	Identificador	“ltxt”	0x6C747874	----
	Longitud	4 bytes	0x14000000	20
	Número de marca	4 bytes	0x01000000	1
	Cantidad de muestras	4 bytes	0x00000000	0
	Propósito del texto	“rgn ”	0x72676E20	----
	País	2 bytes	0x0000	0
	Idioma	2 bytes	0x0000	0
	Dialecto	2 bytes	0x0000	0
	Código de página	2 bytes	0x0000	0
	...	...	...	...
	Identificador	“ltxt”	0x6C747874	----
	Longitud	4 bytes	0x14000000	20
	Número de marca	4 bytes	0x03000000	3
	Cantidad de muestras	4 bytes	0xA6BD0000	48550
	Propósito del texto	“rgn ”	0x72676E20	----
	País	2 bytes	0000	0
	Idioma	2 bytes	0000	0
	Dialecto	2 bytes	0x0000	0
	Código de página	2 bytes	0x0000	0
	...	...	...	...
	Identificador	“ltxt”	0x6C747874	----
	Longitud	4 bytes	0x14000000	20
	Número de marca	4 bytes	0x08000000	8
	Cantidad de muestras	4 bytes	0x00000000	0
	Propósito del texto	“rgn ”	0x72676E20	----
	País	2 bytes	0x0000	0
	Idioma	2 bytes	0x0000	0
	Dialecto	2 bytes	0x0000	0
	Código de página	2 bytes	0x0000	0

**Tabla 3.5.** Ejemplo de fragmentos “labl” de etiquetas asociadas a cada marca. Esta información aparece inmediatamente después del último byte en la tabla 3.4. Sólo aparecen fragmentos con contenido no vacío.

Fragmento	Tipo	Contenido	Valor	Decimal
Etiquetas	Identificador	“labl”	0x6C61626C	---
	Longitud	4 bytes	0x0C000000	12
	Número de marca	4 bytes	0x01000000	1
	Contenido	“Marca 1”	0x4D617263612031	----
	Terminación nula	1 byte	0x00	----
	Identificador	“labl”	0x6C61626C	---
	Longitud	4 bytes	0x0D000000	13
	Número de marca	4 bytes	0x02000000	2
	Contenido	“Marca 1a”	0x4D61726361203161	----
	Emparejador	1 byte	0x00	----
	Terminación nula	1 byte	0x00	----
	Identificador	“labl”	0x6C61626C	---
	Longitud	4 bytes	0x0D000000	12
	Número de marca	4 bytes	0x03000000	3
	Contenido	“Marca 4”	0x4D617263612034	----
	Terminación nula	1 byte	0x00	----
	...	...	...	...
	Identificador	“labl”	0x6C61626C	---
	Longitud	4 bytes	0x0D000000	12
	Número de marca	4 bytes	0x08000000	8
	Contenido	“Marca 6”	0x4D617263612036	----
	Terminación nula	1 byte	0x00	----

### 3.15.2 Formato FLAC

FLAC (Free Lossless Audio Codec, codificador-decodificador de audio sin pérdidas libre<sup>50</sup>) es un formato comprimido sin pérdida que es conveniente para el almacenamiento permanente o documental y para la transmisión de grandes cantidades de audio, ya que permite ahorrar hasta alrededor de un 50 % de capacidad de memoria o de ancho de banda del canal de comunicación, por ejemplo una red de área local inalámbrica o Internet (Salomon, 2007; Marengo et al., 2011; Roveri, 2011; Raynaudo et al., 2013). A diferencia del formato WAV, requiere el uso de un *códec* (codificador-decodificador) para comprimir la señal original PCM y posteriormente recuperarla cuando se requiera.

El método de compresión consiste en subdividir la señal en bloques temporales, aproximar en cada uno de ellos la señal con un modelo predictivo que utilice pocos parámetros, que se incluyen en un encabezamiento, y luego codificar el error de predicción, es decir la diferencia entre la señal original y su aproximación, que es más pequeña que la señal y por lo tanto requiere menos bits.

<sup>50</sup> Este formato se inscribe en la corriente del *software libre*, es decir un software de libre licencia, código abierto y distribución gratuita

**Tabla 3.6.** Ejemplo de fragmentos “note” de notas asociadas a cada marca. Esta información aparece inmediatamente después del último byte en la tabla 3.5. Sólo aparecen fragmentos con contenido no vacío.

Fragmento	Tipo	Contenido	Valor	Decimal
Notas	Identificador	“note”	0x6E6F7465	---
	Longitud	4 bytes	0x0D000000	13
	Número de marca	4 bytes	0x01000000	1
	Contenido	“12:22:35”	0x31323A32323A3335	----
	Emparejador	1 byte	0x00	----
	Terminación nula	1 byte	0x00	----
	Identificador	“note”	0x6E6F7465	---
	Longitud	4 bytes	0x13000000	19
	Número de marca	4 bytes	0x03000000	3
	Contenido	“48550 muestras”	0x3438353530206D75657374726173	----
	Emparejador	1 byte	0x00	----
	Terminación nula	1 byte	0x00	----
	Identificador	“note”	0x6E6F7465	----
	Longitud	4 bytes	0x0B000000	11
	Número de marca	4 bytes	0x08000000	8
	Contenido	“Máximo”	0x4DE178696D6F	----
	Emparejador	1 byte	0x00	----
	Terminación nula	1 byte	0x00	----

Esta descripción, sin embargo, está muy simplificada, ya que lo que en realidad sucede es que *estadísticamente* hay más valores pequeños que altos. En una codificación PCM, la cantidad de bits está determinada por el máximo valor a representar, por lo cual, con que haya un solo error de máxima amplitud ya será necesario utilizar la máxima cantidad de bits por muestra.

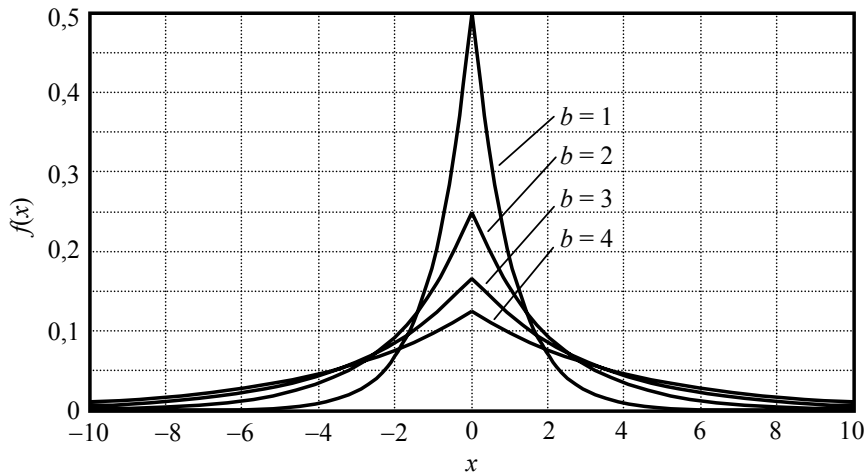
En lugar de esto conviene utilizar un *código de longitud variable* que utilice menos bits para los valores más frecuentes, es decir, los más pequeños. La elección de tal código queda determinada por la distribución estadística de los valores a codificar. Se ha encontrado empíricamente que, para los modelos predictivos típicos, los errores (denominados *residuos*) responden aproximadamente a una distribución de Laplace<sup>51</sup> (ver figura 52) cuya función de densidad de probabilidad es

$$f(x) = \frac{1}{2b} e^{-\frac{|x - \mu|}{b}}, \quad (3.30)$$

<sup>51</sup> Estrictamente, la distribución no es laplaciana, porque al estar trabajando con variables discretas, el error es discreto. Lo que se observa es que el histograma tiene una forma similar a la distribución de Laplace.

donde  $\mu$  es la media y  $b$  es un parámetro de escala proporcional al desvío estándar:

$$\sigma = \sqrt{2} b. \quad (3.31)$$



**Figura 52.** Función de densidad de probabilidad de una distribución de Laplace de media 0 para cuatro valores del parámetro  $b$ .

### 3.15.2.1 Código de Golomb-Rice

Para una variable con distribución laplaciana, uno de los mejores códigos es el de Golomb-Rice. Se trata de un código híbrido, en el cual se divide el número a codificar por un divisor  $D$  convenientemente elegido (en general una potencia de 2). El cociente entero se representa en código unario y el resto en binario. En principio, el código *unario* representa los números con un solo símbolo repetido tantas veces como el número a representar. Sin embargo, dado que su longitud es variable y a priori desconocida, se necesita un símbolo diferente que represente la finalización del número. En versión tipográfica el símbolo podría ser un espacio, pero en un sistema binario el símbolo principal y el de separación deben elegirse entre el 0 y el 1. Si elegimos el 0 como el símbolo a repetir y el 1 como terminación, el 3 se representaría como 0001.

En la tabla 3.7 se muestra, como ejemplo, el código de Golomb-Rice con divisor  $D = 4$  para números entre 0 y 15. Vemos que si bien los valores más altos requieren 6 bits cuando podrían escribirse con sólo 4, los más bajos requieren apenas 3 bits. La idea es que si los valores pequeños son mucho más frecuentes que los más altos, el código conseguirá reducir la cantidad media de bits requeridos. Esto es precisamente lo que sucede en el caso de una distribución laplaciana.

Nótese que, si se conoce el divisor, una cadena de 0's y 1's que representa una sucesión de números puede decodificarse unívocamente. Dada, por ejemplo, la cadena

00100001010100101101010000111

podemos reescribirla, para mayor claridad, como

00100 | 00101 | 0100 | 101 | 101 | 0100 | 00111

Para ello tuvimos en cuenta que el primer número comienza con una cantidad de 0's seguida por un 1 (símbolo separador), y luego dos bits arbitrarios que representan el resto de la división. En este caso hay dos 0's y un 1, seguido por 00. Esto puede repetirse mientras el código sea coherente, es decir, provenga de un codificador de Golomb-Rice. De acuerdo al significado del código, esto implica que los números son

$$2 \times 4 + 0, 2 \times 4 + 1, 1 \times 4 + 0, 0 \times 4 + 1, 0 \times 4 + 1, 1 \times 4 + 0, 2 \times 4 + 3,$$

es decir

$$8, 9, 4, 1, 1, 4, 11$$

**Tabla 3.7.** Ejemplo decodificación de Golomb-Rice para números entre 0 y 15 con divisor  $D = 4$ . El espacio de separación entre el cociente en representación unaria y el resto en representación binaria es al solo efecto de facilitar el análisis visual de la tabla.

$N$	Binario	Golomb-Rice
0	0000	1 00
1	0001	1 01
2	0010	1 10
3	0011	1 11
4	0100	01 00
5	0101	01 01
6	0110	01 10
7	0111	01 11
8	1000	001 00
9	1001	001 01
10	1010	001 10
11	1011	001 11
12	1100	0001 00
13	1101	0001 01
14	1110	0001 10
15	1111	0001 11

La elección del divisor  $D$  depende básicamente del parámetro  $b$  de la distribución de Laplace. Cuanto más pequeño sea  $b$ , mayor será la proporción de valores muy pequeños, por lo tanto convendrá un divisor pequeño. Dado que  $D$  se elige como una potencia de 2, es decir,

$$D = 2^d, \quad (3.32)$$

basta especificar  $d$ , denominado *parámetro de Rice*. FLAC utiliza la siguiente estimación para el valor de  $d$  óptimo (Salomon, 2007; Robinson, 1994):

$$d = \log_2(\ln 2 \ E(|e(k)|)), \quad (3.33)$$

donde  $E(|e(k)|)$  es la esperanza de los errores absolutos de predicción. Para una distribución laplaciana de los errores se cumple

$$E(|e(k)|) = b = \frac{\sigma}{\sqrt{2}}. \quad (3.34)$$

Si, por ejemplo, el desvío estándar de los errores fuera 8, resulta  $d \cong 2$ , de donde  $D = 2^2 = 4$ , que corresponde al ejemplo anterior.

### 3.15.2.2 Modelización predictiva

Con respecto la estrategia de modelización, FLAC utiliza uno de cuatro modelos para predecir la evolución de la señal: 1) la propia señal; 2) un valor constante; 3) una aproximación polinomial; y 4) una aproximación por código de predicción lineal (linear predictive coding, LPC). El modelo elegido es el que comprime más cada bloque.

### 3.15.2.3 Modelización mediante la propia señal (*verbatim*)

Este caso es apto para señales completamente aleatorias, por ejemplo, ruido blanco. En este caso, donde la señal tiene máxima entropía,<sup>52</sup> carece de sentido comprimir ya que al no haber redundancia (es decir, cualquier modelo da errores muy grandes) la compresión será mínima y probablemente quede desvirtuada por la necesidad de incluir los parámetros del modelo. A esto se agrega la carga computacional de codificar y decodificar sin que redunde en una compresión apreciable.

### 3.15.2.4 Modelización constante

Este caso se aplicaría principalmente a intervalos de silencio, que pueden comprimirse simplemente contando las muestras iguales a 0. Este tipo de compresión se conoce como compresión *run-length* y es la que se utiliza en los compresores tipo .zip para archivos generales o .gif para archivos de imagen.

### 3.15.2.5 Modelización polinomial

Este modelo consiste en obtener un polinomio de grado  $n$  que pasa por las  $n + 1$  muestras anteriores ( $n = 0, \dots, 3$ ), extrapolando al instante actual para aproximar el correspondiente valor de la señal (figura 53). Dado que las muestras están separadas entre sí por tiempos uniformes, se puede demostrar que los polinomios adoptan una forma tal que lleva a una expresión muy sencilla del valor de predicción y del error. Llamando  $t$  al tiempo discreto, se trata de predecir la muestra  $y(k)$  en función de las  $n + 1$  muestras anteriores, es decir,  $y(k - 1), \dots, y(k - n - 1)$  mediante un polinomio de grado  $n$ .

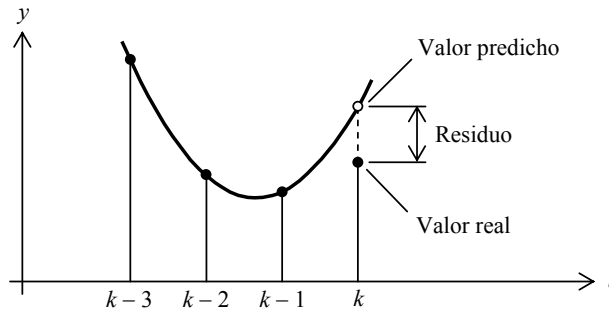
Por ejemplo, para  $n = 2$ , buscamos los coeficientes de un polinomio cuadrático

$$P(t) = a_0 + a_1(t - k) + a_2(t - k)^2 \quad (3.35)$$

tal que

$$\begin{aligned} P(k - 1) &= y(k - 1) \\ P(k - 2) &= y(k - 2) \\ P(k - 3) &= y(k - 3) \end{aligned} \quad (3.36)$$

<sup>52</sup> La entropía de una señal es un indicador de su impredecibilidad y, por lo tanto, de su falta de redundancia. A mayor entropía, menor redundancia, por lo que la compresión será menos efectiva. Para señales discretas se define como  $H = -\sum p_k \log_2(p_k)$ , donde  $p_k$  es la probabilidad del  $k$ -ésimo valor discreto de la señal. Se interpreta como la mínima cantidad media de bits por muestra alcanzable por un código. Los códigos reales en general requieren una mayor tasa de información.



**Figura 53.** Modelo polinomial de segundo grado.

Sustituyendo (3.36) en (3.35) resulta el sistema de ecuaciones

$$\begin{aligned} a_0 - a_1 + a_2 &= y(k-1) \\ a_0 - 2a_1 + 4a_2 &= y(k-2), \\ a_0 - 3a_1 + 9a_2 &= y(k-3) \end{aligned} \quad (3.37)$$

cuya solución es

$$\begin{aligned} a_0 &= 3y(k-1) - 3y(k-2) + y(k-3) \\ a_1 &= \frac{5}{2}y(k-1) - 4y(k-2) + \frac{3}{2}y(k-3) \\ a_2 &= \frac{1}{2}y(k-1) - y(k-2) + \frac{3}{2}y(k-3) \end{aligned} \quad (3.38)$$

El valor predicho en  $t = k$  es

$$\hat{y}(k) = P(k) = a_0 = 3y(k-1) - 3y(k-2) + y(k-3) \quad (3.39)$$

Esta aproximación se denomina de orden 3 ya que usa 3 muestras anteriores. En forma similar se obtienen las aproximaciones polinomiales de los diversos grados, según se indica en la tabla 3.8. Nótese que si bien las aproximaciones son polinomiales, el resultado es, salvo en el caso de orden 0, una *combinación lineal* de las muestras anteriores.

Dado que el concepto de “muestras anteriores” depende del instante específico en el que se quiere predecir una muestra, el modelo se actualiza muestra a muestra. Sólo sirve, por lo tanto, para predecir una muestra al solo efecto de obtener el error de predicción o residuo.

Es natural preguntarse qué grado conviene elegir en cada caso. La respuesta es que el codificador prueba con todos y elige el que produce el menor residuo absoluto acumulado. Aunque puede parecer que es un método de fuerza bruta con alto costo computacional, un hecho interesante es que los residuos para el grado  $n$  se pueden obtener a partir de los del grado  $n-1$ . Llamando  $e_n(k)$  al residuo en la predicción de la muestra  $y(k)$  con un polinomio de grado  $n$ , es decir



**Tabla 3.8.** Expresiones de las aproximaciones polinomiales hasta el grado 3.

Orden	Grado	$\hat{y}(k)$
0	-	0
1	0	$y(k-1)$
2	1	$2 y(k-1) - y(k-2)$
3	2	$3 y(k-1) - 3 y(k-2) + y(k-3)$
4	3	$4 y(k-1) - 6 y(k-2) + 4 y(k-3) - y(k-4)$

$$e_n(k) = y(k) - \hat{y}_n(k), \quad (3.40)$$

entonces resulta

$$\begin{aligned} e_0(k) &= y(k) - y(k-1) \\ e_1(k) &= e_0(k) - e_0(k-1) \\ e_2(k) &= e_1(k) - e_1(k-1) \\ e_3(k) &= e_2(k) - e_2(k-1) \end{aligned} \quad (3.41)$$

Este proceso iterativo permite, por consiguiente, calcular los residuos en forma computacionalmente eficiente.

Una vez determinado el grado  $n$  del polinomio que proporciona el mínimo residuo acumulado, el único parámetro que requiere el decodificador es el grado del polinomio.

### 3.15.2.6 Modelización LPC

La modelización por predicción lineal es una variante más general de la modelización polinomial. Consiste en expresar el valor predicho como combinación lineal de  $p$  muestras anteriores:

$$\hat{y}(k) = \sum_{h=1}^p a_h y(k-h). \quad (3.42)$$

La diferencia con el modelo polinomial está en que los coeficientes  $a_h$  no son fijos sino que se determinan a partir de la señal completa dentro del bloque, eligiéndolos de modo de minimizar el error cuadrático de predicción

$$e^2 = \sum_{k=1}^N \left( y(k) - \sum_{h=1}^p a_h y(k-h) \right)^2, \quad (3.43)$$

donde  $N$  es la cantidad de muestras del bloque. La determinación se realiza mediante el algoritmo de Levinson-Durbin (Proakis et al., 1998).

El usuario puede optar por un valor máximo de  $p$  entre 1 y 32. El codificador determina el valor óptimo de  $p$  a partir del rango dinámico y el tamaño del bloque (Salomon, 2007). Cuanto mayor sea  $p$  mayor es el tiempo de cómputo. Se observa que para  $p \geq 10$  la mejora es despreciable y en general no se justifica ante el mayor tiempo de cómputo.

La minimización del error cuadrático, en el caso de una distribución laplaciana implica también la minimización del error absoluto medio (ecuación (3.34)), por lo cual se aprovecha mejor el código de Golomb-Rice. El precio a pagar es que se requiere una mayor cantidad de información, ya que se deben proporcionar los  $p$  coeficientes del modelo. De todas maneras, para poder trabajar con números enteros (la única manera de no depender de la implementación aritmética de cada procesador) los coeficientes LPC se cuantizan, por lo que el óptimo no es, en general, alcanzado.

### 3.15.2.7 Decorrelación intercanal

Hasta aquí analizamos la técnica de compresión para un único canal de audio. En el caso de dos canales estéreo, en general existe mucha redundancia entre las señales del canal izquierdo,  $y_I$ , y derecho,  $y_D$ , que puede aprovecharse para reducir aún más la tasa de información requerida. Ello se puede lograr reemplazando la señal mediante la transformación izquierdo-derecho a central-lateral, es decir, la semisuma y la diferencia:

$$\begin{aligned} y_C &= \frac{y_I + y_D}{2} \\ y_L &= y_I - y_D \end{aligned} \quad (3.44)$$

Si las señales son parecidas,  $y_L$  resulta más pequeña, lo cual permite una mayor compresión. Sin embargo, ello no siempre sucede, por lo cual FLAC analiza las dos versiones,  $(y_I, y_D)$ ,  $(y_C, y_L)$ , codificando finalmente la que logra mejor compresión. Los canales izquierdo y derecho pueden recuperarse mediante las expresiones

$$\begin{aligned} y_I &= \frac{2y_C + y_L}{2} \\ y_D &= \frac{2y_C - y_L}{2} \end{aligned} \quad (3.45)$$

Dado que FLAC trabaja con enteros, si  $y_I + y_D$  es impar la división por 2 de las ecuaciones (3.44) da fraccionaria y debe truncarse, lo cual llevaría a una recuperación imperfecta. Para evitarla, se puede restar 1 a  $y_I$  durante la conversión a  $(y_C, y_L)$  y luego, si  $y_L$  resulta impar, compensar sumando 1.

### 3.15.2.8 Estructura del archivo FLAC

El método de compresión descrito requiere una serie de parámetros que determinan el significado preciso de los datos comprimidos. Entre ellos se encuentran las características del muestreo (tasa de muestreo, resolución, número de canales), el tipo de modelo adoptado, sus variantes y sus parámetros. A ello se agrega una serie de metadatos que incluyen comentarios, datos de la aplicación de software que genera el archivo (por ejemplo el códec de referencia), datos que identifican las pistas de audio (como en un CD), tablas que facilitan el proceso de búsqueda (ya que al ser la codificación de los residuos de longitud variable y a priori desconocida para el decodificador la búsqueda

requiere la decodificación previa), entre otros. También es posible incluir imágenes embebidas en el propio archivo, por ejemplo las carátulas del envoltorio de un CD, aunque en el caso metrológico podrían ser fotografías del punto de medición o diagramas que muestran su posición en un plano o mapa.

Los detalles del formato que permiten implementar la decodificación de un archivo FLAC se desarrollan en el Apéndice 11.

### 3.16 Archivos de proyecto

Algunos programas de edición de audio digital, como Audacity o Cubase, trabajan con el concepto de *proyecto*. Un proyecto es una estructura de archivos que contienen el audio digital y otra información asociada. Una ventaja del enfoque de proyecto para la tarea de edición de audio es que es posible almacenar algunos *eventos de edición*, lo cual permite un registro y documentación de los procesos efectuados. La principal desventaja es que los formatos de proyecto son propietarios o específicos para cada programa, por lo cual no son interpretados por otros programas. Esto se soluciona con las funciones de *exportación e importación*, que permiten generar o leer archivos de formatos ampliamente difundidos, como los ya vistos WAV y FLAC.

Nos referiremos aquí específicamente al caso de Audacity, por ser éste un software libre, multiplataforma (tiene implementaciones en Linux, Windows y Mac OS X, así como en otros sistemas operativos) y de desarrollo actualmente (2013) muy activo.

Los proyectos en Audacity constan de un archivo con extensión .aup (Audacity Project) y un directorio o carpeta cuyo nombre es igual al nombre del proyecto seguido por \_data. Por ejemplo, el proyecto prueba.aup estará acompañado por el directorio prueba\_data. Dicho directorio contiene una estructura de subdirectorios dentro de los cuales se encuentran los datos de audio en un conjunto de archivos con extensión .au o bien archivos asociados con extensión .auf a archivos de audio externos.

#### 3.16.1 Formato .AUP

Un archivo .au es un archivo de texto en formato XML<sup>53</sup>. Como todo documento XML, está formado por unidades sintácticas denominadas *elementos*, cada uno de los cuales está afectado por unos modificadores denominados *atributos*. A su vez cada elemento puede contener otros elementos anidados (llamados *hijos*). Hay un elemento principal que los contiene a todos, denominado *project*, entre cuyos atributos figura el estado de las ventanas de visualización de las diferentes pistas (posición horizontal, límites de la selección y porcentaje de zoom) la última vez que se guardó, así como la tasa de muestreo del proyecto. Nótese que ésta podría ser diferente de las tasas de muestreo de las distintas pistas, que a su vez podrían diferir entre sí, especialmente cuando se importan en el proyecto archivos de diversos orígenes. Esta posibilidad, sin embargo, no es demasiado prolija ni conveniente en aplicaciones metrológicas, ya que la exportación de un único archivo de sonido correspondiente al proyecto implica operaciones de remuestreo que podrían introducir alteraciones en la señal original. La especificación del formato .aup se detalla en el apéndice 12.

<sup>53</sup> El formato XML (eXtensible Markup Language) es un metalenguaje de marcas apto para definir lenguajes específicos (por ejemplo el aquí utilizado, o los utilizados para documentos de texto, para partituras musicales o para bases de datos). Un lenguaje de marcas es un código que incluye texto y metadatos que indican cuestiones estructurales, semánticas o de formato del texto. Cada archivo XML hace referencia a una definición de tipo de documento (*document type definition*, DTD) que establece la sintaxis del mismo, es decir, la descripción de cada elemento y su secuencia lógica y jerárquica. Ver detalles en el Apéndice 12.

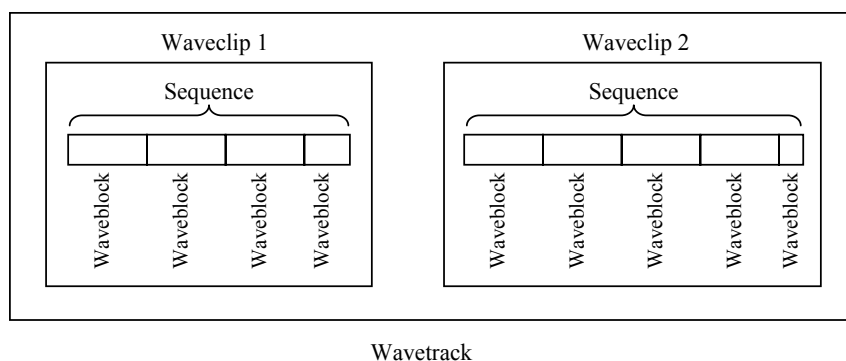
Dentro del elemento *project* aparecen, subordinados, elementos *tags*, *timetrack*, *wavetrack* y *labeltrack*. En *tags* se incluyen metadatos característicos de la industria discográfica, tales como artista, género de la música, etc. En su estado actual carecen de soporte para otros datos que serían de mayor interés para fines metrológicos, como localización espacial y temporal, posición de micrófonos, etc., pero tratándose de software libre es factible que la comunidad colabore para lograr incorporarlos en futuras versiones.

El elemento *timetrack* se refiere a la pista de control de tiempo. Ésta define una envolvente temporal que permite modificar la velocidad de reproducción. Si bien para aplicaciones musicales es una herramienta muy interesante, no reviste interés metrológico, ya que introduce severas alteraciones en el tiempo y en el espectro.

Cada elemento *wavetrack* corresponde a una pista de audio. Contiene el nombre de la pista, el canal de audio al que en principio está asociada,<sup>54</sup> y la información de si la pista está configurada como *solo*<sup>55</sup> o si tiene aplicada *sordina*.<sup>56</sup> También indica la altura en píxeles de la ventana correspondiente a la pista, si la misma se encuentra minimizada y si está seleccionada. Por último, contiene la tasa de muestreo de la pista, la ganancia aplicada, y el *paneo*, como número entre -1 y 1, donde los valores negativos corresponden al canal izquierdo y los positivos al derecho (0 implica que la señal se envía por partes iguales a ambos canales).

En lugar de que cada elemento *wavetrack* contenga un archivo de sonido completo, los proyectos de Audacity subdividen la pista en una serie de elementos *waveclip*. Se trata de fragmentos de audio en general discontinuos. Esto permite ahorrar espacio de almacenamiento cuando entre dos fragmentos hay un silencio. Representar ese silencio implicaría, en formatos no comprimidos, la misma cantidad de bytes que si contuviera un sonido no nulo.

A su vez, cada *waveclip* se divide en pequeños segmentos de audio de un tamaño máximo de  $2^{20}$  bytes, cada uno de los cuales tiene asociado un archivo de sonido propio residente en el directorio de datos del proyecto, o bien es parte de un archivo externo que se ha importado sin permitir su copia al directorio de datos.



**Figura 54.** Estructura de un *wavetrack* en un proyecto de Audacity.

<sup>54</sup> Notar que un canal corresponde a una entrada o salida física (por ejemplo izquierdo y derecho), mientras que una pista contiene una señal susceptible de ser asignada parcialmente a cada uno de los canales mediante el control de *paneo*.

<sup>55</sup> Cuando un canal se configura como *solo* automáticamente se silencian todos los que no estén a su vez configurados como *solo*.

<sup>56</sup> Cuando a un canal se le aplica *sordina*, el mismo queda silenciado.

A continuación se proporciona un ejemplo del contenido de un archivo .aup en el que se intercalan notas y aclaraciones. Las entradas en el documento XML original se transcriben en negrita y las explicaciones en cursiva precedidas por un guión. Para mayor claridad se han separado en líneas e indentado casi todos los atributos de los diferentes elementos. Se han indentado, asimismo, los diversos elementos y subelementos. Para una especificación más formal, ver el apéndice 12.

```
<?xml version="1.0" standalone="no" ?>
```

— *Utiliza la versión 1.0 de XML y no es standalone (requiere definición DTD externa)*

```
<!DOCTYPE project PUBLIC "-//audacityproject-1.3.0//DTD//EN"
"http://audacity.sourceforge.net/xml/audacityproject-1.3.0.dtd" >
```

— *Declara la DTD y en qué dirección web está disponible*

```
<project
  xmlns="http://audacity.sourceforge.net/xml/"
  projname="prueba_data"
  version="1.3.0"
  audacityversion="2.0.3"
  sel0="0.5000000000"
  sel1="3.5000000000"
  vpos="0" h="0.0000000000"
  zoom="154.7142857143"
  rate="44100.0">
```

— *El elemento project contiene, entre sus atributos, el nombre del directorio del proyecto, la versión de Audacity, el estado del entorno de edición, indicando los instantes de tiempo inicial y final del intervalo que se encontraba seleccionado la última vez que se guardó el proyecto, el valor porcentual de zoom respecto a una duración de alrededor de 15 s (depende de la resolución del monitor), la posición (expresada en segundos) del borde del panel de control de las pistas con respecto al instante  $t = 0$  y la tasa de muestreo general.*

```
<tags>
  <tag name="ARTIST" value="Federico Miyara"/>
  <tag name="TITLE" value="Tono puro"/>
  <tag name="YEAR" value="2013-07-10"/>
  <tag name="Software" value="Cool Edit 2000"/>
  <tag name="Copyright" value="----"/>
  <tag name="COMMENTS" value="Prueba de metadatos en WAV"/>
</tags>
```

— *Se incluyen aquí metadatos típicos de la industria discográfica*

```
<wavetrack
  name="Mixto"
  channel="2"
  linked="0"
  mute="0"
  solo="0"
  height="150"
  minimized="0"
  isSelected="0"
  rate="44100"
  gain="0.501187"
  pan="-0.5">
```

— *El elemento wavetrack tiene atributos de nombre de pista, canal, sordina, solo, altura de la ventana de la pista en píxeles, estado de minimización o no, estado de selección o no, tasa de muestreo, ganancia y paneo. Contiene 0 o más elementos waveclip*

```
<waveclip  
  offset="1.00185941">
```

- El elemento *waveclip* tiene como único atributo el desplazamiento temporal (*offset*) en segundos de su comienzo con respecto al instante inicial del proyecto ( $t = 0$ ). En este caso comienza en el instante  $t = 1,00185941$  s.  
Contiene exactamente una secuencia y una envolvente.

```
<sequence  
  maxsamples="262144"  
  sampleformat="262159"  
  numsamples="44100">
```

- El elemento *sequence* (secuencia) tiene como atributos el máximo número de muestras y el mismo número más 11 palabras de 4 bytes del encabezamiento correspondientes a los bloques que contiene, y el número de muestras total que contiene la secuencia. Una secuencia contiene una serie de bloques de audio (*waveblocks*) que configuran un registro sonoro ininterrumpido.

```
<waveblock  
  start="0">
```

- Cada *waveblock* contiene una cantidad de muestras menor o igual a  $2^{20}$ . Su único atributo es el número de muestra inicial con respecto a la muestra inicial del *waveclip*, declarada en su atributo *offset*.  
El *waveblock* contiene la referencia a un archivo *.au* ubicado en el directorio de datos del proyecto

```
<simpleblockfile  
  filename="e0000b18.au"  
  len="44100"  
  min="-0.5"  
  max="0.5"  
  rms="0.237011"/>
```

- El elemento *simpleblockfile* se asocia a un archivo *.au* y sólo contiene atributos. Éstos son el nombre del archivo *.au*, su cantidad de muestras y sus valores mínimo, máximo y rms

```
</waveblock>
```

```
</sequence>
```

```
<envelope  
  numpoints="0"/>
```

- El elemento *envelope* (envolvente) contiene 0 o más puntos de control, formados por un par ordenado (tiempo, valor) que definen una envolvente de amplitud. En este caso no hay puntos de control, por lo que la envolvente es idénticamente 1 en toda la duración de la secuencia.

```
</waveclip>
```

```
<waveclip  
  offset="4.00100000">
```

- Este *waveclip* comienza en el instante  $t = 4,00100000$  s.

```
<sequence  
  maxsamples="262144"  
  sampleformat="262159"  
  numsamples="44761">  
  <waveblock
```

```

        start="0">
        <simpleblockfile
            filename="e0000c89.au"
            len="44761"
            min="-0.501237"
            max="0.501229"
            rms="0.293003"/>
        </waveblock>
    </sequence>
    <envelope
        numpoints="0"/>
    </waveclip>
</wavetrack>

```

```

<wavetrack
    name="Tono 100 Hz"
    channel="2"
    linked="0"
    mute="0"
    solo="0"
    height="150"
    minimized="0"
    isSelected="1"
    rate="48000"
    gain="0.501187"
    pan="0.5">

```

```

<waveclip
    offset="0.00000000">

```

```

    <sequence
        maxsamples="262144"
        sampleformat="262159"
        numsamples="336000">

```

```

        <waveblock
            start="0">

```

— Este primer bloque de audio comienza en el instante inicial del waveclip

```

        <pcmaliasblockfile
            summaryfile="e000039a.auf"
            aliasfile="E:\Sonidos\Audacity\Tono 100 Hz.wav"
            aliasstart="0"
            aliaslen="262144"
            aliaschannel="0"
            min="-0.501251"
            max="0.501221"
            rms="0.354355"/>

```

— En este caso el bloque corresponde a un sonido que no se copió al directorio de datos del proyecto sino que el origen de los datos es un archivo ubicado en E:\Sonidos\Audacity\Tono 100 Hz.wav. No obstante, el proyecto involucra un archivo .auf que contiene el sumario del sonido (versión submuestreada a 1/85 de la tasa original).

```

        </waveblock>

```

```

    <waveblock start="262144">

```

— Este segundo bloque de audio comienza en la muestra 262144 a partir del instante inicial del waveclip.

```

        <pcmaliasblockfile

```

```

summaryfile="e0000cf3.auf"
aliasfile="E:\Sonidos\Audacity\Tono 100 Hz.wav"
aliasstart="262144"
aliaslen="73856"
aliaschannel="0"
min="-0.501251"
max="0.501251"
rms="0.266338"/>

</waveblock>

</sequence>

<envelope numpoints="3">
— En este caso la envolvente tiene 3 puntos de control (elementos controlpoint).
  Entre el segundo y tercer puntos de control hay un descenso brusco de la
  ganancia de 1 a 0,5277...

  <controlpoint t="0.000000000000" val="1.000000000000"/>
  <controlpoint t="3.044321329640" val="1.000000000000"/>
  <controlpoint t="3.044326329640" val="0.527777791023"/>
</envelope>

</waveclip>

</wavetrack>

<labeltrack
name="Etiquetas"
numlabels="2"
height="73"
minimized="0">
— El elemento labeltrack tiene atributos de nombre de pista, cantidad de etiquetas, altura
de la ventana de la pista en píxeles y estado de minimización o no. Contiene 0 o más
elementos label (etiqueta)

<label
t="1.02585411"
t1="2.01662050"
title="Ruido blanco"/>
— El elemento label tiene los siguientes atributos: tiempo inicial, tiempo final y título
(texto de la etiqueta)

<label
t="3.04432133"
t1="3.04432133"
title="Cambio de ganancia"/>
— En este caso los tiempos inicial y final coinciden, indicando que es una etiqueta
asociada a un instante

</labeltrack>

</project>

```

Es interesante observar que el archivo .aup puede editarse e incluso crearse externamente al entorno de Audacity, con lo cual se podrían fijar con gran precisión cuestiones como puntos de control en envolventes que podrían asociarse a cambios de escala en el instrumento que generó la señal.



### 3.16.2 Formato .AU

Los archivos .au residentes en el directorio de datos del proyecto tienen un encabezamiento de longitud fija, un campo de anotaciones de longitud variable y un segmento de datos de audio. En la tabla 3.9 se indica la estructura de un archivo .au. Estos archivos están basados en los archivos .au de la firma Sun Microsystems. Las principales diferencias son que empieza con los caracteres “dns.” en lugar de “.snd” (nótese que se ha elegido un identificador que es la retrógrada exacta del de los archivos de Sun), que la longitud de datos de audio es irrelevante y que siempre hay un solo canal.

**Tabla 3.9.** Estructura de un archivo .au de datos de audio de Audacity

Longitud [byte]	Contenido	Comentarios
4	“dns.”	Identificador asociado a los archivos de sonido de Audacity
4	Offset $A$ en byte del audio	Hexadecimal little endian
4	0xFFFFFFFF	Código que indica que la longitud de datos de audio es irrelevante
4	0x06000000	Código que indica que los datos están en formato IEEE-754 de 32 bit flotante
4	Tasa de muestreo en Hz	Hexadecimal little endian
4	Cantidad de canales	Es siempre 1 para bloques de audio de Audacity
8	“Audacity”	Identificador del programa
12	“BlockFile112”	Identificador del tipo de información
$A - 44$	Campo de anotaciones	Longitud variable múltiplo de 8 byte
$L - A$	Segmento de datos de audio	Muestras de audio en formato IEEE-754 de 32 bit punto flotante

Aunque el formato no se encuentra documentado oficialmente, se consultaron diversos foros y al iniciador del proyecto (Dominic Mazzoni) en relación con el contenido del campo de anotaciones. El mismo está formado por un *sumario*, es decir una versión resumida de la señal consistente en los valores mínimo, máximo, y eficaz (RMS) cada 65536 muestras y luego los valores mínimo, máximo y eficaz cada 256 muestras.

El sumario es útil para graficar el oscilograma de la señal cuando se visualizan intervalos de tiempo muy grandes, ya que graficar millones de muestras en una pantalla que a lo sumo puede tener unos pocos miles de píxeles de resolución horizontal<sup>57</sup> implicaría que muchos de esos valores colapsen en una sola línea vertical. Resulta más atinado representar los valores correspondientes al sumario, eligiéndose los valores cada 256 muestras cuando el zoom es intermedio y los valores cada 65536 muestras cuando el zoom es muy bajo (es decir, se representa un tiempo muy grande, por ejemplo media hora).

Originalmente, el campo de anotaciones estaba previsto para futura inclusión de datos estructurados, como marcadores o puntos de edición. En Audacity dicha información se incluye en el archivo de proyecto .aup.

El tamaño  $L_{\text{archivo}}$  de un archivo .au puede obtenerse sumando la longitud del encabezamiento (44 byte) más la longitud del segmento de audio  $N_{\text{audio}}$  más la longitud del sumario, es decir,

<sup>57</sup> Por ejemplo, un monitor de alta resolución (full HD) tiene 1920 píxeles en dirección horizontal.

$$L_{\text{archivo}} = 44 + 4N_{\text{audio}} + 12 \left\lceil \frac{N_{\text{audio}}}{65536} \right\rceil + 12 \left\lceil \frac{N_{\text{audio}}}{256} \right\rceil$$

donde  $\lceil \rceil$  indica la parte entera.

### 3.16.3 Formato .AUF

Cuando alguna de las fuentes originales de audio de un proyecto de Audacity es un archivo de sonido que ha sido importado sin guardar una copia en el directorio de datos del proyecto (denominado *archivo alias*), Audacity genera uno o más archivos subsidiarios denominados *archivos de sumario*. Cada uno de ellos contiene el sumario (versión submuestreada en un factor 1/85) correspondiente a una porción de audio de hasta  $2^{20}$  bytes. El formato de los mismos se da en la tabla 3.10. Se trata de una versión abreviada del formato .au en la que se eliminó el encabezamiento y el audio propiamente dicho.

**Tabla 3.10.** Estructura de un archivo .auf de datos de audio de Audacity

Longitud [byte]	Contenido	Comentarios
8	“Audacity”	Identificador del programa
12	“BlockFile112”	Identificador del tipo de información
$A - 44$	Sumario	Longitud variable múltiplo de 8 byte

### 3.16.4 Acceso externo a los datos de audio

Es posible acceder a los datos de audio de un proyecto de Audacity mediante software externo, por ejemplo los paquetes matemáticos matriciales Scilab, Octave o Matlab. Para ello existen dos posibilidades. La primera es exportar desde Audacity el proyecto como archivo .wav y posteriormente utilizar la función wavread de dichos paquetes para adquirir la señal en forma de un vector columna que contiene las muestras normalizadas entre  $-1$  y  $1$  (o matriz de dos columnas en el caso de archivos estereofónicos). La segunda posibilidad consiste en leer el archivo de proyecto .aup y extraer del mismo la información sobre la ubicación de cada bloque de audio y otros atributos de edición como el paneo y la ganancia, para luego reconstruir la señal dentro del software matemático.

## 3.17 Medios de registro y almacenamiento

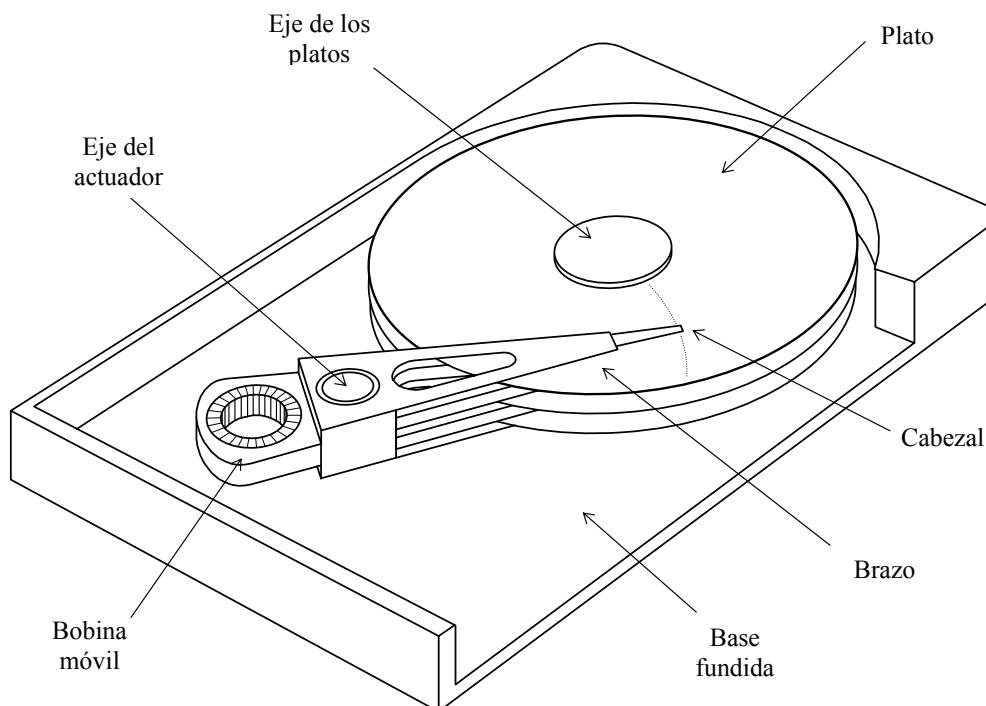
Coexisten en la actualidad diversos medios de registro y almacenamiento de audio digital. Aclaremos que por *registro* entendemos el proceso de transferencia inmediato, en tiempo real o con baja *latencia*,<sup>58</sup> de la señal digital generada por el conversor analógico/digital al medio, mientras que por *almacenamiento* entendemos la conservación a largo plazo de dicha información.

Los medios más utilizados son los discos duros (o discos rígidos), las memorias flash, los discos ópticos tales como el CD, el DVD y el Minidisc, y las cintas de audio digital (DAT).

<sup>58</sup> La *latencia* es el retardo que existe entre el instante en que el sonido llega al micrófono y el instante en el que el mismo es transferido al medio o, viceversa, el retardo entre que el sonido es leído del medio en el que se encuentra almacenado y es emitido al aire por un dispositivo reproductor.

### 3.17.1 Disco duro

Los *discos duros*, también denominados *discos rígidos*, constituyen actualmente (2013) el medio más difundido de almacenamiento masivo de información digital en computadoras. Se utilizan también en algunos grabadores de audio, sobre todo equipos multipista, y en algunas cámaras de video. También forman parte de unidades portátiles con conexión USB. Consisten en uno o más platos giratorios de aluminio o vidrio recubiertos con una película finamente pulida de material magnético, sobre los cuales “flotan” por efecto aerodinámico unos diminutos cabezales de lectura y escritura (ver figura 55). La altura de vuelo o flotación de los cabezales es inferior a  $1\ \mu\text{m}$ , para lo cual es preciso no sólo que la superficie sea extremadamente plana y lisa, sino además que esté perfectamente limpia. Para evitar la deposición del polvillo que hay en el aire se requiere que el conjunto esté contenido dentro de un receptáculo con aire altamente filtrado.



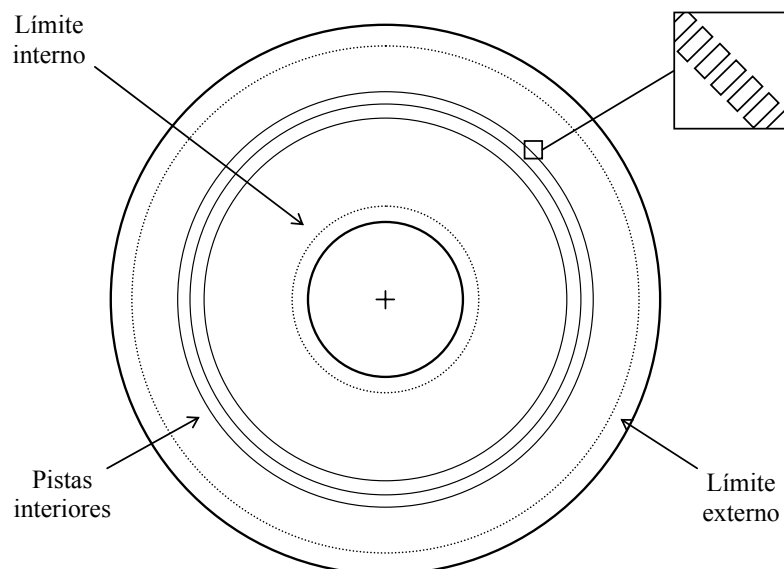
**Figura 55.** Esquema de la disposición interna de una unidad de disco duro de tres platos y seis cabezales (se ha removido una parte del lateral). En línea de puntos se muestra el recorrido pseudo radial del cabezal sobre la superficie externa del disco superior.

Los principios físicos de registro y lectura de información en el disco duro son similares a los de la cinta magnética. Sólo hay diferencias en los aspectos mecánicos. En la grabación analógica, una cinta magnética se mueve en forma moderadamente rápida frente a un cabezal fijo en contacto con ella. En el disco duro, el material magnético se mueve

mucho más rápidamente,<sup>59</sup> sustituyendo el movimiento lineal de la cinta por uno rotativo. El cabezal en este caso se mueve, pero lo hace radialmente y sólo para ubicarse a una distancia dada del eje, es decir, para realizar una búsqueda aleatoria. Como ya se indicó, no hay contacto entre el cabezal y el plato. Esto es para evitar la erosión que se produciría debido a la alta velocidad relativa entre el cabezal y el disco y al hecho de que el plato gira permanentemente (aunque no se esté leyendo o escribiendo información).<sup>60</sup>

### 3.17.1.1 Estructura física de una unidad de disco duro

En la figura 55 se muestran los componentes mecánicos principales de la unidad de disco duro. En primer lugar tenemos los discos propiamente dichos, o *platos*. En general, para aumentar la capacidad total de la unidad hay dos o más platos. Habitualmente, los platos tienen ambos lados aprovechables (es decir, ambos poseen recubrimiento magnético). La información se distribuye en cada plato a lo largo de una serie de circunferencias concéntricas denominadas *pistas* (ver figura 56). Cada bit individual ocupa una microscópica porción de la pista. Debido a la gran capacidad de los discos duros, el número de pistas es muy grande (algunas decenas de miles), y dentro de cada pista, a su vez, hay una gran cantidad de bits.



**Figura 56.** Estructura de pistas en un plato de una unidad de disco duro. Se ha ampliado un fragmento de una pista. Ni las pistas ni su espaciamiento están en escala.

Cada plato es accedido, para la escritura y la lectura, por uno o más cabezales montados en el extremo de sendos brazos. Estos brazos son todos solidarios y giran al-

<sup>59</sup> En la periferia de un disco de 3,5 pulgadas que gira a 7200 rpm, la velocidad lineal de la superficie respecto al cabezal alcanza los 5,33 m/s.

<sup>60</sup> Compárese con el DAT, en el cual la velocidad entre la cinta y el cabezal es de sólo 3 m/s y el rozamiento se produce únicamente mientras la cinta está en movimiento.

rededor de un eje común impulsados por un actuador. En los primeros discos el actuador era un motor eléctrico paso a paso. Este tipo de motor recibe pulsos eléctricos, y por cada pulso avanza un ángulo fijo (un submúltiplo de un giro completo). Mediante un mecanismo reductor se logra reducir el ángulo de manera que cada pulso permita saltar de una pista a la siguiente. Este tipo de acceso a las pistas se denomina *posicionamiento absoluto*. Tiene la ventaja de ser simple, dado que para ir de una pista a otra sólo hace falta entregar un número de pulsos igual a la cantidad de pistas a saltar más 1. Sin embargo, a medida que la tecnología avanza y se incrementa la densidad de pistas (es decir, el número por unidad de distancia), aumentan los problemas de precisión en el posicionamiento. Particularmente, al variar la temperatura ambiente tanto el plato como el brazo y otras partes del mecanismo experimentan dilataciones o contracciones. Al ser de materiales diferentes, las dilataciones o contracciones no coinciden, produciendo errores de alineación entre el cabezal y la pista.

En las unidades actuales el actuador está formado por una bobina (solidaria a la estructura sobre la que se montan los brazos) que se mueve dentro de un campo magnético creado por un potente imán permanente (no indicado en la figura). Al hacer circular corriente por la bobina, ésta genera su propio campo magnético (es decir, se convierte en un electroimán) que es atraído o repelido (según el sentido de la corriente) por el campo magnético del imán permanente, de la misma manera en que lo hace un galvanómetro de bobina móvil. Esta técnica no permite un posicionamiento directo muy preciso, por lo que se debe complementar con un sistema de servocontrol. Un *servocontrol*, o simplemente *servo*, es un mecanismo realimentado en el cual se mide, en cada instante, la posición real del elemento móvil a posicionar y se la compara con la posición deseada. Si ésta aún no se alcanzó, se continúa aplicando fuerza motriz. Si, en cambio, es rebasada, se aplica una fuerza motriz inversa. En el límite se alcanza un equilibrio en el que la posición final coincide precisamente con la deseada. Este tipo de acceso se denomina *posicionamiento relativo*.

### 3.17.1.2 Aspectos geométricos de un disco duro

Una característica importante de las unidades de disco es su *geometría*, es decir, la forma en que se distribuye la información en la superficie de los platos. Esta información necesita ser direccionada, es decir localizada espacialmente. Ello se logra por medio de la especificación del *cilindro*, el *cabezal* y el *sector* correspondientes a la información. Se denomina *cilindro* a todo conjunto de pistas de igual radio nominal sobre los diferentes platos de una unidad de disco. Este concepto se ilustra en la figura 57. La razón para hablar de *cilindro* y no directamente de *pista* es que todos los cabezales se mueven en forma solidaria, de manera que en cualquier posición del actuador es posible acceder a cualquiera de las pistas del cilindro correspondiente.

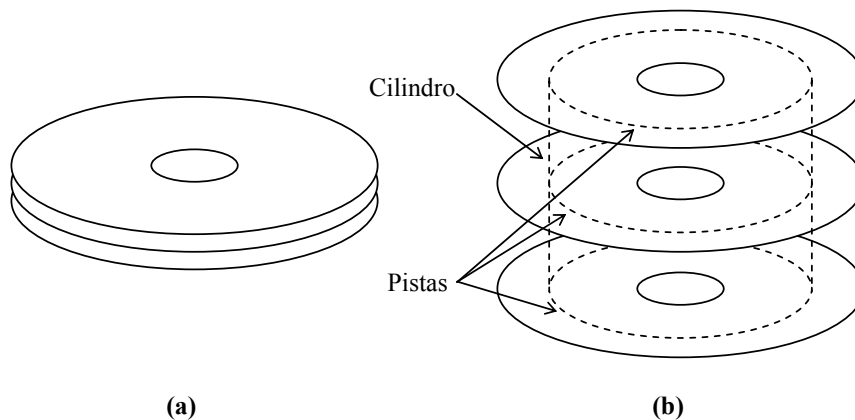
La especificación del *cabezal* se refiere a cuál de todos los cabezales que se encuentran sobre un determinado cilindro está siendo utilizado para leer o escribir. En general cada plato tiene las dos superficies (superior e inferior) recubiertas magnéticamente, y por lo tanto es accedido por dos cabezales. La excepción puede ser el último plato en el caso de unidades con una cantidad impar de superficies.

Por último, un *sector* es una porción angular del plato (figura 58) cuyas pistas contienen una cantidad fija de bytes denominada *bloque*<sup>61</sup> (en general, 512 ó 4096). Esta subdivisión de la información en bloques se debe a que las dificultades de sincronismo y de procesamiento del código corrector de errores son las mismas para un byte individual

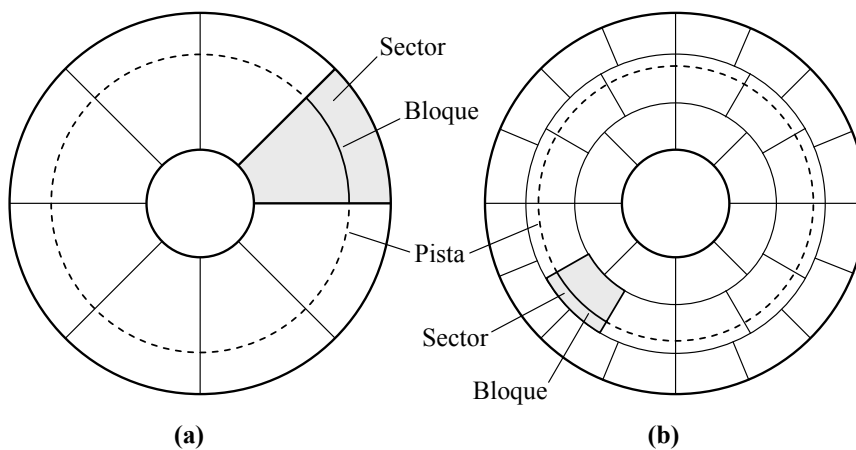
<sup>61</sup> Muchas veces se utiliza la palabra “sector” para referirse indistintamente a un sector o a un bloque.

que para un grupo completo de bytes, por lo cual resulta más eficiente realizar la transferencia en grupo. De hecho el bloque se elige como la menor cantidad de bytes para la que estos procesos resultan eficientes.

Además de la información del usuario, cada bloque contiene un encabezamiento que permite direccionarlo en forma unívoca dentro del disco y un código de detección y corrección de errores (EDC y ECC), necesarios ya que existe una cierta probabilidad de errores por pequeñas fallas en el material o en el proceso de escritura. En general esta información extra implica incrementar en alrededor de un 10 % la capacidad física del disco con respecto a la realmente aprovechable.



**Figura 57.** (a) Los platos en su posición normal. (b) Los platos separados para permitir visualizar las pistas correspondientes a un mismo cilindro.



**Figura 58.** Subdivisión de la superficie de un plato en sectores. (a) Igual cantidad de sectores en todas las pistas. (b) Más sectores en las pistas más externas.

La información a almacenar se guarda transitoriamente dentro de la sección electrónica de la misma unidad, en una memoria temporaria denominada *buffer*. Una vez que el *buffer* se ha llenado se vuelca su contenido en el disco. Análogamente, durante la lectura se transfiere el contenido de un bloque al *buffer* y, una vez completada la operación, se vuelca la información a la unidad central de proceso (*CPU*).

En los primeros discos, todas las pistas contenían la misma cantidad de sectores (figura 58 a), por lo cual las pistas externas tenían menor *densidad de bits* (cantidad de bits por unidad de longitud) que las internas. Si bien esto simplificaba enormemente el direccionamiento de los cabezales, también implicaba un importante desperdicio de espacio. Esta situación fue cambiando a medida que el requisito de capacidad de las unidades de disco se hacía mayor. Las unidades actuales (año 2013) tienen una cantidad variable de sectores, como se indica en la figura 58 b, lo cual se denomina *grabación zonificada*, *ZBR* (en inglés, *zoned bit recording*).

Los discos duros son relativamente delicados ya que el tamaño del cabezal es tan pequeño que podría dañarse fácilmente ante un golpe. Por este motivo, la manipulación de los discos duros, especialmente las unidades portátiles con conexión USB o las incluidas en *notebooks* o *laptops*, debe hacerse en forma cuidadosa, evitando movimientos bruscos. Como contrapartida, permiten una tasa de transferencia de datos bastante alta y una elevada confiabilidad en el almacenamiento a mediano plazo, pudiendo conservar la información inalterable durante 5 años o más.

Todos los sistemas de archivos prevén ir ocupando los espacios del disco que han sido borrados, y como los espacios así generados no necesariamente alcanzan para el próximo archivo a guardar, es muy frecuente que el archivo se fragmente en dos o más porciones. Con el tiempo, puede llegar a suceder que los archivos tengan un grado de fragmentación tan grande que, 1) se vea comprometida la velocidad de acceso, ya que el cabezal deberá desplazarse varias veces para acceder a las diferentes posiciones que ocupan los fragmentos y 2) se someta el sistema mecánico del cabezal a una excesiva sobrecarga, disminuyendo la vida útil. Por ese motivo se recomienda realizar periódicamente una operación de desfragmentación del disco duro (el sistema operativo tiene una utilidad para ello).

### 3.17.2 Memoria flash

La memoria flash es un medio de registro de audio digital muy conveniente por su reducido tamaño, bajo consumo energético, robustez frente a golpes y vibraciones, elevada capacidad de almacenamiento, acceso aleatorio y rapidez de transferencia. Sin embargo, no es excesivamente confiable para el almacenamiento a largo plazo y además admite una cantidad limitada de operaciones de escritura en cada posición de memoria, si bien en las de mejor calidad hoy en día se alcanza un orden especificado de 1 000 000 operaciones. Otra desventaja es que el precio por cantidad de información almacenada es mayor que en el disco duro, pese a que este último utiliza una tecnología mecánica muy sofisticada.

Desde hace algunos años se vienen utilizando memorias flash en grabadores portátiles en reemplazo del disco duro. Los formatos más difundidos son la SD (Secure Digital), la SDHC (Secure Digital High Capacity) y la CF (Compact Flash). En la actualidad alcanzan capacidades de hasta 256 GB, siendo los modelos más comunes de 4 GB a 16 GB, que permiten grabar entre 6 h y 25 h de audio con calidad de CD (16 bit estéreo, muestreado a 44 100 Hz).

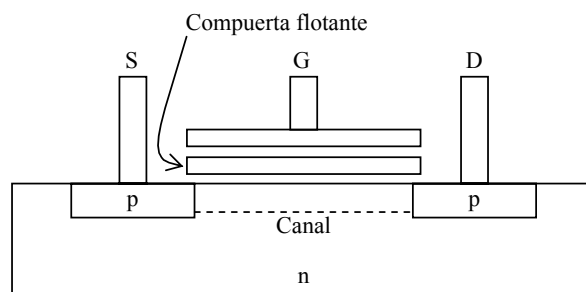
La velocidad de transferencia depende de la especificación de *clase* de la memoria, cifra expresada en MB/s que se indica en la etiqueta de la memoria. Así, una memo-

ria clase 6 permitirá una transferencia de 6 MB/s. Las memorias más comunes son de clase 4, valor más que suficiente para transferencia de audio. En efecto, una señal de audio digital de calidad CD requiere una tasa de transferencia de 176,4 kB/s.

La tasa de transferencia limitada podría constituir un problema en grabadores multipista, ya que permiten grabar decenas de pistas simultáneas por lo cual exigen una tasa de información muy elevada. Una aplicación metrológica donde esto puede ser útil es en el uso de técnicas de *beamforming*.<sup>62</sup>

La memoria flash se basa en el almacenamiento de cargas eléctricas en un conductor “flotante”, es decir separado de otras partes conductoras por un dieléctrico de resistencia extremadamente alta. Si bien existe una filtración de carga a través del dieléctrico, ésta se produce muy lentamente, por lo que potencialmente la información puede durar años. Esta duración, no obstante, puede verse afectada por condiciones de alta temperatura, presencia de campos electromagnéticos o radiación ionizante así como cambios imprevistos en la alimentación (caso de un corte de energía).

El elemento básico de las memorias flash es un tipo de transistor de efecto de campo metal-óxido-semiconductor (MOSFET) de compuerta flotante (FGMOS), cuya estructura se muestra esquemáticamente en la figura 59. Dicha compuerta flotante se encuentra separada del canal del MOSFET y de la compuerta de control por sendas películas aislantes de óxido de silicio. En condiciones normales, al aplicar una tensión positiva a la compuerta de control, el canal se vuelve conductor. Pero si se aplica una tensión relativamente alta (5 V), la densidad de corriente por el canal aumenta, permitiendo que algunos electrones de alta energía salten del canal hacia la compuerta flotante (fenómeno conocido como *inyección de electrones calientes*). Ésta queda, así, polarizada de tal forma que rechaza los portadores del canal, volviéndolo permanentemente aislante. El carácter conductor o aislante del canal se asocia naturalmente con un 1 o un 0 lógico. Para volver al estado primitivo se aplica una tensión de  $-5\text{ V}$  a la compuerta de control, con lo cual los electrones atrapados en la compuerta flotante están sometidos a un campo eléctrico tan intenso que pueden volver al canal, tornándolo nuevamente conductor.



**Figura 59.** Estructura de un transistor FGMOS utilizado en las memorias flash. S es la fuente (se conecta a masa), D el drenador (la salida) y G la compuerta de control.

<sup>62</sup> Las técnicas de *beamforming* (conformación de haz) consisten en utilizar un arreglo de varios micrófonos y aplicar distintas ponderaciones a cada uno de manera de crear un patrón direccional altamente selectivo. Como ese patrón puede cambiarse durante el postprocesamiento, es una técnica que permite presentar imágenes de la distribución de fuentes sonoras.



Estos transistores FGMOS se integran en una matriz que permite el acceso individual a cada uno de ellos mediante un elaborado sistema de direccionamiento.

Hay dos tipos de memorias flash, conocidas como NOR y NAND. Si bien los detalles de su implementación escapan a los propósitos de este libro, digamos que las NOR permiten un acceso aleatorio a cada byte durante la lectura y la programación, pero sólo permiten el borrado por bloques. Las NAND, por otra parte, logran mayor densidad superficial de información, pero al precio de que el acceso para la lectura y programación sea por *página*, donde cada página está formada por entre 512 y 4096 bytes, mientras que el borrado se realiza por bloque, donde cada bloque contiene entre 32 y 128 páginas. Esto no constituye un problema insoluble, dado que si sólo se requiere borrar unos pocos bytes, bastará leer la totalidad del bloque que lo contiene y guardarlo transitoriamente en una memoria RAM, para luego volver a programar el bloque con los bytes modificados.

Este comportamiento es muy similar al de la lectura de sectores en los discos rígidos, donde no es factible leer, grabar o borrar un byte individual.

En la actualidad, la mayoría de las memorias flash disponibles en el mercado son de tipo NAND, debido a que, como se mencionó, permiten mayor densidad de información.

En el caso de las memorias flash, no sólo no es necesaria la desfragmentación sino que además es muy contraproducente, ya que obliga a una enorme cantidad de operaciones de borrado y escritura, que podrían redundar en una disminución de la vida útil del dispositivo.

### 3.17.3 Discos ópticos

La tecnología del disco óptico se inicia en la década de 1980 con la introducción del disco compacto (*compact disc*, CD). Se trata de discos de policarbonato, un polímero de excelente estabilidad dimensional que permite además un moldeado de precisión. La filosofía general de los discos ópticos es proporcionar dos reflectividades diferentes, asignando a una de ellas un valor lógico 0 y a la otra, 1. El disco se explora mediante un haz láser muy delgado que al incidir sobre la superficie activa es reflejado con mayor o menor intensidad generando en un fototransistor dos valores diferenciados asociables a los dos estados lógicos.

La forma de lograr la diferencia en la reflectividad difiere según el tipo de disco. En los discos de lectura solamente producidos industrialmente (CD de audio, DVD de video, CD-ROM de software u otros contenidos digitales) existe una sucesión de fosas o cavidades cuya profundidad es  $1/4$  de la longitud de onda de la luz incidente. Cuando la luz incide en el fondo de una fosa o fuera de ella, la reflexión es completa gracias al acabado metalizado. En cambio, cuando incide en el borde de una fosa, la parte que se refleja en el fondo retorna con un desfase de  $1/2$  longitud de onda con respecto a la que se refleja en la superficie. Se produce una interferencia sustractiva entre ambas ondas, disminuyendo la luminosidad reflejada. Resulta así que los bordes corresponden a un 1 y el resto a un 0. No puede haber, por lo tanto dos 1's seguidos, lo que obliga a una recodificación de los datos digitales mediante la denominada *modulación ocho a catorce* (*eight to fourteen modulation*, EFM).<sup>63</sup>

En los discos grabables una sola vez, CD-R, la técnica de escritura consiste en que un láser de mayor potencia modifica las características de un pigmento de la familia de las *cianinas*, las *ftalocianinas* o los *azos*, descomponiéndolo a una temperatura de 250 °C. Donde el pigmento se ha quemado disminuye la reflectividad, provocando que

<sup>63</sup> Consiste en reemplazar cada byte por 14 bits que cumplen con la condición de que no puede haber dos 1's seguidos ni más de diez 0's seguidos.

al ser posteriormente iluminado por un láser de menor potencia se lea un 1. Si bien en este caso podría haber dos o más 1's seguidos, para compatibilidad con los CD convencionales se sigue utilizando la modulación ocho a catorce. Las cianinas utilizadas en los primeros CD-R eran químicamente inestables y particularmente sensibles a la exposición a luz ultravioleta. Los pigmentos utilizados actualmente son más estables, aunque la exposición a la luz ultravioleta sigue causando degradaciones. Un disco con fines de preservación a largo plazo debería almacenarse protegido de la luz solar.

Finalmente, los discos regrabables CD-RW utilizan una técnica de cambio de fase. Se trata de la irradiación de un material reflectivo policristalino (una aleación de plata, indio, antimonio y telurio) con un láser de alta potencia hasta fundirlo a una temperatura de entre 500 °C y 700 °C. Al solidificarse no vuelve su estado policristalino cristalino sino que pasa a un estado amorfo que resulta opaco (poco reflectante). Para restaurar el material a su estado policristalino original se lo irradia con menos potencia, de modo de alcanzar los 200 °C, temperatura a la cual recupera su estado policristalino original de alta reflectancia.

### 3.17.3.1 Disco compacto (CD o CD-DA)

Estudiaremos ahora con algún detenimiento el formato digital del disco compacto de audio digital. El soporte consiste en un disco de policarbonato transparente de 120 mm de diámetro (en el mini-CD, de 80 mm) y 1,2 mm de espesor (figura 60). El disco se coloca en el reproductor introduciendo el orificio central en el eje de tracción, operación que en la mayoría de los reproductores se lleva a cabo automáticamente. La lectura se realiza por medio de un diodo láser, el cual emite un haz de luz que es reflejado en mayor o menor grado, indicando la presencia de un 1 ó un 0. El diodo no entra en contacto con la superficie del disco por lo cual no existe desgaste ni en el elemento captor ni en la superficie del disco, permitiendo una alta durabilidad.

Contrariamente a lo que sucedía en los discos analógicos de vinilo, la lectura se realiza desde adentro hacia afuera, y la velocidad de giro no es constante, sino que se va reduciendo. Otra diferencia es que la velocidad de giro difiere de un disco compacto a otro. Esto es debido a que la velocidad se utiliza como variable de ajuste para que el área disponible se llene casi totalmente, sin importar la duración del material grabado (que puede llegar a un máximo de 74 min 33 s). Para un disco compacto dado, lo que se mantiene constante es la *velocidad lineal* (*constant linear velocity*, CLV), y no la velocidad de rotación. La velocidad lineal se mantiene constante gracias a un control de velocidad o servomecanismo que utiliza señales de sincronismo registradas en el propio disco.

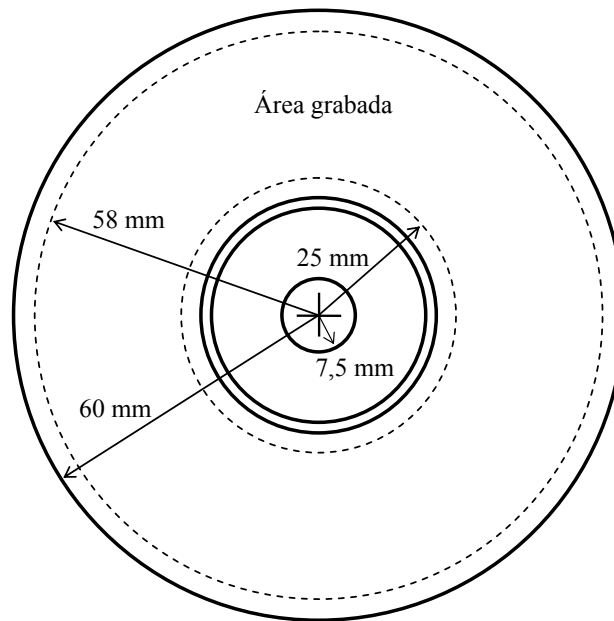
La velocidad lineal en los discos compactos más cortos es mayor (cerca de 1,4 m/s) y en los más largos es menor (alrededor de 1,2 m/s). La relación entre la velocidad lineal  $v$  en m/s y la de rotación  $\omega$  en rpm puede obtenerse del siguiente modo. Para un radio dado  $r$ , cada vuelta tendrá una longitud

$$L = 2 \pi r,$$

por lo cual la cantidad de vueltas por segundo será  $v/L$ , y la cantidad de vueltas (o revoluciones) por minuto, 60 veces mayor. Entonces

$$\omega = 60 \frac{v}{L} = \frac{60 v}{2\pi r}.$$

Si tomamos un disco promedio, en el cual la velocidad lineal sea de 1,3 m/s, la velocidad de rotación al comienzo del disco (es decir, en su vuelta interna, de radio 25 mm) será



**Figura 60.** Dimensiones de un disco compacto. Las circunferencias punteadas delimitan el área grabada. Las circunferencias próximas al orificio proporcionan un cambio de nivel para separar el área grabada del lector.

$$\omega = \frac{60 \cdot 1,3}{2 \cdot 3,14 \cdot 0,025} = 497 \text{ rpm} ,$$

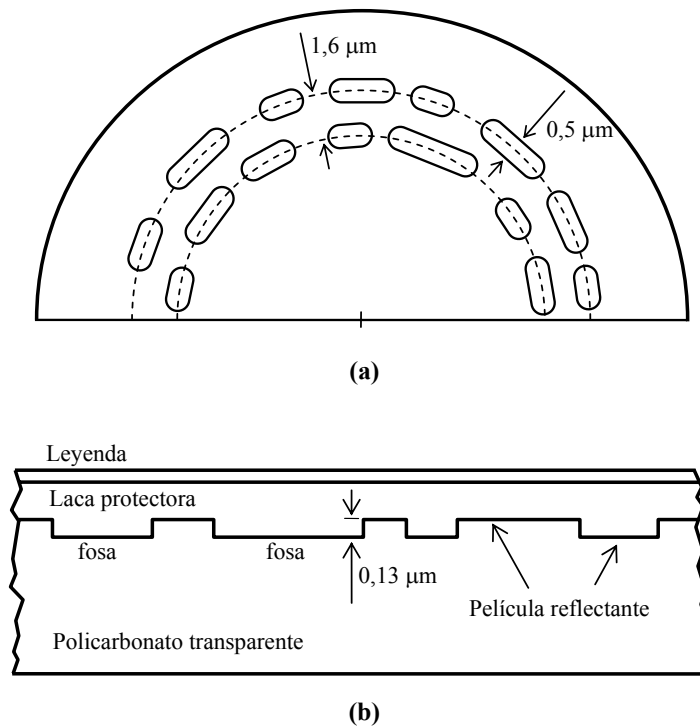
mientras que al terminar (en su vuelta externa, de radio 58 mm), será

$$\omega = \frac{60 \cdot 1,3}{2 \cdot 3,14 \cdot 0,058} = 214 \text{ rpm} .$$

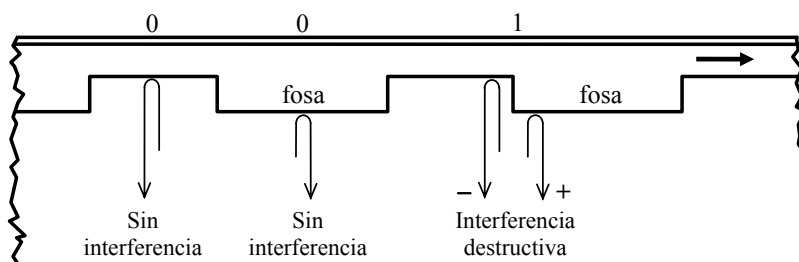
Se observa que las velocidades de giro son mucho más altas que para el tradicional disco de vinilo (33 rpm ó 45 rpm).

Veamos ahora cómo se codifica la información digital en un disco compacto. Se utiliza una estructura de fosas de forma ovooidal (figura 61a) estampadas sobre la superficie superior del disco. Dicha superficie es luego metalizada con plata o aluminio, con el propósito de obtener un acabado reflectante, y posteriormente recubierta por una laca protectora, sobre la cual se imprimen las leyendas que identifican al disco (figura 61b). La exploración del disco por medio del láser se efectúa desde la cara inferior, es decir, la cara sin leyendas. La profundidad de las fosas está ajustada a un valor entre  $0,11 \mu\text{m}$  y  $0,13 \mu\text{m}$ , elegido de modo que el haz de luz infrarroja (de  $780 \text{ nm}$ ) reflejada sobre una fosa se desfase con respecto a la onda reflejada fuera de ella. Esto hace que cuando se proyecta un haz justo sobre el borde de una fosa, dado que parte de la luz se refleja dentro y parte fuera de la fosa, existirá una interferencia destructiva (figura 62). En consecuencia, cada vez que el borde de una fosa pase frente al fotosensor, éste recibirá luz

atenuada, y en cambio cada vez que esté frente al interior o frente al exterior de una fosa recibirá luz intensa. Como ya se mencionó, se interpreta cada borde de una fosa como un 1 y el resto como 0.



**Figura 61.** Estructura de fosas de un disco compacto. (a) Disco visto desde abajo. (b) Corte transversal de una pequeña porción del disco a lo largo de una hilera de fosas. (Diagramas fuera de escala.)



**Figura 62.** Mecanismo de lectura de la información digital contenida en un disco compacto.

Uno de los problemas que debieron resolverse en el diseño del CD es el de la robustez frente a corrupción de datos. A diferencia de un disco duro, que se encuentra encerrado en una cavidad con aire libre de partículas de polvo, el CD es un medio susceptible de

ser manipulado por cualquier persona, incluso por quienes no son demasiado cuidadosos, por lo cual la propensión a sufrir alteraciones es alta.

Dado que el espacio que ocupa un bit individual es diminuto, una pequeña raspadura, una impresión digital<sup>64</sup> o una mancha de suciedad son capaces de alterar o hacer ilegibles unos cuantos bits espacialmente vecinos. Este tipo de error se denomina *error de ráfaga* o también *error de salva*. Uno de los grandes desafíos es lograr un código lo suficientemente robusto como para ser tolerante frente a errores de ráfaga.

La solución utiliza dos estrategias. La primera es utilizar un código con corrección de errores aislados, lo cual se consigue agregando redundancia. A modo de ejemplo, si tuviéramos palabras de 2 bit,

00   01   10   11,

podríamos agregar otros tres bits elegidos apropiadamente, por ejemplo,

00000   01101   10110   11011,

donde se han subrayado los tres bits agregados. Por cada palabra original se presentan seis posibilidades, una sin alteración y cinco con un solo bit cambiado. En total son 24 posibilidades. Se puede observar que, para cada una de estas palabras de 5 bit, cualquiera de estas seis alternativas no puede provenir de ninguna de las otras palabras, por lo cual es posible reconstruir la palabra original y, a partir de ella, deducir la palabra de 2 bit representada.

Este ejemplo sencillo puede generalizarse mediante un código propuesto por Reed y Solomon en 1960 (Pohlman, 2020). Este código toma una palabra original de  $k$  símbolos<sup>65</sup> y le adiciona  $2t$  símbolos elegidos apropiadamente para obtener una palabra de  $n = k + 2t$  símbolos que es resistente a errores en hasta  $t$  símbolos (Pohlman, 2002). Este código se designa como RS( $n, k$ ).

La segunda estrategia consiste en reordenar la información de forma tal que los símbolos originalmente adyacentes se encuentren alejados entre sí. Dado que los errores de ráfaga afectan a símbolos adyacentes, los símbolos que ahora son adyacentes pertenecen en realidad a diferentes palabras, por lo que en cada palabra se verán afectados muy pocos símbolos. Ésta es una situación ideal para ser manejada por el código de Reed-Solomon, ya que permite corregir unos pocos errores por palabra.

La combinación de ambas estrategias da como resultado un código tolerante a errores moderados de ráfaga.<sup>66</sup> Se lo denomina *código de intercalación cruzada de Reed-Solomon* (*Cross Interleaved Reed-Solomon Code*, CIRC) en honor de sus autores.

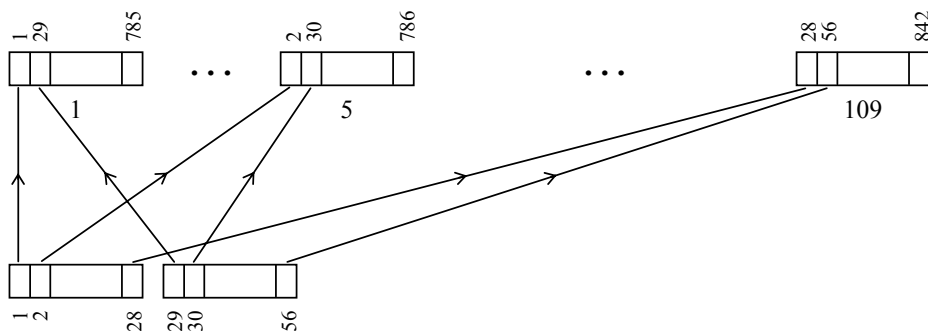
En el caso de los CD se parte de un conjunto de 6 muestras estéreo sucesivas (denominado *trama*), que contiene  $6 \times 2 \times 2$  bytes, es decir, 24 bytes. En una primera etapa se lleva a cabo una intercalación de las muestras, lo cual supone alterar el orden sucesivo original  $D_1 I_1 D_2 I_2 D_3 I_3 D_4 I_4 D_5 I_5 D_6 I_6$  por  $D_1 D_3 D_5 I_1 I_3 I_5 D_2 D_4 D_6 I_2 I_4 I_6$ . Este proceso se completa con un retardo de una muestra para las muestras pares. El resultado se pasa

<sup>64</sup> De todas maneras, el CD es bastante insensible a manchas más o menos transparentes. Debido al índice de refracción del policarbonato, el haz dentro del material es cónico. El vértice (de aproximadamente 1  $\mu\text{m}$ ) se encuentra en la superficie interna activa, pero en la superficie externa su diámetro es de unos 0,7 mm, por lo cual la contaminación de la superficie no es tan perjudicial.

<sup>65</sup> Cada símbolo puede tener  $s$  bits. Por ejemplo, puede ser un byte formado por 8 bit

<sup>66</sup> Desde luego, no existe ningún mecanismo corrector de errores tolerante a un error de ráfaga masivo. Este método se orienta a los errores que típicamente pueden producirse, por ejemplo pequeñas manchas de suciedad o raspaduras.

por un codificador RS(28, 24) que agrega 4 bytes para obtener una nueva trama de 28 bytes (a este codificador se lo designa habitualmente como C2). En una segunda etapa se aplica una nueva intercalación en la cual cada uno de los 28 bytes irá a parar a una trama diferente, estando dichas tramas separadas un espacio equivalente a 4 tramas. En otras palabras, cada trama original se distribuye en un espacio equivalente a  $27 \times 4 + 1$ , es decir 109 tramas (ver figura 63). La tercera etapa involucra un nuevo codificador de Reed-Solomon RS(32,28) (designado como C1) que agrega otros 4 bytes. La nueva codificación tiene 8 bytes más que el original por cada trama, por lo que permite corregir hasta 4 bytes erróneos. La mínima proximidad entre dos bytes originalmente adyacentes será ahora  $32 \times 4$  bytes, por lo que un error de ráfaga que alterara hasta  $4 \times 32 \times 4$  bytes (es decir, hasta 512 bytes o 4096 bit) estaría cambiando sólo 4 bytes de la trama con la doble intercalación y la doble codificación Reed-Solomon, es decir, una alteración que el código es capaz de detectar y corregir.



**Figura 63.** Intercalación cruzada en un CD luego de la aplicación del codificador de Reed-Solomon RS(28,24). Abajo, tramas originales. Arriba, tramas luego de la intercalación.

A cada trama de 32 bytes que resulta luego de la codificación CIRC (24 bytes originales más 8 bytes de redundancia para la corrección de errores) se le agregan dos elementos adicionales. El primero es un byte denominado *subcódigo*, cuyos bits se suelen designar como P, Q, R, S, T, U, V, W. Está destinado a alojar algunos metadatos y se establece que los bits homólogos de 98 tramas sucesivas forman una palabra de subcódigo de 98 bit, por ejemplo,  $P_1P_2 \dots P_{98}$ . Los subcódigos P y Q se destinan a información sobre la distribución de las pistas (comienzo y fin, número de pista), tabla de contenido, código de catalogación ISRC, etc. También contiene un código de redundancia cíclica CRC para detectar errores.

El segundo elemento es una *palabra de sincronismo* de 24 bit, elegida como un patrón imposible de formar con ninguna otra distribución de datos:

10000000001000000000010 .

La finalidad es recalibrar frecuentemente (cada  $6/44100 \text{ s} = 0,13605 \text{ ms}$ ) el reloj del sistema para evitar que un corrimiento lento se acumule gradualmente y llegue a provocar errores catastróficos.

Se tiene así un código que contiene 256 bit de audio codificado CIRC, 8 bit de subcódigo y 24 bit de sincronismo, totalizando 288 bit. Esta información no puede transferirse en forma directa al CD por dos razones. La primera es que no hay garantía de que no aparezcan dos o más 1's seguidos, y, según vimos, la tecnología utilizada no permite tener dos 1's seguidos ya que el 1 corresponde a una transición entre el fondo y la superficie de una fosa y no es posible tener dos transiciones seguidas. Por otra parte, también existe el riesgo contrario de que haya demasiados 0's seguidos, lo cual implicaría la presencia de gran contenido de baja frecuencia en las transiciones. Esto complica la operación del sistema de sincronismo. Es necesario, por lo tanto, transformar la secuencia original de 0's y 1's mediante lo que se denomina un *código de canal de grupo*, es decir, uno que codifique grupos de bits en otros grupos aptos para ser transmitidos por un determinado canal de transmisión (en este caso la superficie activa del disco) respetando las limitaciones de ese canal.

El código de canal utilizado en el CD es la *modulación de 8 a 14* (EFM), que sustituye cada byte por un código de 14 bit de manera de cumplir la regla 2 a 10, según la cual no puede haber menos de 2 ni más de 10 ceros seguidos. De las 16384 combinaciones posibles de 14 bit hay 267 que satisfacen esta regla, entre las cuales se eligieron 256 para configurar un "diccionario". La palabra de sincronismo no se modula ya que cumple con la regla 2 a 10.

La modulación 8 a 14 garantiza el cumplimiento de la regla 2 a 10 dentro de cada palabra válida de 14 bit, pero no asegura que se cumpla entre palabras vecinas. Por ejemplo una palabra podría terminar en 1 y la siguiente empezar en 1. Para evitar esto y además reducir prácticamente a 0 la componente de continua (requisito necesario para un sistema de sincronismo robusto frente a alteraciones del medio, por ejemplo la presencia de grasitud en la superficie), por cada palabra de 14 bit se agregan 3 bit elegidos dinámicamente para evitar estos problemas. También se agregan 3 bit a la palabra de sincronismo. Resulta así que el código de canal de cada trama tendrá una longitud

$$(256 + 8) \times \frac{14 + 3}{8} + 24 + 3 = 588 \text{ bit}$$

lo que configura un factor de redundancia de  $588 / 192 = 3,0635$ , es decir se triplica la cantidad necesaria de bits. Ello es necesario para lograr un sistema robusto frente a errores de ráfaga y compatible con las características y limitaciones del medio.

El código CIRC no sólo permite corregir errores de ráfaga no demasiado severos sino que además detecta muchos errores que no puede corregir debido a que es demasiada la información perdida. Esta detección implica también su localización y la posibilidad de disimularlos mediante técnicas de interpolación. En otras palabras, si a pesar de todos los recursos disponibles alguna muestra se detecta como errónea sin posibilidad de corrección, se la sustituye por el valor medio entre dos muestras adyacentes. Esto es bueno para la reproducción de audio, pero podría alterar la señal en caso de realizar copias del audio.

Es importante tener esto en cuenta, ya que uno de los aspectos más destacables del audio digital era la posibilidad de obtener copias idénticas al original. El CD de audio, debido a que prioriza la experiencia de escucha por sobre la integridad absoluta de los datos, puede no ofrecer una garantía completa de ello.

Cuando el CD se va a utilizar con fines de archivo o repositorio, conviene observar ciertas normas para su manipulación y almacenamiento. Éstas incluyen:

1) No tocar ninguna de las superficies en forma directa, ya que la inevitable grasa de los dedos podría incrementar la probabilidad de pérdida de información. También la acidez de la transpiración podría, a la larga, dañar la laca. Conviene en general tomar el CD por el borde.

2) Evitar el almacenamiento en cajas delgadas (*slim*), ya que el espacio es tan exiguo que la presión al apilar varios CD hace que las superficies del disco toquen las caras internas de la caja.

3) Evitar lugares húmedos que podrían favorecer el desarrollo de hongos y otros microorganismos. De ser posible, utilizar algún producto desecante como el cloruro de calcio o el sílica gel, en bolsas protegidas. Tener en cuenta que si bien el cloruro de calcio es más económico puede absorber tanta agua que se disuelve en ella, por lo cual si no está físicamente aislado de aquello que debe desecar podría impregnarlo.

4) En caso de tener que limpiar el CD, hacerlo con un paño suave y no abrasivo, preferentemente en dirección radial y ejerciendo un mínimo de presión. Si debe usarse algún detergente, debería ser muy suave y con base acuosa. Nunca usar alcohol ni solventes, ni intentar limpiar el lado de la laca: es innecesario ya que el haz láser penetra por el lado opuesto, y peligroso porque la laca es muy delgada y protege directamente la superficie activa.

5) Verificar con alguna frecuencia la integridad de los datos. Como los daños son siempre progresivos, la probabilidad de perder datos aumenta con los años.

6) Utilizar para su lectura y/o reproducción una lectora de buena calidad en la que no se hayan reproducido CD's sucios o con daños en su superficie, ya que podrían haber depositado en la óptica residuos capaces de alcanzar la superficie de otros CD's.

### **3.17.3.2 Disco compacto para datos generales (CD-ROM)**

El CD-ROM es una variante del CD que permite almacenar datos generales, no necesariamente de audio (Pohlmann, 2002). El formato es algo diferente y permite otras capas adicionales de control de errores. La información se organiza en 98 tramas de 24 bytes formando un bloque de datos de un total de 2352 bytes, de los cuales los primeros 12 constituyen un patrón de sincronismo. Luego hay un encabezamiento del bloque formado por 4 bytes organizados de la siguiente forma: 1 byte indica el minuto, 1 byte indica el segundo dentro de ese minuto, 1 byte indica el número de bloque dentro de ese segundo, y el último byte indica el modo, es decir el tipo de datos y de corrección de errores.

Cabe aclarar que cuando se habla de minutos y segundos, es en realidad una manera de referir por analogía al CD de audio, donde se supone que 1 s contiene 44 100 muestras estereofónicas de 16 bit, es decir, 176 400 bytes. Esta cantidad equivale a 75 bloques de 2352 bytes.

En cuanto al modo, hay tres modos: 0, 1 y 2. El modo 0 corresponde a un bloque de datos nulos. El modo 2 destina la totalidad de los  $2352 - 12 - 4 = 2536$  bytes restantes a datos del usuario. En este caso no hay corrección adicional de errores con respecto al CD. En el modo 1, el más interesante para este tipo de disco, se reservan 304 bytes para códigos adicionales de corrección de errores, restando sólo 2048 bytes para datos disponibles para el usuario. Aunque no profundizaremos en los detalles de la implementación de este control de errores, cabe comentar que el mismo provee una tasa de errores extremadamente baja (del orden de 1 bit entre  $10^{15}$  bit. Esto hace ideal este medio para almacenar información crítica, por ejemplo datos binarios de software, cuya alteración más insignificante podría tornar inútil a un programa.



### 3.17.3.3 Disco compacto grabable de escritura única (CD-R)

Los discos anteriormente descriptos son de fabricación industrial y exigen una cantidad mínima de réplicas (del orden de 100) para que se justifiquen económicamente, ya que se requiere la fabricación de costosos moldes que sólo se amortizan en la producción en escala. Por ese motivo no están al alcance de quien busca el almacenamiento de una copia única o de pocas copias de un material de audio digital cuya finalidad no es la diseminación pública.

En esos casos se recurre a los discos grabables CD-R, (Recordable CD) a veces también denominados CD-WO (*write-once CD*, es decir, CD de escritura única). Se trata de un tipo especial de discos que poseen un surco pregrabado en espiral de  $0,6\text{ }\mu\text{m}$  de ancho y un paso entre vueltas de  $1,6\text{ }\mu\text{m}$ . El surco está cubierto por un pigmento del tipo de las *cianinas*, *ftalocianinas* o *azos* (Pohlmann, 2002). Estos pigmentos reflectantes pierden su reflectancia al ser descompuestos en forma permanente por acción del calor de un láser. El surco pregrabado es necesario para poder disponer de un servomecanismo que permita posicionar con precisión la óptica móvil (láser y fotosensor) y, sobre todo, hacer compatible el disco para la lectura de en un lector de CD's pregrabados. Más aún, para lograr un adecuado sincronismo, el surco posee una ondulación con equivalente a una frecuencia de 22 050 Hz cuando se lo recorre a la velocidad normal (denominada 1x).

Debido a las previsibles diferencias en los procesos de fabricación (incluyendo la composición exacta del pigmento, su espesor, etc.), los discos CD-R poseen, antes de la primera vuelta útil del surco (más precisamente entre  $-35\text{ s}$  y  $-13\text{ s}$  con respecto a la primera vuelta), un espacio para calibración de la potencia óptica (OPC, *optimal power calibration*) denominado *área de calibración de potencia* (PCA, *power calibration area*). Allí se hacen varias tentativas de grabación con diversas potencias, cuyos resultados son luego leídos para determinar cuál de ellas produce la grabación más confiable, que es la que finalmente se selecciona. Hay espacio para realizar este ensayo 100 veces. Es importante tener esto en cuenta, ya que el CD-R permite grabar en varias sesiones antes de cerrar el disco, y en cada una se repite el proceso de optimización.

Un poco antes de la PCA, es decir antes de los  $-35\text{ s}$ , existe un *área de memoria de programa* (PMA, *program memory area*) destinada a una tabla de contenido provisoria que almacena las pistas ya grabadas mientras la grabación de la sesión aún no finaliza. Al terminar se transfieren a la *tabla de contenido* (TOC, *table of contents*), ubicada al comienzo del disco.

Los discos grabables admiten grabación en una sesión única o en multisesión. La sesión única completa el disco de una sola vez, incluyendo el área de entrada (*lead-in*) y el área de salida (*lead-out*). Una vez finalizado el disco, aunque no se hubiera consumido la totalidad del espacio disponible, ya no se podrán grabar más datos. La sesión múltiple permite tratar cada sesión (que puede contener más de una pista) en forma independiente, cada una con su área de entrada (con la correspondiente tabla de contenido) y de salida. Esto lleva a que por cada sesión se consuman alrededor de 13,5 Mb. En general estos discos tienen mayores problemas de compatibilidad con los reproductores de CD's pregrabados.

La vida útil de los discos CD-R, si son adecuadamente manipulados y conservados, depende esencialmente del tipo de pigmento. Las cianinas tienen una vida media estimada de entre 10 años y 100 años, mientras que las ftalocianinas tienen una vida media estimada de 240 años. En todos los casos se supone que el disco está protegido de la radiación ultravioleta.

Tanto para el respaldo de datos de audio como para su transporte, es preferible utilizar el modo de datos que el modo de CD de audio, opción que se selecciona en el software de grabación, ya que debido al control de errores reforzado la información almacenada es mucho más confiable.

#### **3.17.3.4 Disco compacto regrabable (CD-RW)**

El disco compacto regrabable CD-RW (también llamado CD-E) utiliza, como el CD-R, un surco premoldeado que sirve de guía. El medio de grabación, o *capa de registro*, es una aleación de plata, indio, antimonio y telurio que ofrece la característica de que calentada a una temperatura entre 500 °C y 700 °C pierde el estado policristalino y pasa a un estado amorfo, reflectivamente opaco, pero vuelto a calentar hasta 200 °C y enfriado con relativa lentitud recupera su estado policristalino reflectante original. Dicha capa se encuentra entre dos películas dieléctricas que obran de aislante térmico. El láser se utiliza pulsado entre el estado de alta potencia o media potencia que se requieren para la escritura o el borrado y el de baja potencia que se utiliza en la lectura, lo que permite un mejor control de la temperatura y un monitoreo y calibración en tiempo real de la potencia óptica del láser, ya que en el estado de baja potencia puede realizar la lectura de lo ya grabado y determinar si hace falta modificar la potencia.

Desafortunadamente, las diferencias en reflectividad entre los estados policristalino y amorfo no es tan alta como la que se logra en el CD pregrabado y en el CD-R, por lo que la compatibilidad de los CD-RW con los lectores o reproductores tradicionales de CD es menor, aunque no ofrecen dificultades en los lectores designados como RW.

Este medio no es tan utilizado como el CD-R, por varias razones. Primero porque es más costoso. Segundo porque sólo tiene interés en los casos en que se deban hacer continuamente registros de escaso valor documental y que, por lo tanto, pudieran descartarse rápidamente. Tercero porque están siendo rápidamente reemplazados por el *pen-drive*, dispositivo de memoria flash con conexión USB, como medio de almacenamiento masivo transitorio o para transferencia de una computadora a otra sin los trastornos y los riesgos de seguridad de la interconexión mediante redes locales.

#### **3.17.3.5 Minidisc (MD)**

Si bien el Minidisc (MD) se utiliza ahora muy poco, es un caso interesante por las tecnologías involucradas. Se trata de un pequeño disco óptico de 64 mm de diámetro que se encuentra protegido por una cubierta plástica. El grabador de Minidisc posee un mecanismo que abre una tapa deslizante, similar a la de los tradicionales diskettes, para acceder a la superficie. El disco es de policarbonato y contiene una película magnética amorfa de una aleación de terbio, hierro y gadolinio de alta coercitividad (entre 2500 Oe y 5000 Oe)<sup>67</sup>

La técnica de grabación es termo-magneto-óptica (Pohlmann, 1993). Posee un cabezal magnético frente a una cara del disco y un cabezal óptico láser frente a la otra. El cabezal magnético establece un campo magnético que es, en sí mismo, insuficiente para modificar el estado de magnetización de la película magnética. En la fase de escritura, el láser calienta el material magnético hasta su temperatura de Curie (alrededor de 180 °C), donde el material pierde su coercitividad (desciende hasta unos 80 Oe) y es magnetizado o desmagnetizado fácilmente por el cabezal. La combinación del campo magnético y

---

<sup>67</sup> La *coercitividad* o *fuerza coercitiva* es la intensidad de campo magnético inverso que debe aplicarse a un material ferromagnético que ha sido magnetizado llevándolo primero a la saturación magnética y luego retirando la excitación, para desmagnetizarlo. Cuanto más alta sea, más difícil será provocar un cambio en su magnetización.

el láser permite un tamaño de marca magnética mucho menor que si se usara un cabezal solo, dado que el tamaño con que se lo puede fabricar es mucho mayor que el tamaño requerido. El láser permite el calentamiento de un punto muy pequeño del sustrato magnético.

La fase de lectura utiliza el denominado *efecto Kerr*, por el cual un campo magnético introduce una rotación en la polarización del haz láser. Este cambio de polarización se convierte en un cambio de intensidad que es detectado en forma análoga a la del CD convencional pregrabado. En este caso se utiliza el mismo láser, pero con una potencia mucho menor, con lo que el sustrato magnético no se calienta y por lo tanto no corre riesgo de desmagnetizarse.

La codificación de la información enviada al el disco obedece a los mismos principios del CD, es decir, hay una codificación de corrección de errores CIRC y una modulación EFM. Sin embargo, existe un paso previo adicional, que es la *compresión de audio con pérdida*. Esto es necesario ya que uno de los objetivos originales del diseño fue lograr que el Minidisc pudiera almacenar la misma cantidad de audio que un CD, es decir, 74 min. El problema es que el área disponible resultaba mucho menor, por lo cual la única forma posible era desarrollar un sistema de compresión perceptual de la información que preservara la calidad auditiva de la señal a costa de sacrificar datos.

Un compresor perceptual utiliza inteligentemente el fenómeno de enmascaramiento para eliminar aquellas partes de la señal que quedarían enmascaradas por otras y que, por lo tanto, no serán audibles. Para lograrlo, primero se descompone la señal en una serie de bandas espectrales (denominadas *sub-bandas*). Dado que cada sub-banda es más angosta que la banda original de audio, puede ser submuestreada sin dejar de cumplir con el teorema de muestreo (ver Apéndice 10). Por lo tanto, a pesar de que hay más señales (una por cada sub-banda), cada una tiene una menor tasa de muestreo, de modo que la tasa de muestreo global se mantiene igual a la original. Lo mismo sucede con la resolución requerida para representar cada banda, por lo cual no hay todavía reducción de datos.

El siguiente paso es determinar, mediante un algoritmo que implemente los modelos psicoacústicos de enmascaramiento, el umbral enmascarado de cada sub-banda causado por el contenido de las otras.<sup>68</sup> Una vez hecho esto, y considerando que el umbral correspondiente a la frecuencia de cada sub-banda equivale a un ruido, se puede determinar la relación señal/ruido entre el contenido real de la sub-banda y su umbral enmascarado. Por ejemplo, si en una banda el umbral enmascarado fuera de 40 dB y el nivel de la señal fuera 70 dB, tendríamos una relación señal/ruido de 30 dB, que se puede representar con  $n = 30/6 = 5$  bit. Habremos ahorrado, así, 11 bit con respecto a lo que hubiera sido la representación original con 16 bit.

¿Qué sucedería si hubiéramos mantenido los 16 bit? Simplemente estaríamos representando fenómenos sonoros en esa banda que resultan inaudibles por estar debajo del umbral.

En el caso del Minidisc se utiliza el sistema de compresión llamado ATRAC (*Adaptive Transform Audio Coding*, es decir, Codificación de Audio mediante Transformada adaptativa) que reduce la tasa de bits desde la original correspondiente a 44100 Hz, 2 canales y 16 bit por muestra, es decir,  $2 \times 16 \times 44100 = 1\,411\,200$  bit/s, hasta 292 000 bit/s. ATRAC divide la señal de audio original en tres bandas: [0 kHz, 5,5125 kHz], [5,5125 kHz, 11,025 kHz] y [11,025 kHz, 22,05 kHz], mediante

<sup>68</sup> Los modelos de enmascaramiento proporcionan fórmulas para obtener el umbral auditivo en cualquier frecuencia a partir de los niveles de las otras bandas.

dos *filtros espejo en cuadratura*,<sup>69</sup> uno de ellos que separa las bandas por encima y por debajo de 11,025 kHz y el otro que divide esta última entre las frecuencias por encima y por debajo de 5,5125 kHz. Cada una de estas bandas se submuestra para alcanzar la tasa de muestreo crítica correspondiente al ancho de banda correspondiente, esto es, 22 050 Hz para la banda superior y 11 025 Hz para las dos bandas inferiores. Se aplica aquí el teorema de muestreo para señales pasabanda (ver Apéndice 10).

Cada banda es sometida a una *transformada coseno discreta modificada* (MDCT), una forma de descomposición espectral especialmente apropiada para el tratamiento de cuadros de audio solapados (en el Minidisc se solapan al 50 %). Al igual que la *transformada discreta de Fourier*, la resolución espectral varía inversamente con la resolución temporal, debido al principio de incertidumbre tiempo-frecuencia. Por esa razón, el tamaño o longitud del cuadro se elige, en función del comportamiento transitorio de la señal. En aquellos fragmentos de audio en los que se detecta un comportamiento transitorio muy rápido (por ejemplo en un sonido impulsivo) y, por consiguiente, se requiere buena resolución temporal, se toma una longitud de análisis de 32 muestras. Para la banda superior equivale a un tiempo de 1,451 ms, mientras que para las dos bandas inferiores equivale a 2,902 ms.

Cuando, en cambio, la señal varía lentamente, se toma una longitud del cuadro de análisis de 256 muestras para la banda superior y 128 muestras para las dos bandas inferiores. En las tres bandas equivale a 11,61 ms. Esta adaptabilidad de algunos parámetros del codificador es interesante porque contempla no sólo el enmascaramiento espectral sino el enmascaramiento temporal.<sup>70</sup> Cabe destacar, asimismo, que la selección de los tamaños largo y corto de la ventana se hace en forma independiente para cada una de las tres bandas, de modo que podría suceder, por ejemplo, que la banda inferior tenga un transitorio rápido mientras que las otras dos sean lentamente variables (por ejemplo en un estampido de baja frecuencia).

Una vez realizada esta conversión de la señal del dominio tiempo al dominio frecuencia, viene la etapa de cuantificación selectiva o asignación de bits basado en el modelo psicoacústico de umbral enmascarado. El resultado es que cada coeficiente espectral está representado con menos bits, cantidad que se actualiza cada 11,61 ms. Es interesante destacar que el algoritmo en sí no es parte de la especificación ATRAC, lo cual ha permitido una evolución a lo largo de los años con notables mejoras en los resultados.

Los valores así obtenidos se empaquetan en tramas de 512 datos, incluyendo un encabezamiento que informa si se usó el modo corto o largo, la cantidad de bits y el factor de escala que permite asignar un valor de referencia al bit menos significativo. Luego de esto se pasa a la codificación CIRC y la modulación EFM, momento en el que la información se puede transferir al disco.

<sup>69</sup> Los filtros espejo en cuadratura son filtros que dividen el espectro de una señal digital a la mitad de la frecuencia de Nyquist pero lo hacen de tal manera que si cada banda se submuestra a la mitad, los alias residuales que quedan en cada una de las bandas son tales que al sumar las señales se cancelan, recuperándose completamente la señal

<sup>70</sup> El enmascaramiento temporal se refiere a que un sonido fuerte puede enmascarar a uno más débil que sucedió unos milisegundos antes. Este comportamiento parece paradójico, pero no lo es. Se explica por el retardo que existe entre que el oído interno capta el sonido y se forma la sensación en la corteza cerebral. En este caso se aprovecha el fenómeno de esta forma: si se toma una ventana larga y se produce un transitorio al final, al cuantizar aparece un ruido de cuantización en toda la ventana. Si la ventana es larga, ello implica que el ruido empieza mucho antes que el transitorio y se percibe claramente como un soplo. Si la ventana es corta el soplo será enmascarado temporalmente por el transitorio.

### 3.17.3.6 Digital Versatile Disc (DVD)

El *Digital Versatile Disc* (Disco Versátil Digital, DVD)<sup>71</sup> es un disco de iguales dimensiones y aspecto que el CD (en sus dos versiones de 80 mm y 120 mm de diámetro) pero con una mayor capacidad de almacenamiento. Cada sector de un DVD posee 2048 bytes de datos de usuario que se incrementan a 2418 bytes al incluir los datos complementarios de control de errores. A diferencia del CD, el DVD puede tener una cara o dos caras, y además cada una de ellas puede tener uno o dos estratos o capas activas, siendo semitransparente la más próxima a la superficie externa expuesta al haz láser y totalmente reflectante la más profunda. Cada capa permite en principio una capacidad de 4,7 GB, por lo cual los discos de doble capa y doble cara llegan a tener capacidades de 17 GB. La gran capacidad se debe a que la longitud del láser es menor (635 nm ó 650 nm)

El DVD admite las variantes de una escritura (DVD-R) y grabables (DVD-RW), lo cual constituye un interesante medio para almacenar no sólo gran cantidad de archivos de audio, sino además archivos de varias horas de duración ininterrumpida.

### 3.17.3.7 Blu-ray Disc (BD)

El *Blu-ray Disc* (BD) es otro disco óptico, denominado así pues utiliza un láser azul<sup>72</sup> de 405 nm que reduce el tamaño del punto de enfoque y por lo tanto de las fosas. Esto hace posible ampliar la capacidad hasta 25 GB (o incluso 33 GB) por cada capa. Así, existen discos Blu-ray de 25 GB, 50 GB, 100 GB y 128 GB de capacidad. Las características físicas del Blu-ray son similares al CD y el DVD, pero el Blu-ray tiene un recubrimiento protector de gran resistencia a las raspaduras de unos 2 µm de espesor, formado por una resina de sílice coloidal disperso, endurecible mediante luz ultravioleta. Este material exhibe, además, un buen rechazo a la adherencia de la grasitud de huellas dactilares, muy interesante dado que en este caso el haz tiene un diámetro pequeño (de 0,13 mm) sobre la superficie externa del disco y es por lo tanto más afectado por la contaminación de la superficie.

La corrección de errores que utiliza el BD se denomina *código de larga distancia* (*long distance code*), similar a un código entrelazado de Reed-Solomon con un entrelazado a una distancia equivalente a 64 kB (Wikipedia, 2013c; Blu-ray Disc Association, 2012).

También existen variantes grabables (BD-R) y grabables/borrables (BD-RE) con iguales capacidades que los anteriores. Nuevamente, se trata de un medio muy interesante para almacenar grandes cantidades de audio, aunque la tecnología es todavía demasiado nueva para poder adelantar cuál es la confiabilidad en términos reales de almacenamiento a largo plazo.

### 3.17.3.8 Holographic Versatile Disc (HD)

Por último, tenemos el *Holographic Versatile Disc* (Disco Versátil Holográfico, HVD). Es un disco del tamaño de un CD o aun menor con capacidad para almacenar varios TB de información. Utiliza para ello una técnica denominada *holografía colinear*. Utiliza dos haces láser, rojo y verde, colimados mediante un dispositivo óptico de modo de formar un único haz. El rojo se utiliza para leer información requerida por el servomecanismo que controla la velocidad y la alineación del haz con respecto a la pista, que se encuentra en una capa profunda similar a la de un CD. El haz verde lee una capa casi superficial que codifica, mediante franjas de interferencia láser, la información. Este medio podría constituir, en un futuro próximo, un medio revolucionario para

<sup>71</sup> Cuando se introdujo en 1995 se denominaba Digital Videodisc. Sin embargo, ni esta designación ni la actualmente en boga fueron nunca oficiales.

<sup>72</sup> La ortografía *Blu* fue elegida, al parecer, para poder registrar la marca Blu-ray.

almacenar grandes volúmenes de información a largo plazo con fines documentales. Si bien ha sido presentado al público en algunas exposiciones de tecnología e incluso normalizado por la ECMA (ECMA-377), aún no ha tenido difusión comercial, posiblemente por su elevado costo.

#### **3.17.4 Cinta de audio digital (DAT)**

Mencionaremos brevemente el sistema de grabación digital en cinta magnética, comúnmente conocido como DAT. Debido a los requisitos de ancho de banda del audio digital (luego de la aplicación de redundancia se acerca a los 4 MHz), no es posible realizar la grabación con un cabezal magnético fijo sobre una cinta de baja velocidad. El sistema DAT utiliza un cabezal rotativo que gira a 2000 rpm realizando un barrido helicoidal sobre la cinta. El resultado es una serie de pistas que cruzan la cinta de abajo hacia arriba con un ángulo de unos  $6^\circ$  respecto al eje longitudinal de la cinta. La separación entre ellas es de  $13,5\text{ }\mu\text{m}$ . Igual que en el CD, se utiliza codificación CIRC, pero en lugar de modulación ocho a catorce usa *modulación ocho a diez*.

En la actualidad ya prácticamente ha caído en desuso al ser sustituido por los grabadores en tarjetas de memoria flash que lo superan ampliamente en comodidad y bajo consumo. Una desventaja muy grande del DAT, así como de cualquier sistema basado en cinta es la imposibilidad de acceso aleatorio ya que el medio es intrínsecamente secuencial.

Como ventaja debe mencionarse que la mayoría de los equipos DAT portátiles han sido de excelente calidad, con excelentes especificaciones y resultados en los ensayos de laboratorio. Los que se utilizan actualmente son los equipos adquiridos antes de la proliferación de los grabadores en memoria flash.

#### **3.18 Sistemas de archivos**

En la mayoría de los medios de almacenamiento masivo de datos existe un mínimo para la cantidad de información que es posible intercambiar, tanto en el proceso de escritura como en el de lectura, debido a características o limitaciones estructurales del medio o a la necesidad de un mecanismo de corrección de errores. En un disco duro tradicionalmente ha sido un sector de 512 bytes, pero en las unidades más nuevas y de mayor capacidad es de 4096 bytes. Los sistemas operativos pueden introducir agrupamientos o bloques todavía mayores, denominados *clusters* o *unidades de asignación*, cada uno de los cuales abarca una cierta cantidad de sectores (por ejemplo, 64 sectores para un disco de gran capacidad). Esto se debe a la forma en que se guardan los archivos en el disco y a las limitaciones que esto impone.

En efecto, el acceso a cualquier bloque de información requiere direccionarlo, y el direccionamiento requiere espacio, sobre todo si los archivos no se encuentran íntegramente contenidos en un conjunto de bloques sucesivos. Esta situación se da cuando se empiezan a borrar archivos, ya que el espacio liberado puede no ser suficiente para el próximo archivo a escribir, por lo tanto una parte del mismo se alojará en dicho espacio liberado y el resto en una ubicación no consecutiva. En un disco de gran capacidad, la cantidad de sectores puede ser tan grande que su direccionamiento llega a ser prohibitivo. Por ejemplo, en un disco de 1 TB tendremos alrededor de  $2 \times 10^9$  sectores, cada uno de los cuales requerirá  $\log_2(2 \times 10^9)$ , es decir 32 bits o 4 bytes para su representación, por lo cual el direccionamiento de la totalidad del espacio disponible requiere una suerte de índice de 8 GB, es decir un 0,8 % del espacio total disponible. La agrupación de los sectores en unidades de asignación de mayor capacidad reduce el tamaño de dicho índice. Por ejemplo, tomando una unidad de asignación de 64 sectores, el índice se reducirá

a unos 125 Mb, es decir, un 0,0125 %. Una desventaja es que cuando se requiere guardar archivos pequeños se desperdicia espacio, ya que cualquier archivo menor de 32 kB requerirá un cluster completo. Este espacio desperdiciado se denomina *slack*.

Un enfoque más moderno consiste en brindar la posibilidad de reservar espacio contiguo en un disco para un archivo, independientemente de que éste sea ocupado. Esto se realiza mediante el concepto de *extent* (Wikipedia, 2013; McDougall, 1999) y evita la fragmentación del archivo. Esto es muy útil en grandes archivos que deben ser leídos secuencialmente, como los archivos de audio o de video digital.

Existen diferentes sistemas de archivo. Entre ellos, los más difundidos son los FAT 32, NTFS, HFS+, EXT3 y EXT4, exFAT y UDF. Veremos una breve introducción a FAT 32, por ser a la vez el más sencillo, el de mayor compatibilidad con todos los sistemas operativos, y el que se implementa por lo general en las tarjetas de memoria flash, tales como la SD y sus variantes (Longares, 2011; Microsoft, 2000; Kholodov, 2010). Esto último es de especial interés ya que los nuevos grabadores digitales utilizan cada vez más este medio de registro de audio digital.

### 3.18.1 FAT 32

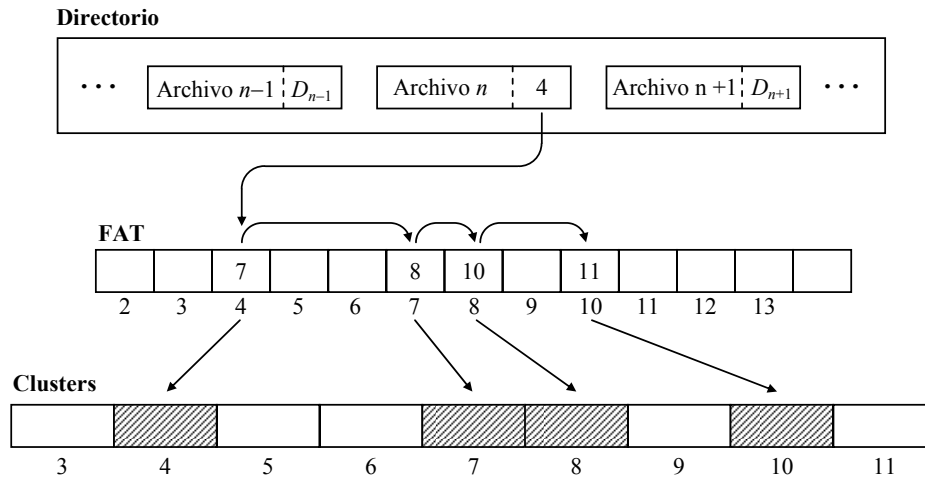
FAT32 (*File Allocation Table* de 32 bit) es una actualización de los antiguos sistemas de archivo FAT del sistema operativo DOS (Disk Operating System). La idea de este sistema es la existencia de una tabla de asignación de archivos (la FAT) formada por sucesivas entradas de 32 bit (4 byte), cada una de las cuales está asociada a un cluster. Para cada unidad de almacenamiento (disco duro, tarjeta SD), cada cluster está formado por una cantidad determinada de sectores. A su vez, cada archivo está formado por una cierta cantidad de clusters que pueden o no ser adyacentes. La organización jerárquica de los archivos se realiza a través de un árbol de directorios, donde para el usuario cada directorio es un contenedor<sup>73</sup> que puede contener archivos u otros directorios (generalmente llamados subdirectorios). A nivel físico, sin embargo, un directorio es sólo un archivo cuyo contenido son las *entradas de directorio*, es decir unos pocos campos que indican el nombre del archivo o subdirectorio (11 caracteres ASCII), su tamaño total en bytes y el número del primer cluster del archivo, algunos atributos y la fecha y hora de creación, última escritura y último acceso. Cada entrada de directorio ocupa 32 bytes, salvo en los nombres largos, donde ocupa un múltiplo de 32 bytes.

El número del primer cluster del archivo es en realidad un puntero doble, ya que por un lado indica dónde empieza el contenido del archivo, pero por otro lado indica la entrada homóloga de la FAT. El contenido de esta entrada de la FAT es también un puntero de doble propósito: indica el número correlativo de la próxima entrada en la FAT y del correspondiente cluster del archivo. Este proceso se ilustra en la figura 64.

Como cada cluster existente en el disco o unidad está representado por una entrada en la tabla, el anterior es un método simple de localizar todos los clusters que constituyen un archivo, no importa cuán fragmentado se encuentre éste. Dado que FAT 32 usa palabras de 32 bit para representar los números de clúster, la máxima cantidad disponible es  $2^{32} = 4\,294\,967\,296$ , que corresponden , para clusters de 32 kB a 140 TB. Sin embargo realmente se utilizan 28 bit, por lo que la unidad más grande que puede direccionarse es de 8,8 TB. Sin embargo, el archivo más grande que puede representarse es de 4 GB, ya que el campo del directorio correspondiente al tamaño de archivo es de 32 bit. Esto podría imponer una limitación para grabaciones sonoras a largo plazo. Un archivo estéreo con una resolución de 16 bit requiere 635 MB por

<sup>73</sup> Esta imagen es tan común que se suele denominar *carpetas* a los directorios. Preferimos denominarlos directorios ya que transmite mejor la idea de una lista con información sobre ítems de una clase.

hora de grabación, con lo cual la grabación ininterrumpida más larga en sistema FAT es de unas 6 h 45 min.



**Figura 64.** Diagrama esquemático del sistema FAT de direccionamiento de clusters. En la entrada de directorio correspondiente al archivo  $n$  aparece el número del primer cluster del archivo,  $D_n = 4$ , que es igual al número de entrada de la FAT correspondiente. En esa entrada se encuentra el siguiente número de entrada de la FAT, 7. NOTA: el directorio es también un archivo y su contenido forma parte de uno o más clusters. Aquí se lo separó para mayor claridad.

Los nombres de archivo o de directorio admiten dos versiones: una corta y una larga. La corta es de 11 caracteres ASCII correspondientes a mayúsculas, números y unos pocos caracteres adicionales, como + - \_ ~ ( ). Se supone convencionalmente que hay un punto (no representado en el campo del nombre) entre los 8 primeros caracteres y los tres últimos, que determinan el carácter de *extensión* para estos últimos.

La versión larga de los nombres fue agregada a posteriori a los primeros sistemas FAT, por lo cual, para lograr retrocompatibilidad con software desarrollado para aquellas versiones, se conservó la entrada de directorio con nombre corto y se agregaron nuevas entradas que admiten mayor cantidad de caracteres expresables con dos bytes en código Unicode (aquí la interpretación de los mismos depende de la selección de idioma y código de página de cada una computadora en particular). Un software antiguo mostrará los nombres de archivo “Este es un archivo largo.txt” y “Este es un archivo más largo.txt” como “ESTEES~1.TXT” y “ESTEES~2.TXT”, donde existe una serie de convenciones como eliminar los espacios y puntos, convertir a mayúsculas, y agregar ~ $n$  donde es un número correlativo para diferenciar todos los archivos que empiezan igual y cuya diferencia está más allá de los 8 caracteres.

### 3.18.2 NTFS

NTFS (New Technology File System) es el utilizado por Windows de Microsoft a partir de las versión NT (NT, XP, Vista, etc.). Es una derivación del HPFS (High Performance File System). A diferencia de FAT 32, este sistema de archivos permite en forma



nativa nombres de archivo largos de hasta 255 caracteres de 2 bytes cada uno. El sistema respeta las mayúsculas y minúsculas pero no las distingue (cosa que sí sucede en sistemas operativos como Unix o Linux). Cualquier caracter es admisible salvo ? " / \ < > \* | y :, ya que estos son utilizados con funciones especiales, por ejemplo formar parte de la ruta a través del árbol de directorio, como comodines, etc. La descripción siguiente está basada en (Kozierok, 2004).

NTFS utiliza un cluster de entre 512 bytes y 4096 bytes. La organización de la información de un archivo es por atributos, donde incluso el contenido del archivo es un atributo. Los atributos comienzan guardándose en un archivo oculto del sistema denominado *Master File Table* (MFT, tabla maestra de archivos). Al crearse cada archivo se genera un registro en esta tabla de un tamaño igual al cluster del volumen pero con un mínimo de 1024 bytes. Este registro contiene algunos atributos del archivo, denominados *atributos residentes*. En el caso de archivos pequeños, el atributo de contenido del archivo podría caber dentro del propio registro asignado, reduciendo el espacio desperdiciado. En general, no obstante, el espacio del registro no alcanza para la totalidad de los atributos, lo que motiva que los mismos continúen como *atributos no residentes* fuera de dicho registro. Pueden almacenarse en otros registros dentro de la propia MFT, o bien fuera de ella, en los denominados *extents*. Cabe notar que el universo de posibles atributos no se restringe a unos pocos establecidos por el sistema, sino que el usuario podría crear nuevos atributos.

El tamaño de la MFT depende, en principio, de los archivos que se encuentren almacenados en el sistema, por lo cual podría crecer considerablemente. Es incluso posible, como en cualquier archivo, que se encuentre fragmentada, pero al formatear una unidad NTFS, se reserva un espacio de 1/8 de la capacidad total del disco para lograr que la MFT crezca sin fragmentación y así evitar degradar el rendimiento del sistema. Aunque parece un desperdicio de espacio, no lo es, dado que parte o todo el contenido de los archivos puede estar dentro de la tabla. Esto último empieza a suceder cuando el resto del espacio disponible se llenó. También puede ocurrir que el espacio reservado se llene íntegramente con la MFT antes de que el espacio libre se haya agotado. En ese caso la MFT continuará creciendo, ya fragmentada, donde quede lugar vacante. Es importante tener esto en cuenta, ya que en un disco en el que la capacidad ha sido ocupada en una fracción 7/8, el rendimiento empieza a bajar.

La estructuración de los archivos se basa en una estructura de datos típica de las bases de datos relacionales denominada *árbol B+* (*B+ tree*), cuya principal característica es la posibilidad de almacenar los datos en bloques (clusters) de manera de lograr gran eficiencia en su recuperación, donde los bloques tienen punteros hacia otros bloques. Se logran tiempos de acceso secuencial, inserción y borrado que varían logarítmicamente con la cantidad de datos. Los detalles de este proceso pueden consultarse en la bibliografía específica, por ejemplo (Bayer et al, 1970; Bayer et al. 1972, Wikipedia, 2013a, Wikipedia, 2013b).

El registro de un archivo en la MFT comienza con un encabezamiento que incluye punteros a los atributos, entre ellos el atributo de información estándar (fechas y horarios y los atributos tradicionales presentes en la FAT), el atributo de nombre de archivo (que puede incluir más de un nombre, tal como el nombre corto, etc.), el atributo de datos, y el atributo de descriptor de seguridad, que incluye el control de acceso (por ejemplo los permisos de lectura, escritura, ejecución, borrado, e incluso el permiso de cambiar los permisos). También existe el concepto de *auditoría*, que permite llevar un registro histórico de diferentes tipos de eventos, como accesos a directorios y archivos o inicios de sesión. Una última característica de interés en materia de seguridad es la encriptación, dis-

ponible a partir de la versión 5.0 de NTFS. Éste es un dispositivo de seguridad de mayor confiabilidad que los permisos, ya que utilizando herramientas de acceso de bajo nivel al disco, es teóricamente posible acceder a cualquier información no encriptada.

Otra posibilidad que brinda NTFS es la compresión de archivos o directorios. Si bien se produce una disminución del rendimiento por el uso de recursos para comprimir y descomprimir, en algunos casos ello se ve compensado por el hecho de que al ser menor el tamaño y la fragmentación del archivo, el tiempo de acceso al disco se reduce. En esta línea, también permite una forma muy básica y rápida de compresión que es el manejo de *archivos raros* (*sparse files*), es decir, archivos que tienen grandes zonas “vacías”, es decir, con bytes nulos. En el caso de audio, por ejemplo, archivos con grandes intervalos de silencio. En el uso de software matricial (por ejemplo Scilab) se deben inicializar vectores o matrices de grandes dimensiones para mejorar el rendimiento. Esta inicialización normalmente es con valores nulos. El sistema de archivos permite reemplazar ese contenido nulo por metadatos, y aprovechar ese espacio de disco con otros datos.

Este sistema de archivos propietario tiene también una contrapartida libre con licencia GPL (*General Public License*, Licencia Pública General), el NTFS-3G, que permite leer y guardar archivos en unidades NTFS en Linux, OS X, Android y otros sistemas operativos. (Tuxera, 2013).

### 3.18.3 EXT4

EXT4 (*Fourth Extended File System*) es el sistema de archivos empleado por las distribuciones actuales de Linux y por el sistema operativo Android, en el que se basan las computadoras de tipo *tablet* y los teléfonos celulares inteligentes.

Permite el direccionamiento de unidades hasta  $2^{60}$  bytes (es decir, 1000 TB) y archivos de hasta 16 TB. Si bien los discos actuales (2013) son de unos pocos TB, indudablemente la capacidad de almacenamiento continuará creciendo. Para ponerlo en perspectiva, un archivo de audio de 16 TB equivale a 1154 días corridos, es decir 3 años y 2 meses de grabación estéreo con calidad CD, apto para los más ambiciosos estudios a largo plazo. Por otra parte, cada directorio admite hasta 64 000 subdirectorios.

También permite el uso de *extents*, que, como ya se ha dicho, corresponden a un grupo de sectores adyacentes sucesivos que, a diferencia de los clusters, no necesariamente tienen un tamaño fijo. Esto es especialmente útil en audio digital, donde el acceso a grandes archivos muy fragmentados supone una pérdida de rendimiento debido a la necesidad de reposicionar constantemente el cabezal de lectura y escritura del disco duro. EXT4 admite extents de hasta 128 MB.

Es posible reservar espacio de disco para un archivo sin necesidad de hacer crecer efectivamente el archivo mediante el agregado artificial de ceros. Esto es muy útil cuando se prevé que un archivo va a ir creciendo, por ejemplo cuando bajo el mismo nombre se vayan guardando versiones cada vez de mayor extensión. Esta reserva de espacio permite minimizar la fragmentación. Evita también que otros procesos que pudieran estar ejecutándose en segundo plano lleguen a consumir el espacio necesario, impidiendo así completar una larga operación (por ejemplo, el guardado en disco de una matriz con los espectros FFT de sucesivos cuadros de un extenso archivo de sonido).

Por otra parte, EXT4 incorpora la posibilidad de optimizar la distribución de un archivo en el disco mediante la técnica de asignación demorada. Consiste en no guardar directamente al disco un archivo que está siendo todavía modificado por algún proceso sino guardarlo temporalmente en un caché en memoria hasta que se sepa más del archivo. Entonces el sistema de archivos busca la mejor ubicación, es decir, la que al momento de guardar requiere la menor fragmentación posible. Esto significa no empezar a

escribir el archivo en el primer bloque vacío (disponible) sino en el más conveniente. Los pequeños huecos que van quedando se rellenan con archivos más pequeños o con archivos más grandes cuando ya no quedan espacios grandes contiguos. Existe algún riesgo de pérdida de datos en caso de un colapso del sistema o un corte del suministro eléctrico

La medición de tiempos en cualquier sistema de archivos está afectada por el problema de *granularidad*, es decir, la mínima resolución. En FAT 32, por ejemplo, es de 10 ms. En EXT4, se extiende a los nanosegundos por un lado. También se extiende el límite de los años representables al año 2242.

Por último, el sistema de archivos provee mecanismos para acelerar las operaciones de chequeo del sistema de archivos (Wikipedia 2013d).

### 3.18.4 HFS+

HFS+ (*Hierarchical file System*) es el sistema que utilizan los sistemas OS X de Macintosh. Este sistema de archivos utiliza sectores de 512 bytes agrupados en unidades de asignación (Wikipedia, 2013e) direccionadas mediante 32 bit, por lo que la cantidad máxima de bloques direccionables es  $2^{32}$ . La cantidad efectiva de bloques de una unidad depende de la capacidad total de la misma. Existe un archivo de asignación (*Allocation File*), también denominado *bitmap* (no debe confundirse con el formato BMP de imágenes) que asocia un bit con cada bloque de asignación que indica su estado de ocupación: 0 para un bloque libre y 1 para uno ocupado. El tamaño real se redondea hacia arriba para ocupar un número entero de sectores (Developer Connection, 2007).

Existe también un archivo de catálogo en la forma de árbol B\* (*B\*-Tree*). El árbol B\* es una forma de organizar la información en *nodos*, cada uno de los cuales tiene 4 kB. Los nodos contienen un descriptor, un grupo de registros que contienen la información propiamente dicha, y una serie de *offsets* que direccionan los registros. El descriptor contiene información del tipo de nodo, su nivel en el árbol, el número de registros que contiene y los enlaces hacia atrás y hacia adelante. La finalidad de un árbol B\* es hacer más eficiente la recuperación, el borrado y la sustitución de información. Varias estructuras dentro de HFS+ se organizan como árboles B\*.

Los archivos en HFS+ están en realidad bifurcados, de modo que por cada archivo hay una bifurcación de recursos (*resources fork*) y una bifurcación de datos (*data fork*). La bifurcación de recursos está encabezada por un mapa de recursos. Por ejemplo, para un archivo ejecutable los recursos pueden describir los menús, diálogos, íconos, utilización de memoria, el propio código, etc. Los archivos de documento tienen en general la bifurcación de recursos vacía. Los ejecutables, en cambio pueden carecer de datos. (Developer Connection, 2007)

Los nombres de archivos y carpetas pueden tener 255 caracteres Unicode. Algunos atributos del sistema de archivos son un alias, un ícono personalizado para archivos y carpetas, el hecho de estar localizado en el escritorio, carácter de archivo de sistema, la existencia o no de un recurso INIT, es decir, modificación de la rutina del sistema, el carácter de *bundle* que permite que una carpeta aparezca como un único archivo. El máximo tamaño de archivo es de  $2^{63}$  bytes y en general. A diferencia de otros sistemas de archivo, siendo esta cifra tan grande, quedan limitados sólo por el tamaño de la unidad (Developer Connection, 2004).

### 3.18.5 UDF

UDF (*Universal Disk Format*) es el sistema de archivos utilizado por los DVD y otros medios ópticos. Está definida por la Optical Standard Technology Association

(OSTA, 2000), que a su vez es una derivación de lo estipulado por ECMA (ECMA-167). Los nombres de archivo pueden utilizar diversos códigos, como el ASCII, el Unicode y el ECMA-094 y alcanzar los 255 bytes. La estructura de directorio admite profundidades cuya ruta tenga hasta 1023 bytes. El tamaño de archivo puede ser de 16 exabytes ( $2^{60}$  bytes). Permite registrar las fechas de creación, archivado, modificación, cambio de atributos y acceso.

Estructuralmente posee sectores que dependen del medio específico, organizados en bloques y *extents* de hasta  $2^{30} - 1$  bytes. Dependiendo de si el medio admite escritura, el sistema permite la lectura y escritura aleatorias. El sistema tiene tres variantes. La primera, designada como *plana* (*plain*), es aplicable a discos que permiten acceso aleatorio de lectura y escritura, como los DVD-RW. La segunda, conocida como de *escritura incremental*, se aplica a discos de escritura solamente, como CD-R y DVD-R. Utiliza una *tabla de asignación virtual* (*virtual allocation table*). En este caso debe haber una precompilación del material para ser registrado incrementalmente sobre el disco. Sin embargo, se permite la escritura en paquetes en diferentes sesiones incrementales, sustituyendo la tabla en cada nueva sesión. Se permite el borrado pero sólo el borrado lógico, es decir, la nueva tabla no contendrá una entrada que direcciona al espacio ocupado originalmente. Sin embargo, dado que cada sesión utiliza una nueva versión de la tabla, se consume espacio ya que la vieja tabla no puede borrarse físicamente. En un CD-R cada reescritura implica perder 13 MB. La tercera variante (*sparse*) es aplicable a los discos regrabables como el DVD-RW. Consiste en minimizar las reescrituras para evitar el desgaste del medio. En este caso, al reescribir un archivo la operación se realiza preferentemente en un espacio vacío. La tabla correspondiente lleva un registro de los errores causados por el desgaste. (Wikipedia, 2013f).

### 3.19 Preservación a largo plazo

Un problema de gran importancia es la preservación de las señales sonoras (y, en general, todo tipo de información digital) a largo plazo con varios objetivos: 1) Amortizar el costo y el esfuerzo que implica cualquier medición; 2) Documentar campañas de medición realizadas; 3) Preservar testimonios audibles del paisaje sonoro.

Los principios físicos de la lectura de un disco de vinilo, por ejemplo, o de una grabación magnética analógica bien preservada, son ampliamente comprendidos y fáciles de implementar. Así, las primeras grabaciones sonoras realizadas por Thomas Alva Edison en 1877 sobre lámina de estaño se han conservado y restaurado (National Park Service, s/f). En cambio, las modernas técnicas digitales, incluyendo los complejos códigos de corrección de errores, las representaciones espectrales, los códigos comprimidos con estrategias de difícil comprensión, la atomización de archivos, etc., pueden implicar un verdadero desafío futuro para quien se proponga recuperar la información respectiva.

A ello se agrega un problema todavía más grave, como es el hecho de que los medios de almacenamiento comúnmente disponibles son muy delicados debido al tamaño microscópico de sus bloques de información, lo que los hace fácilmente propensos a sufrir pérdidas masivas de datos.

Otra particularidad es que la información digital es virtual, es decir, no está asociada indisolublemente al medio en el que se almacena. Esto implica que la preservación del objeto físico, el contenedor, no necesariamente garantice la preservación de los datos. Así, mientras que a lo largo de la historia se ha conservado mucha información que a posteriori pudo ser recuperada y descifrada, actualmente se genera mucha más información pero se conserva muy poca en soportes físicos resistentes. En su mayoría se conservan en dispositivos de corta vida útil, como discos rígidos, discos ópticos y memorias flash.

Se han identificado varios enfoques para la preservación a largo plazo de la información digital (Peres, 2007; Thibodeau, 2002). El primero es no hacer nada, dejar que los futuros *arqueólogos digitales* se encarguen de descubrir la manera de hacerlo. Esto es prácticamente una garantía de que la información se pierda definitivamente, ya que los costos que tal recuperación detectivesca implica pueden ser tan altos que no se justificaría aplicarlos salvo en casos muy especiales.

El segundo es el enfoque *museológico*: preservar en estado de óptimo funcionamiento tanto los medios de soporte de la información como los equipos de cómputo necesarios para ello, incluyendo los periféricos y el software requerido. Una dificultad para esto es la *obsolescencia planificada* que hace que los equipos, incluso bien conservados, dejen de funcionar correctamente al cabo de unos pocos años. La posibilidad de repararlos se complica cada vez más dado que el servicio técnico en la actualidad se reduce a remplazar partes que no funcionan por otras nuevas, que serán imposibles de conseguir cuando los equipos se vuelvan obsoletos.

Otra técnica es la *emulación y virtualización* de los equipos antiguos, es decir, disponer de software que permita emular un hardware para poder correr el viejo software. En cualquier caso, esto no exime de tener que conservar los datos digitales, tanto la información que se desea preservar, como el código de los programas emuladores.

La *migración* es otra técnica posible y admite dos variantes. En la primera, se conserva la información digital exactamente como es, lo que obliga a disponer del software necesario para decodificarla o transferirla a un nuevo formato. En la segunda, se realiza una conversión a un formato más actualizado que no produzca cambios en la información esencial. Aquí es importante definir qué es *esencial* y qué no en cada caso. Por ejemplo, se puede conservar un archivo de texto puro sin formato, se puede conservar también el formato incluyendo la diagramación, o se puede conservar información estructural (por ejemplo la categorización en capítulos y secciones, etc.). A largo plazo podría parecer que determinada información accesoria se vuelve irrelevante, pero si lo que se desea hacer es un estudio de carácter histórico esa información podría ser esencial.

Otra técnica consiste en el uso de medios permanentes, que tengan una longevidad estimada de décadas o incluso siglos. Es muy difícil conseguir en la actualidad tales medios en el mercado destinado al consumo, precisamente porque de eso se trata el consumo, de una vida efímera de los productos tecnológicos que garanticen su recambio periódico.

Por último, podemos mencionar la redundancia y la dispersión geográfica de la información, como medios de hacer la preservación menos dependiente de la integridad de una copia en particular. Consiste en guardar varias copias, de ser posible en lugares geográficos diferentes. En cierta medida esta técnica tiene como antecedente la publicación de libros, especialmente los de grandes tiradas, ya que es mucho más difícil la destrucción masiva de todos los ejemplares.

Como se puede apreciar, estas técnicas no son excluyentes, y de hecho algunas se basan en las otras. Podemos ver que en la medida que la tecnología no desarrolle una solución satisfactoria y a un costo razonable para este problema,<sup>74</sup> la migración frecuente, sumada al uso de los mejores medios de almacenamiento y a la distribución de la información parecen ser una solución al menos provisoria. Se pueden aprovechar los servicios denominados de *almacenamiento en la nube*, es decir en servidores indeterminados de Internet. La seguridad de tales sistemas no está completamente garantizada

<sup>74</sup> Esta problemática parece circunscribirse a los círculos académicos, ya que en las sociedades industriales se desalienta el apego a la historia, a sus objetos y a sus enseñanzas, de modo que son pocos los individuos que sienten la necesidad de esa preservación.

tanto en lo que se refiere a la integridad del hardware y los medios de almacenamiento (hoy por hoy, discos rígidos o unidades de estado sólido) como a la inviolabilidad y confidencialidad de la información, por lo que es necesario mantener copias locales y encriptar los archivos.

Aunque no es factible aquí dar una especificación completa, un intento serio de preservar las señales por décadas requiere establecer un protocolo en donde se estipule, como mínimo:

- 1) Planificación a mediano plazo de las acciones a tomar
- 2) Uno o más responsables
- 3) Selección de repositorios
- 4) Previsión de inversión económica
- 5) Documentación exhaustiva de todo lo que se realice, incluyendo descripción y especificación de equipos y tecnologías aplicadas, normas relativas a los formatos a utilizar, uso en lo posible de formatos lo más sencillos que se pueda, organización y administración de la redundancia y distribución geográfica, verificación periódica de la integridad de los datos, reescritura de los archivos para minimizar el efecto del envejecimiento y debilitamiento de marcas magnéticas, electrostáticas u ópticas en los dispositivos de almacenamiento, descripción detallada de los sonidos, incluyendo ubicación geográfica y temporal del registro, motivos del mismo, referencia a documentos en los que se hayan estudiado o analizado.

### **3.20 Conclusión**

Hemos cubierto las técnicas de registro digital, incluyendo la digitalización, los medios en los que se almacena la información, los sistemas de archivo más frecuentemente utilizados y algunos de los formatos binarios detallados de los mismos. Esto último permite al usuario la *interoperabilidad*, es decir la posibilidad de leer, crear o modificar un archivo de sonido (o de cualquier otro tipo) mediante un software diferente del habitualmente utilizado para crearlo o manipularlo. En particular, permite acceder a algunas características como las marcas, lo cual podría servir para automatizar ciertas funciones de análisis de la señal sonora, por ejemplo incorporando señales de calibración, o asignando ubicaciones en un sistema de posicionamiento geográfico global (GPS), o asociando otra información como videos, imágenes o textos descriptivos.

## Capítulo 4

# EDICIÓN DE AUDIO DIGITAL

### 4.1 Introducción

Al adquirir la señal del transductor por medio de un grabador digital o de una computadora portátil dotada de una placa de sonido, se obtiene en general un archivo .wav grabado en algún medio, fijo o removible, típicamente el disco duro de una notebook o una tarjeta SDHC de memoria flash.

Dependiendo del protocolo de medición que se haya seguido podríamos tener distintos archivos para los diferentes eventos de interés, o bien un único archivo de mayor extensión que los contenga a todos. Dichos eventos podrían corresponder a pequeños intervalos que contienen el ruido de determinadas fuentes predominantes o de interés, intervalos de una duración predeterminada (según algún protocolo) con independencia de las fuentes sonoras que intervienen, señales de calibración, comentarios orales efectuados por el operador o incluso señales auxiliares, por ejemplo señales de sincronismo (a la manera de la *claqueta*<sup>75</sup> usada en las tomas de cine y video) o señales provenientes de contadores de vehículos, dispositivos GPS, etc.

En cualquier caso será necesario someter los archivos de sonido a diferentes procesos, tales como cálculos de niveles sonoros, determinación de niveles estadísticos y análisis de espectro, entre otros. Para ello es conveniente realizar una labor previa de edición del material, algunos de cuyos objetivos son:

- 1) Aplicar constantes de calibración que den sentido a los valores numéricos abstractos con que se representan digitalmente las muestras.
- 2) Aislar o circunscribir fragmentos de señal que contengan rasgos característicos, por ejemplo, determinado tipo de fuente.
- 3) Demarcar e identificar eventos con miras a su clasificación o categorización.
- 4) Eliminar porciones irrelevantes
- 5) Transcribir comentarios formulados oralmente o por escrito mediante marcas con anclaje temporal.
- 6) Sincronizar dos o más registros simultáneos.
- 7) Asociar marcas específicas con la hora estándar u oficial.
- 8) Aplicar filtros específicos para eliminar componentes indeseadas
- 9) Cambiar la tasa de muestreo o la resolución en bits.
- 10) Mezclar o combinar diversas señales.

Éstas y muchas otras operaciones pueden realizarse por medio de editores de audio digital. Existen muchos editores, que en su mayoría están orientados a su utilización en producción musical para la industria discográfica. Cabe mencionar Cool Edit, Goldwave, Adobe Audition, Sound Forge, Cubase, Nuendo, Pro Tools, Free Audio Editor, Wavosaur, Wave Editor, Wave Pad, Power Sound Editor, Music Editor Free, Traver-

---

<sup>75</sup> La *claqueta* es una pequeña pizarra con el borde superior articulado que permite identificar la escena y al mismo tiempo producir un sincronismo entre la imagen y el chasquido que produce al golpear las dos partes articuladas.

so DAW, Ardour, Sound Engine, Expstudio Audio Editor, Wavesurfer, entre muchos otros (lista vigente en 2013). Algunos de ellos son gratuitos, otros admiten un periodo de prueba sin cargo (Tan, s/f), y otros tienen versiones de demostración (*demo*) que no son completamente funcionales.

## 4.2 Audacity

Un editor que se destaca frente a todos ellos es Audacity, por ser software libre de código abierto (*open source*),<sup>76</sup> multiplataforma,<sup>77</sup> bajo licencia GPL<sup>78</sup> y en continua expansión a través de un número creciente de colaboradores a lo largo de todo el mundo. Además de una lograda interfaz gráfica para el usuario cuenta con una potente funcionalidad que le permite competir con otros softwares comerciales. Así, por ejemplo, admite una cantidad ilimitada de pistas de audio y de etiquetas e incluye multitud de efectos, filtros, generadores y herramientas de análisis, así como la posibilidad de ampliación mediante el agregado de *plugins* que pueden implementarse en C o en un lenguaje interpretado llamado Nyquist,<sup>79</sup> directamente por el usuario.

Es siempre arriesgado abarcar en un libro demasiados detalles de un software que se actualiza con frecuencia, ya que incluso la próxima versión podría experimentar cambios que harían obsoleta buena parte de la explicación. La decisión de incluir, sin embargo, algunas de las características y funciones de la versión vigente al momento de escribir esta reseña (2013) obedece a tres razones. La primera es que varias de las características seleccionadas son bastante comunes a muchos programas y, ciertamente, a las últimas versiones de Audacity. La segunda razón es que aun si hubiera cambios importantes en versiones futuras, todas las versiones anteriores siguen disponibles en el sitio web del proyecto (<http://audacity.sourceforge.net/>) por lo cual el lector podría descargar la aquí descrita (2.0.3) a fin de poner en práctica lo que aquí se comenta. Por último, es interés del autor dar difusión a este software por ser un software libre que, además de haber alcanzado ya un grado de madurez importante, es mucho lo que promete.

NOTA: Lo que sigue no pretende ser un tutorial del software sino una explicación de muchos aspectos a tener en cuenta en cualquier editor para su uso en un sistema de medición. La explicación de algunas funciones de Audacity es sólo un complemento. Para una descripción más concisa de todas las funciones se sugiere consultar el manual accesible desde el menú de ayuda.

Nos referiremos a algunas de las posibilidades que ofrece Audacity en la versión disponible al momento de escribir esta reseña (2.0.3). Tengamos en cuenta, antes de empezar, que si la instalación se realizó en inglés, siempre es posible cambiar de idioma mediante *Edit / Preferences... / Interface*.<sup>80</sup> Una vez allí, el selector de idioma (Language) permite optar entre más de 50 idiomas. Supondremos de ahora en más, que hemos seleccionado el español.

---

<sup>76</sup> Al ser de código abierto es posible obtener los archivos de código fuente en C++, para su estudio, análisis o eventual modificación.

<sup>77</sup> Existen versiones para Linux, Windows, y Macintosh

<sup>78</sup> GPL son las siglas de *General Public License* (Licencia Pública General), un tipo de licencia impulsada por la Free Software Foundation (Fundación para el Software Libre) como forma de favorecer el intercambio solidario y el acceso irrestricto al uso del software y a la posibilidad de estudiarlo y modificarlo, con la sola condición de que toda distribución del trabajo derivado debe efectuarse bajo la misma licencia GPL. Pueden encontrarse mayores detalles en <http://www.fsf.org/es>

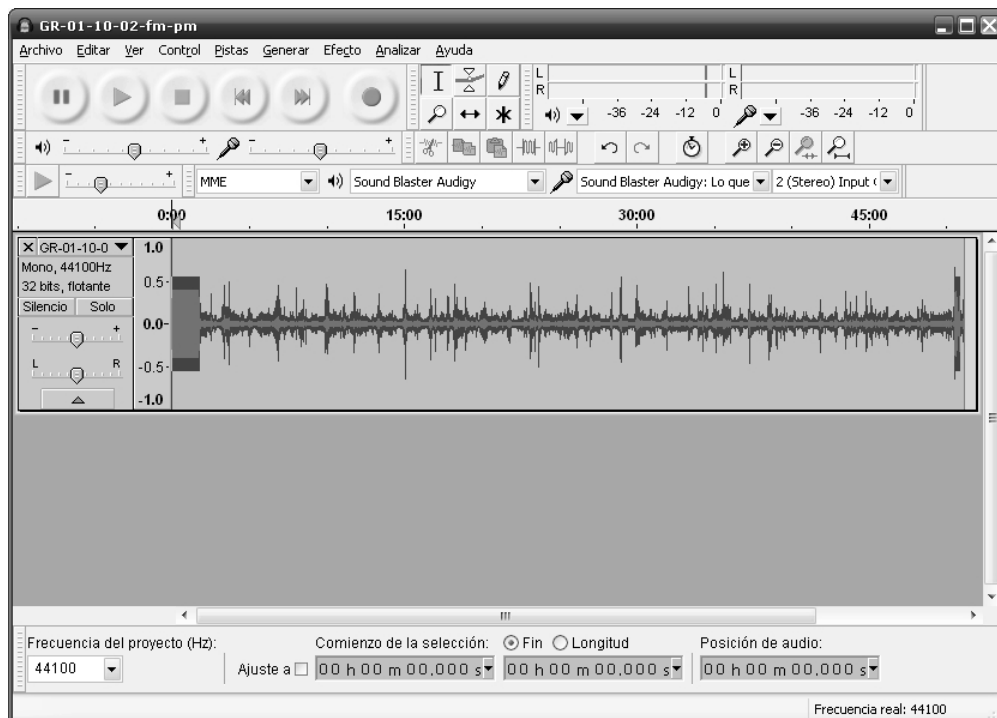
<sup>79</sup> Nyquist fue originalmente diseñado para generar y administrar instrumentos musicales virtuales, pero también ofrece gran potencial para el análisis y procesamiento de señales.

<sup>80</sup> Utilizaremos la convención de separar con / los ítems seleccionados de cada menú o submenú de aquellos seleccionados en cualquier submenú



### 4.2.1 Abrir un archivo preexistente

Aunque el formato .wav no es el formato nativo de Audacity, se pueden *importar* archivos de este tipo ya sea abriéndolos directamente mediante *Archivo / Abrir...* o mediante *Archivo / Importar / Audio...* Una tercera forma es arrastrar el archivo desde un explorador a la interfaz de Audacity. En la figura 65 se ve un ejemplo de una forma de onda así importada. Según se observa, la señal se representa normalizada entre  $\pm 1$ . Cuando, como en este caso, se muestra una gran cantidad de audio, el programa realmente representa el *sumario* del archivo, y no la propia onda. Como se indicó en 3.16.2, el sumario es una versión resumida del archivo que contiene el mínimo, el máximo y el valor eficaz cada 256 muestras y cada 65536 muestras. Se representa la versión de 256 muestras cuando el zoom es mediano y la de 65536 cuando es muy bajo (se representan tiempos muy largos). En el oscilograma, los picos negativos y positivos se muestran en azul oscuro y, en tono más claro,  $\pm$  el valor eficaz (gris oscuro y claro en la figura).



**Figura 65.** Ventana principal de Audacity (el original es en colores). Se observan áreas de menús, de controles de transporte que permiten grabar, reproducir, pausar o detener, de herramientas interactivas, de pista (donde se muestra el oscilograma) y de selección.

En cualquiera de los dos casos el proyecto adopta la tasa de muestreo del archivo original, pero el formato numérico pasa a ser de punto flotante de 32 bit, según se muestra en el panel de control a la izquierda de la pista. Es posible conservar la resolución original, pero para ello se debe modificar el valor por defecto mediante el editor de preferencias (*Editar / Preferencias... / Calidad*).

Esta decisión puede ser conveniente si los procesos de edición previstos no introducirán cambios en la forma de onda; por ejemplo: separar una porción de la señal para su análisis posterior, agregar marcas, concatenar dos trozos.

Cuando la señal vaya a ser sometida a diversos cálculos, como por ejemplo multiplicarla por una constante o filtrarla, conviene trabajar en 32 bit, ya que los errores aritméticos por truncamiento se darán en un dígito mucho menos significativo. Para ilustrar el concepto, supongamos que multiplicamos el número 3,2456 cuatro veces por 0,65321 (una operación típica en los filtros digitales). El valor exacto es

$$3,2456 \times 0,65321^4 = 0,385\,974\,948\,890\,394\,580\,168\,251\,860\,56$$

Si cada resultado intermedio se trunca a 10 cifras significativas después de cada operación, al final tendremos

$$3,2456 \times 0,65321^4 \cong 0,385\,974\,947\,9$$

donde se han subrayado las cifras incorrectas. Si ahora cada resultado intermedio se trunca a 5 cifras significativas después de cada operación, resultará

$$3,2456 \times 0,65321^4 \cong 0,385\,9\bar{5}$$

Comparando con el cálculo exacto hecho con precisión ilimitada, vemos que en el primer caso tenemos 8 cifras exactas, mientras que en el segundo tenemos 4. Si truncamos la primera aproximación a 5 cifras, serán todas correctas, mientras que en la segunda aproximación sólo 4 cifras son exactas.

Esta situación, ilustrada en código decimal, también se verifica en código binario, de allí que sea preferible trabajar con 32 bit, ya que cuando al final truncemos a sólo 16 bit para obtener una representación estándar el resultado será más exacto.

Ahora bien, según lo analizado en la sección 3.7, antes de realizar el truncado final, que conceptualmente es una recuantización con menos bits (ya que pasaremos de 32 bit a 16 bit) será preciso agregar una señal de *dither* a fin de descorrelacionar.

Audacity, por defecto agrega *dithering* a la señal antes de recuantizar, es por ello que no conviene convertir a 32 bit una señal de 16 bit si no se va a modificar, ya que antes de recuantizar se le agregará un dithering que implica una modificación innecesaria de las muestras originales. El agregado de dithering se puede desactivar en el menú *Edición / Preferencias / Calidad / Fusión* (obsérvese que se ha traducido *dither* como *fusión*).

#### 4.2.2 Grabar sonidos

Otra forma de cargar una señal para luego trabajar con ella es grabarla en forma directa. Para ello, la señal del transductor, que en nuestro caso será la salida calibrada de audio de un sonómetro, debe conectarse a la entrada de línea de una placa de sonido, ya sea interna o, preferentemente, externa (pues son de mejor calidad).

Antes de proceder a grabar es necesario cerciorarse de que la señal que llega a la placa de sonido no es ni demasiado pequeña como para que el ruido de digitalización la deteriore significativamente, ni demasiado grande como para saturar.

El ajuste de nivel se realiza mediante el deslizador con un ícono de micrófono mientras se monitorea visualmente el vúmetro con el mismo ícono. Para hacer visible la indicación del vúmetro se abre un menú contextual marcando con el botón secundario del mouse sobre el vúmetro y activando la monitorización. Si la señal será monofónica,

se deberá seleccionar en la casilla correspondiente de la barra de herramientas de dispositivos.<sup>81</sup> Una vez ajustado el nivel, se oprime el botón rojo de grabación. El software genera automáticamente una nueva pista

#### 4.2.3 Generar señales

Por último, mediante el menú Generar se pueden obtener señales tales como tonos puros, ondas triangulares y cuadradas y barridos tonales (*chirp*), que resultan útiles como señales de prueba. La variedad de funciones disponibles no es aún muy grande, pero siempre es posible crearlas utilizando el lenguaje Nyquist. Como se verá más adelante, es especialmente interesante que se pueda reproducir una pista mientras simultáneamente se graba otra pista, lo cual resulta útil para ensayos acústicos.

#### 4.2.4 Agregar pistas

Una vez que se dispone de una pista es posible agregar pistas nuevas por varios mecanismos. El primero es importar otro archivo .wav.<sup>82</sup> Automáticamente se generará una pista nueva que contendrá la nueva señal. Cabe aclarar aquí que Audacity permite importar archivos con diferentes tasas de muestreo en un mismo proyecto. Sin embargo, cada proyecto tiene una tasa de muestreo global que se define por defecto en el editor de preferencias (*Editar / Preferencias... / Calidad*). Esta corresponde a la primera pista del proyecto, pero se puede modificar en el panel inferior de la ventana. Esta posibilidad podría ser muy útil en el contexto de la composición de música electroacústica, o de la simulación de paisajes sonoros, al permitir combinar fuentes de procedencia heterogénea. Pero no es recomendable en un contexto metrológico ya que en general serán señales de calidades diferentes.

#### 4.2.5 Guardar un proyecto

Una vez que se dispone del material sonoro necesario, se puede proceder a guardar el proyecto, generándose un archivo .aup y un directorio de igual nombre terminado en \_data que contiene, en un árbol de subdirectorios, uno o más archivos .au o .auf según la modalidad de guardado que se elija. Existen dos posibilidades: guardar una copia real del audio, en cuyo caso se generan archivos .au según el formato descrito en 3.16.2, o no guardarla, en cuyo caso se crean archivos .auf con el formato indicado en 3.16.3. En este último caso se crea, además, una *dependencia*, que no es otra cosa que un vínculo al archivo original que aparece en el archivo de proyecto. El programa advierte al usuario antes guardar. Si bien el archivo .wav original no será modificado por Audacity, sí podría modificarse mediante algún programa externo, lo cual podría causar incoherencias que podrían afectar la integridad del proyecto. Las dependencias de un proyecto pueden consultarse mediante *Archivo / Comprobar dependencias...*

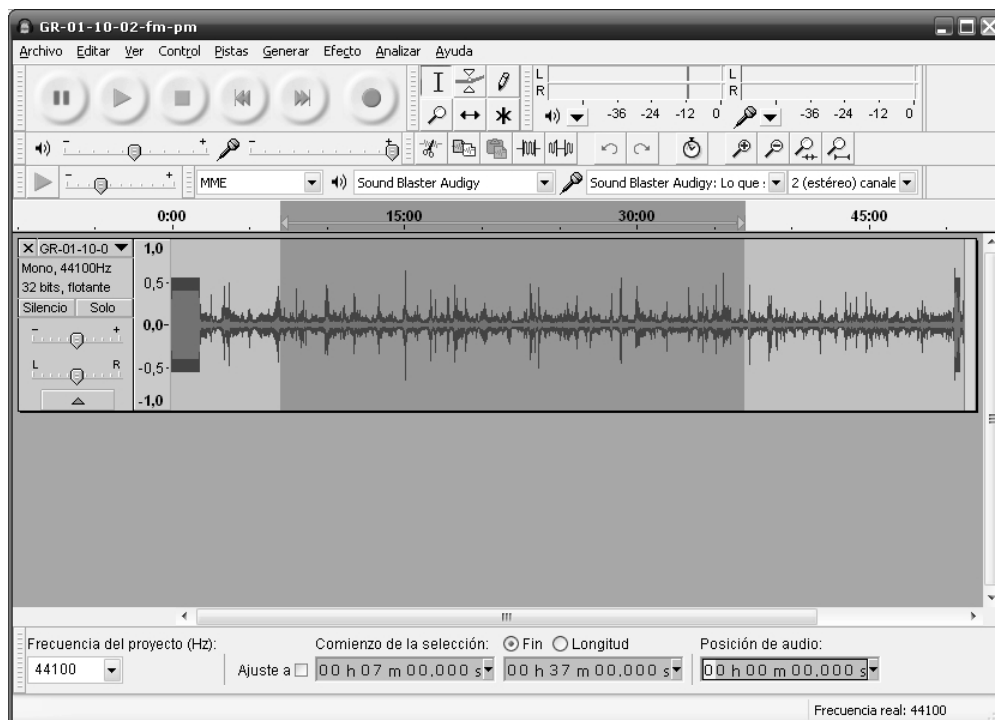
De todas formas hay un límite para el tipo de edición que se puede realizar sin que el programa copie los archivos de audio. Si la edición implica el cambio efectivo de la onda, por ejemplo, reemplazar alguna parte, o filtrarla, al guardar el proyecto se generarán los archivos .au correspondientes (y desaparecerá la dependencia), a menos que se trate de un cambio en la envolvente, la ganancia o el paneo, ya que éstos quedan registrados en el archivo de proyecto .aup.

<sup>81</sup> Al pasar el puntero del mouse sobre cualquier elemento activo de la interfaz, Audacity muestra ayudas breves contextuales. Esto permite ir descubriendo funciones o recordar dónde estaban.

<sup>82</sup> Tener en cuenta que en este caso sólo se puede utilizar la función Importar, ya que si se intenta Abrir el archivo a agregar se abrirá en otra ventana independiente.

#### 4.2.6 Selección

Una de las operaciones preliminares para cualquier evento de edición es seleccionar una porción de la señal. En Audacity ello puede hacerse de dos maneras. La primera es identificar visualmente la porción deseada, posicionar el puntero del mouse al principio y arrastrar presionando el botón principal hasta el final. El resultado se muestra en la figura 66. La segunda es identificar los tiempos inicial y final e ingresarlos en las casillas correspondientes de la barra de selección que se encuentra debajo de las pistas.<sup>83</sup> También se puede optar por definir el tiempo inicial o la longitud de la selección, para lo cual basta marcar el botón correspondiente. Incluso si se utilizó el primer método, después se pueden ajustar los límites de la selección entrando en las casillas correspondientes. También pueden hacerse ajustes con el mouse. Basta acercar el mouse a uno u otro extremo hasta que aparezca un puntero en forma de mano, y luego arrastrar.



**Figura 66.** Selección de una porción de una pista de sonido.

Una selección puede guardarse momentáneamente mediante *Editar / Guardar selección* y luego recuperarse mediante *Editar / Restaurar Selección*. En caso de haber más de una pista, se puede copiar la selección hecha en una de ellas oprimiendo Shift y haciendo clic con el mouse en el panel de control a la izquierda de otra pista.

<sup>83</sup> Esta posición podría variar, ya que en Audacity la mayoría de las barras de herramientas pueden redistribuirse y admiten la funcionalidad de acoplamiento (*docking toolbars*). Es decir que al tomarlas con el mouse por su asa a la izquierda se pueden liberar y desplazar hasta una nueva área de acoplamiento. Al soltarlas se acoplan permanentemente quedando fijas hasta que se las vuelva a desplazar. Las áreas de acoplamiento se identifican porque aparece superpuesta a la barra un triángulo azul.

Una función interesante en el contexto de la selección es la detección de cruces por cero. Una vez seleccionada una porción o región de la señal, mediante *Editar / Encontrar cruces en cero* (o con la tecla abreviada Z) los límites de la selección se mueven hasta cruces por 0 cercanos. Esto es útil para evitar saltos que podrían alterar, por ejemplo, el espectro.

#### 4.2.7 Marcas y etiquetas

Una herramienta sumamente útil son las pistas de etiquetas. Se pueden crear tantas pistas de etiquetas como se desee. Cada pista contendrá etiquetas asociadas a intervalos o a instantes. Se crean mediante *Pistas / Añadir nueva / Pista de etiqueta*. A cada pista se le puede cambiar el nombre por defecto (“Pista de etiqueta”) por un nombre personalizado. Una vez creada, se pueden agregar etiquetas de dos formas. En la primera se selecciona una región a la que se desea asociarle una etiqueta y, mediante *Pistas / Añadir etiqueta en la selección*, se genera la etiqueta. Inicialmente está vacía, pero se puede escribir directamente en la casilla el texto deseado. La casilla se ira ensanchando conforme se agregue más texto.

La segunda forma de crearla es mediante *Pistas / Editar etiquetas...* que nos ofrece una tabla con el nombre de la pista de etiqueta, el contenido de la etiqueta, el instante inicial y el instante final. Allí no sólo se pueden editar los campos correspondientes a etiquetas previas sino que se pueden insertar nuevas etiquetas. Al hacerlo la etiqueta estará ubicada en 0, por lo que habría que editar los tiempos para ubicarla en la posición o en el intervalo deseado. La figura 67 muestra dos etiquetas agregadas en el archivo de las figuras anteriores

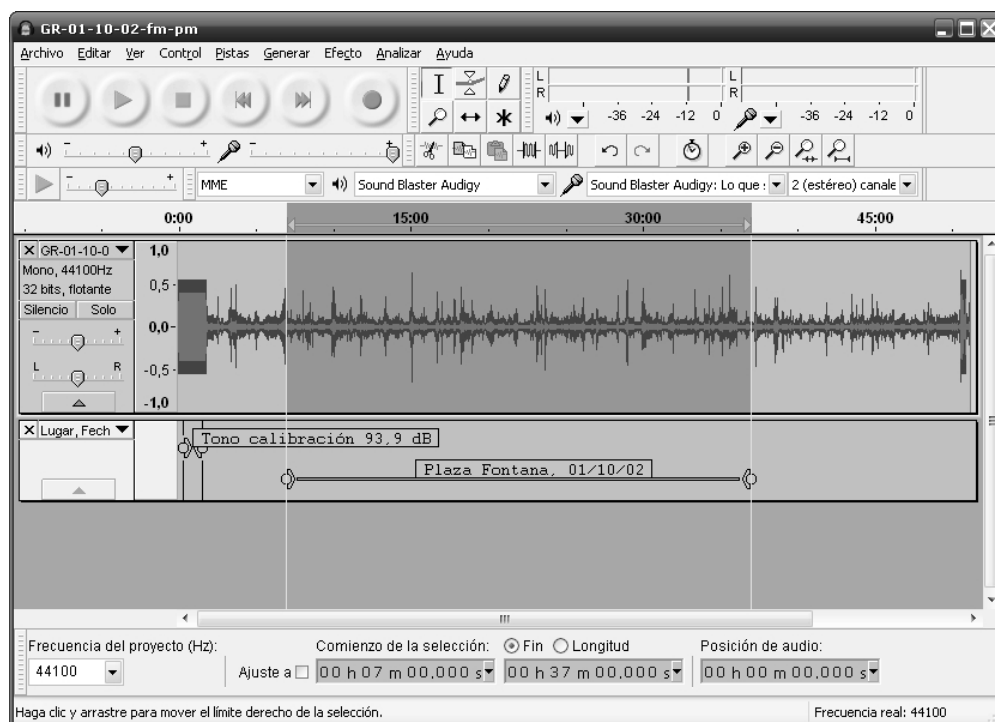


Figura 67. Selección de una porción de una pista de sonido.

Las etiquetas pueden utilizarse para determinar un protocolo de análisis que después podría automatizarse mediante una lectura del archivo XML del proyecto (.aup) y detección automática de palabras clave. Así, podría asociarse una pista de etiquetas distinta a cada categoría de información que nos interese registrar. Por ejemplo, podría haber una etiqueta para fecha y hora, una para coordenadas GPS, una para tonos de calibración, una para intervalos de medición, otra para eventos de sincronismo, otra para descripción de eventos estructurados (por ejemplo, conteo automático de vehículos), otra para descripción de eventos no estructurados (por ejemplo sonidos sin una clasificación previa, como un ladrido, un grito, un objeto que cae).

Dado que la estructura de un archivo XML está diseñada precisamente para poder extraer y clasificar información estructurada de una forma sencilla, puede levantarse esa información directamente por software, detectando la presencia de las palabras clave que definen una categoría de información, y luego dentro de ella, detectando los elementos y, si corresponde, sus atributos o parámetros asociados.

#### 4.2.8 Selección en presencia de etiquetas

Las etiquetas también facilitan el repetir selecciones, ya que al arrastrar el puntero del mouse con el botón principal oprimido hacia el borde de una selección esta atrae y fija el borde de la selección exactamente al borde (a veces llamado *ajuste magnético*). Al tomar y arrastrar el otro extremo de la selección hacia el otro borde vuelve a suceder lo mismo, con lo cual queda seleccionada la región correspondiente a la etiqueta. Para extender la selección a otras pistas bastará oprimir Shift y marcar con el mouse dentro el panel de control que se encuentra a la izquierda de la pista a seleccionar.

#### 4.2.9 Tono de calibración

Cuando se utiliza un transductor (micrófono) como generador de señal, el proceso que conduce finalmente a un archivo de sonido está mediado por una serie de componentes que van modificando la señal. El primero es el propio transductor, cuya relación entre la presión sonora (entrada) y la tensión eléctrica (salida) supondremos lineal a través de una constante denominada *sensibilidad*:

$$v(t) = S_m p(t). \quad (4.1)$$

Luego tenemos la ganancia del acondicionador de señal del sonómetro,  $G_a$ , la del preamplificador del dispositivo digitalizador,  $G_d$ , la constante del conversor analógico digital,  $K_{A/D}$ , y la constante de normalización  $K_n$  utilizada por el software. En consecuencia, el valor numérico adimensional  $x(n)$  con el que el software representa el valor de presión sonora en el instante  $t_n$  es

$$x(n) = K_n K_{A/D} G_d G_a S_m p(t_n) = K p(t_n). \quad (4.2)$$

Es difícil en la práctica llevar un control de todas las constantes que forman  $K$ . Particularmente, las constantes  $G_d$  y  $K_{A/D}$  del conversor no suelen estar especificadas ni son susceptibles de ajuste alguno, a diferencia de  $G_a$  que es ajustable mediante un preset o de  $K_n$ , que está bajo el control del programador.

Algunas de esas constantes pueden estar sujetas, incluso, a derivas por cambios en las condiciones ambientales, como temperatura, presión atmosférica y humedad relativa (mediano plazo) o por envejecimiento (largo plazo). No obstante, las variaciones son pequeñas durante tiempos de medición típicos de algunas decenas de minutos.

La manera de obtener  $K$  en forma indirecta es grabando una señal de calibración de la cual se conozca el nivel de presión sonora con una incertidumbre pequeña. Esta señal podría ser, en principio, cualquier señal de la cual podamos obtener una lectura en el sonómetro que pueda asociarse al valor correspondiente obtenido de la grabación. En la práctica, sin embargo se utiliza un tono de calibración generado por un calibrador acústico. Éste es un dispositivo que produce, en un acoplador acústico en el que se introduce el micrófono, un tono puro de frecuencia y nivel de presión sonora especificados.

Existen aquí dos posibilidades. La primera es tomar algún valor instantáneo característico, como por ejemplo el pico de la señal de calibración registrada. La segunda, tomar el valor eficaz en un periodo de algunos segundos de duración. Esta última posibilidad es preferible, ya que se reduce la influencia del ruido acústico<sup>84</sup> y eléctrico.

La forma de calcularlo es aproximando numéricamente la integral de la ecuación (1.1) mediante

$$X_{\text{ef}} = \sqrt{\frac{1}{N_2 - N_1 + 1} \sum_{n=N_1}^{N_2} x^2(n)}. \quad (4.3)$$

donde  $N_1$  es el índice de la muestra inicial y  $N_2$  el de la muestra final del tono de calibración grabado. Llamaremos  $X_{\text{cal}}$  a dicho valor.

Suponiendo que el nivel de presión sonora sea  $L_{\text{cal}}$ , que corresponde a una presión sonora

$$P_{\text{cal}} = P_{\text{ref}} 10^{L_{\text{cal}}/20} \quad (4.4)$$

donde  $P_{\text{ref}} = 20 \mu\text{Pa}$ , podemos obtener la constante  $K$  como

$$K = \frac{X_{\text{cal}}}{P_{\text{cal}}} = \frac{X_{\text{cal}}}{P_{\text{ref}} 10^{L_{\text{cal}}/20}}. \quad (4.5)$$

Podemos aplicar ahora esa constante, junto con la ecuación (4.2), para determinar la presión sonora correspondiente al instante  $t_n$ :

$$p(t_n) = \frac{1}{K} x(n). \quad (4.6)$$

Naturalmente, esta ecuación también se cumple para valores eficaces, como se deduce fácilmente de (4.3):

$$P_{\text{ef}} = \frac{1}{K} X_{\text{ef}}. \quad (4.7)$$

Si, como es habitual, preferimos trabajar con niveles de presión sonora en lugar de presiones sonoras instantáneas, entonces

<sup>84</sup> El acoplador acústico tiene cierto grado de aislación acústica, pero los sonidos ambientales, aun atenuados, contaminan valores característicos como el de pico.

$$L_p = 20 \log \frac{X_{ef} / K}{P_{ref}} = 20 \log X_{ef} - 20 \log X_{cal} + L_{cal} \quad (4.8)$$

Resulta así que podemos convertir un valor eficaz determinado numéricamente a partir de la señal digital en el correspondiente nivel de presión sonora. Para ello solo necesitamos contar con una porción de la señal de calibración y conocer el nivel de presión sonora que tenía.

Podemos implementar esto en forma directa en Audacity con el siguiente procedimiento:

- 1) Abrir la ventana *Analizar / Contraste*
- 2) Identificar el tono de calibración, cuyo oscilograma se caracteriza por tener una amplitud marcadamente constante;
- 3) Seleccionar una porción de varios segundos y aplicar la función de detección de cruces por cero. Esto reducirá el error en el cálculo del valor eficaz, ya que el error se anula cuando el tiempo coincide con una cantidad entera de semiciclos (ver figura 3);
- 4) En la ventana de contraste, marcar el botón inferior de los dos rotulados “Selección de medida”

El resultado es un nivel en dB que equivale a  $20 \log X_{ef}$ . Es un nivel referido a 1, es decir, la máxima señal normalizada, por lo tanto siempre es  $\leq 0$ . En el ejemplo de la figura 65 el resultado es  $-8,2$  dB.

Es muy importante no caer en el error de creer que éste es directamente el nivel de presión sonora. En realidad, el nivel de presión sonora del calibrador era  $93,9$  dB.

Si ahora repetimos las operaciones 2) a 4) con la selección marcada por la otra etiqueta, marcando el botón superior “Selección de medida”, obtenemos  $-30,2$  dB. Aplicando la ecuación (4.8), resulta:

$$L_p = -30,2 \text{ dB} - (-8,2 \text{ dB}) + 93,9 \text{ dB} = 71,9 \text{ dB},$$

de modo que el nivel de presión sonora equivalente en el periodo indicado es de  $71,9$  dB. El casillero rotulado “Diferencia” calcula la diferencia entre ambos valores por lo cual también puede utilizarse, si se suma el nivel del tono de calibración.

Dos observaciones finales. Primero, para tiempos largos (varias horas), puede convenir repetir la calibración con alguna frecuencia (por ejemplo, cada 30 min), aunque las variaciones suelen ser muy pequeñas. Esto es para tener en cuenta derivas de la calibración del instrumento, sobre todo si ha habido variaciones graduales de temperatura o humedad relativa.

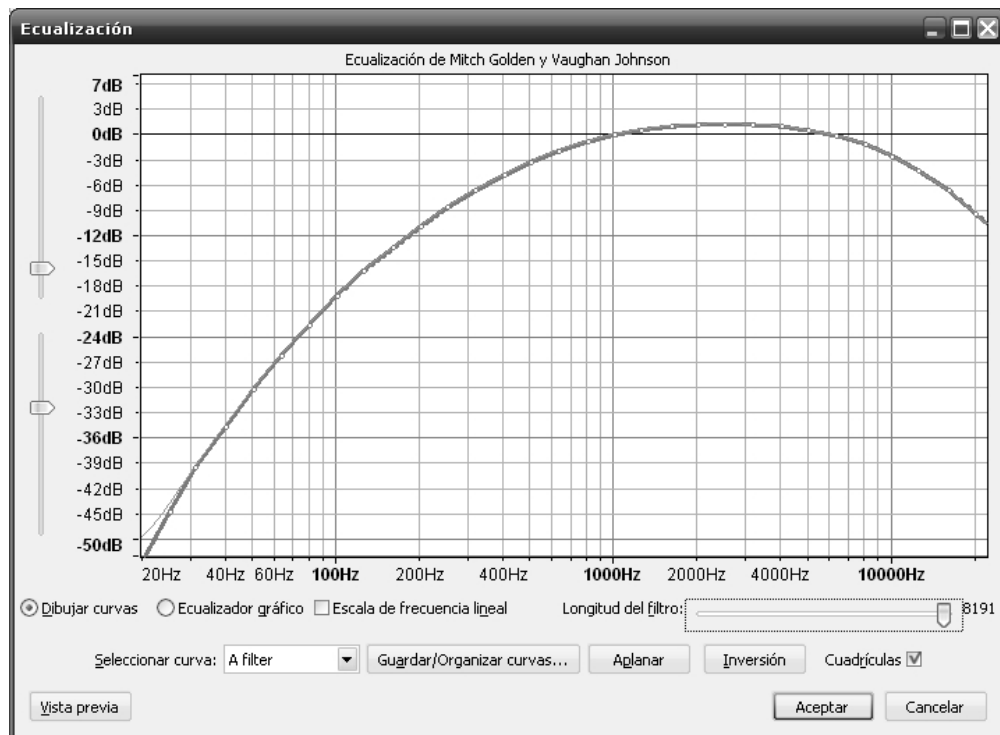
Segundo, es conveniente, inmediatamente después de grabar el tono de calibración, grabar oralmente el valores leídos como documentación sonora de lo realizado.

#### 4.2.10 Filtros FFT

Aunque nos referiremos en detalle a los filtros FFT en el capítulo sobre análisis de espectro, veremos aquí las posibilidades que ofrece Audacity al respecto. Conceptualmente, los filtros FFT posibilitan implementar especificaciones de respuesta en frecuencia prácticamente arbitrarias, con lo cual permiten la realización de filtros específicos como los filtros A y C o los de banda de octava o tercio de octava, entre otros. Una vez



seleccionada una porción de señal. La forma de invocar esta función es mediante *Efecto / Ecualización...*. Se abre una ventana como la indicada en la figura 68.



**Figura 68.** Ventana de selección y aplicación de filtros FFT.

En ella se puede dibujar una curva tomando y arrastrando puntos con el mouse o bien seleccionar una curva filtrante mediante *Seleccionar curva*. La curva gruesa (en azul en el original) con puntos vacíos representa la especificación deseada del filtro. La curva delgada (en verde en el original) representa la respuesta realmente obtenida. En baja frecuencia esta respuesta difiere un poco de la deseada. Se aproxima tanto más cuanto mayor sea la longitud del filtro. Los deslizantes laterales permiten mover los límites superior e inferior.

La curva mostrada en la figura corresponde a un filtro A (ver figura 5 y tabla 1.1), que no está originalmente entre los provistos. Mediante *Guardar/Organizar curvas...* se accede a una ventana de diálogo donde se pueden importar curvas que se hayan guardado previamente en un archivo con extensión .xml de acuerdo al siguiente formato:

```
<equalizationeffect>
  <curve name="nombre">
    <point f="f1" d="d1" />
    <point f="f2" d="d2" />
    ...
    <point f="fn" d="dn" />
  </curve>
</equalizationeffect>
```

donde *nombre* es el nombre del filtro,  $f_1, \dots, f_n$  representan valores numéricos de frecuencias en Hz y  $d_1, \dots, d_n$  representan los correspondientes valores numéricos de la respuesta expresados en dB. En particular, el filtro A queda definido como:

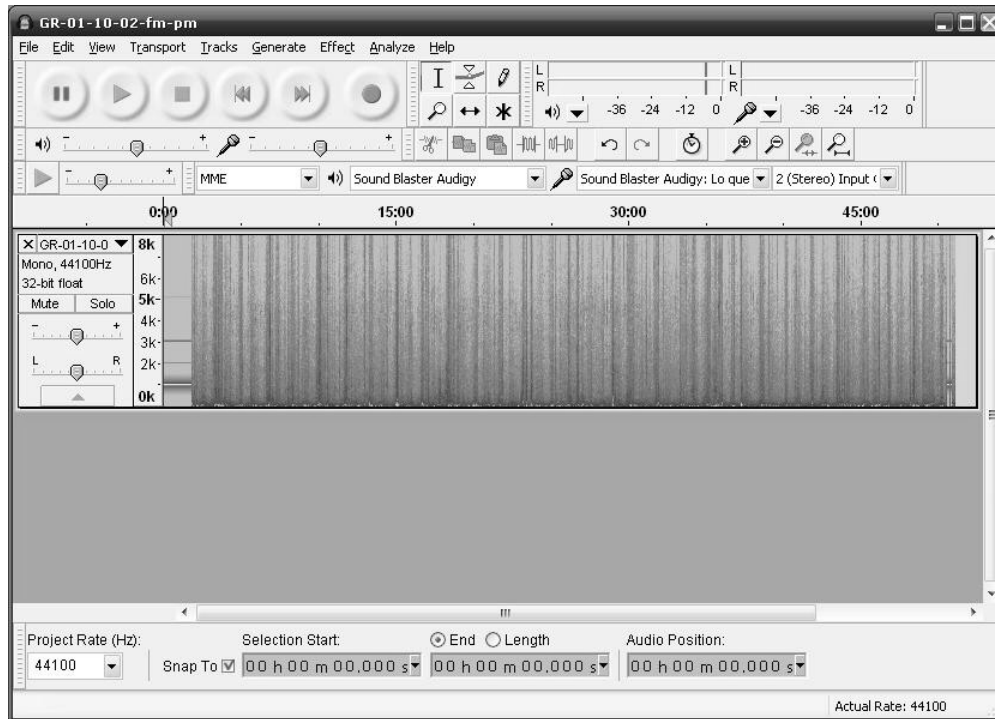
```
<equalizationeffect>
  <curve name="A filter">
    <point f="20.0000000000" d="-50.5000000000"/>
    <point f="25.0000000000" d="-44.7000000000"/>
    <point f="31.0000000000" d="-39.4000000000"/>
    <point f="40.0000000000" d="-34.6000000000"/>
    <point f="50.0000000000" d="-30.2000000000"/>
    <point f="63.0000000000" d="-26.2000000000"/>
    <point f="80.0000000000" d="-22.5000000000"/>
    <point f="100.0000000000" d="-19.1000000000"/>
    <point f="125.0000000000" d="-16.1000000000"/>
    <point f="160.0000000000" d="-13.4000000000"/>
    <point f="200.0000000000" d="-10.9000000000"/>
    <point f="250.0000000000" d="-8.6000000000"/>
    <point f="315.0000000000" d="-6.6000000000"/>
    <point f="400.0000000000" d="-4.8000000000"/>
    <point f="500.0000000000" d="-3.2000000000"/>
    <point f="630.0000000000" d="-1.9000000000"/>
    <point f="800.0000000000" d="-0.8000000000"/>
    <point f="1000.0000000000" d="0.0000000000"/>
    <point f="1250.0000000000" d="0.6000000000"/>
    <point f="1600.0000000000" d="1.0000000000"/>
    <point f="2000.0000000000" d="1.2000000000"/>
    <point f="2500.0000000000" d="1.3000000000"/>
    <point f="3150.0000000000" d="1.2000000000"/>
    <point f="4000.0000000000" d="1.0000000000"/>
    <point f="5000.0000000000" d="0.5000000000"/>
    <point f="6300.0000000000" d="-0.1000000000"/>
    <point f="8000.0000000000" d="-1.1000000000"/>
    <point f="10000.0000000000" d="-2.5000000000"/>
    <point f="12500.0000000000" d="-4.3000000000"/>
    <point f="16000.0000000000" d="-6.6000000000"/>
    <point f="20000.0000000000" d="-9.3000000000"/>
    <point f="25000.0000000000" d="-12.3000000000"/>
    <point f="31500.0000000000" d="-15.7000000000"/>
    <point f="40000.0000000000" d="-19.4000000000"/>
    <point f="48000.0000000000" d="-22.3000000000"/>
    <point f="50000.0000000000" d="-23.0000000000"/>
    <point f="63000.0000000000" d="-26.8000000000"/>
    <point f="80000.0000000000" d="-30.9000000000"/>
    <point f="100000.0000000000" d="-34.7000000000"/>
  </curve>
</equalizationeffect>
```

De igual manera pueden implementarse filtros C, de banda de octava, de tercio de octava, de bandas críticas psicoacústicas, así como muchos otros que puedan resultar de interés para aplicaciones específicas. En un futuro próximo se prevé la posibilidad de agregar filtros variables en el tiempo (Boggino, 2013; Bergero, 2013).

#### 4.2.11 Espectrograma y análisis de espectro

Dos herramientas infaltables en todo software de edición de sonido son el espectrograma y el análisis de espectro. El espectrograma muestra en el espacio correspondiente a la pista un gráfico en colores o tonos de gris en el cual el eje horizontal es el de

tiempo y el vertical el de frecuencia, mientras que el color o tono de gris representa la magnitud en una escala convencional que va desde los colores oscuros o fríos hasta los más cálidos. Internamente es en realidad una escala de 5 colores que se interpolan. En la figura 69 se muestra el espectrograma correspondiente a la figura 65. Para acceder a esta forma de presentación se debe pulsar sobre la flecha en el panel de control a la izquierda de la pista, y seleccionar *Espectro* de la lista que se despliega.



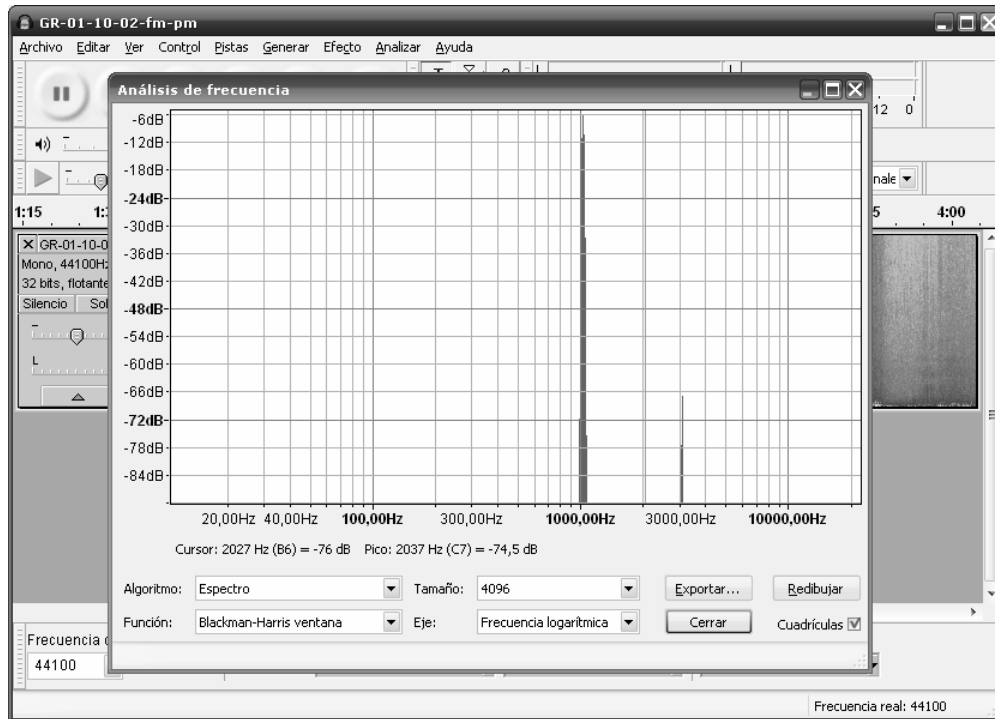
**Figura 69.** Espectrograma de la señal de la figura 65.

Si bien se requiere algo de práctica, el espectrograma es muy útil para detectar visualmente eventos. En la figura se aprecia en seguida la presencia del tono de calibración al principio, donde se ve una línea horizontal intensa centrada en 1 kHz más algunos armónicos más débiles en 2 kHz, 3 kHz y 5 kHz. El resto son eventos de ruido de tránsito que exhiben un espectro amplio que cubre casi todo el espectro, con énfasis en las bajas frecuencias. Notar que la escala por defecto se extiende entre 0 y 8 kHz, pero se puede aumentar o reducir el nivel de zoom vertical haciendo clic sobre la escala correspondiente manteniendo oprimida o no la tecla Shift. También es posible seleccionar una región vertical del espectro arrastrando con el mouse entre dos frecuencias. La segunda frecuencia puede estar incluso fuera de la escala original.

En general, abarcando hasta 5 kHz se logran apreciar los detalles más relevantes, por ejemplo, los armónicos de la mayoría de los sonidos tonales, como voces humanas y animales, motores, vehículos, entre otros. Para poder apreciar los patrones característicos de eventos cortos, como trinos de pájaros, palabra hablada, bocinas, ruidos mecánicos, ruidos impulsivos, etc. es preciso adecuar los niveles de zoom horizontal y vertical.

También es posible aumentar el tamaño vertical de la ventana de la pista, con lo cual se obtendrá mayor nivel de detalle.

La otra herramienta de presentación de características espectrales es el análisis de espectro de una porción seleccionada de la onda. Para ello, después de efectuar la selección, mediante *Analizar / Análisis de espectro...* se abre una ventana como la indicada en la figura 70 que muestra el espectro. En este caso se observa el tono de calibración.



**Figura 70.** Ventana de análisis de espectro de la señal de la figura 65, donde se ha seleccionado parte del tono de calibración.

Una posibilidad interesante que ofrece esta ventana es la de exportar el espectro en la forma de un archivo de texto en el cual se detallan todas las líneas espectrales. El siguiente es un extracto de lo que sucede cerca de 1000 Hz (la frecuencia nominal del tono de calibración):

Frecuencia (Hz)	volumen (dB)
968,994141	-97,818932
979,760742	-65,876839
990,527344	-30,771830
1001,293945	-14,120676
1012,060547	-6,435168
1022,827148	-5,694033
1033,593750	-11,738699
1044,360352	-26,038492
1055,126953	-54,824665
1065,893555	-98,254608

Vemos que la energía se concentra fuertemente entre las bandas de 1001,293945 Hz y 1033,593750 Hz. Más adelante se explicará cómo se interpreta un espectro FFT.

También es posible cambiar el tamaño del espectro, es decir la cantidad de muestras utilizadas para calcularlo. Este valor impacta directamente en la resolución que se obtiene. Cuanto mayor sea la longitud (siempre son potencias de 2) más detallado será el espectro obtenido. En el estado actual del programa no es posible ir más allá de 16384, pero este límite un parámetro que se puede modificar con relativa facilidad editando el código fuente, posibilidad que existe por tratarse de un software libre de código abierto.

#### 4.2.12 Reducción de ruido

Por último, al igual que muchos programas de edición de audio digital, Audacity tiene una rutina de reducción de ruido. Si bien las aplicaciones metrológicas de la reducción de ruido son escasas, ya que toda reducción de ruido también altera en alguna medida la señal útil, haremos una rápida descripción de esta herramienta.

Ante todo es necesario aclarar que el objetivo de la reducción de ruido es eliminar una componente de ruido sistemática agregada ya sea por el canal de transmisión o por el medio de grabación. Es necesario disponer de un intervalo de tiempo en el que sólo se encuentre presente el tipo de ruido a eliminar. En la figura 71 se muestra la ventana que se abre mediante *Efecto / Reducción de ruido...*



**Figura 71.** Ventana de selección de parámetros para el algoritmo de reducción de ruido.

El procedimiento consiste en seleccionar primero un intervalo en el que se haya grabado sólo el ruido de fondo a eliminar y oprimir *Obtener perfil de ruido*. El software realiza un análisis de espectro promedio del fragmento de ruido y lo mantiene en reserva. Luego se selecciona la porción de señal de la que se quiere eliminar el ruido y se eligen los parámetros indicados. Al aceptar se pone en funcionamiento el algoritmo y el ruido se reduce.

El reductor de ruido es esencialmente una compuerta de ruido multibanda. Tanto para la obtención del perfil de ruido como para el análisis de la señal utiliza una FFT de 2048 puntos que brinda un espectro de 1024 bandas.<sup>85</sup> El perfil de ruido es básicamente un percentil alto (un valor cercano al máximo) del espectro durante el tiempo que dura el fragmento de ruido.

Para reducir el ruido se toman ventanas de 2048 puntos de señal, se calcula el espectro de 1024 bandas y se determina, en forma separada para cada banda, si la señal supera un umbral basado en el perfil del ruido de fondo y en el parámetro *Sensibilidad*, que indica a cuántos dB por encima del perfil correspondiente a cada banda está el umbral. Si la señal supera al umbral, no se hace nada. Si, en cambio, es menor que el umbral, procede a reducir el valor de esa banda en la cantidad especificada por el parámetro *Reducción de ruido* de la ventana de diálogo del reductor.

Para reducir la aparición de *artefactos*<sup>86</sup> se toman dos precauciones. La primera es no eliminar o reducir una banda suelta. Esto es porque salvo en situaciones muy particulares un tono está representado por más de una banda,<sup>87</sup> por lo que si se elimina una sola de ellas tenderá a producirse una notoria modulación en amplitud. El “ancho de banda” para determinar la similitud de la reducción de ruido en bandas adyacentes lo fija el parámetro *Suavizado de frecuencia*.

La segunda precaución es evitar un cambio brusco en el tiempo, es decir, que de repente desaparezca una componente de la señal, ya que ello produciría un transitorio perceptible. El parámetro que regula esto es *Tiempo de ataque o decaimiento*.

Una vez reducidas las bandas que corresponda, se obtiene la FFT inversa, recuperando una versión de la señal sin el ruido. Dado que las ventanas son muy cortas, es preciso empalmarlas. Ello se hace por solapamiento mediante un *crossfade* entre ventanas adyacentes (ver filtros FFT) (Mazzoni, s/f; Crook, 2011).

<sup>85</sup> Veremos más adelante que como la FFT es un espectro complejo, para obtener cada armónico real se requiere superponer dos líneas conjugadas.

<sup>86</sup> Se denomina *artefacto* en el contexto de las señales de audio a cualquier perturbación audible causada por algún procesamiento. Ejemplo: los clics y la aparición de tonos parásitos.

<sup>87</sup> Un tono es representable con una única línea espectral si su amplitud es constante y su frecuencia coincide exactamente con la frecuencia de una línea espectral, es decir  $kF_s/N$  (donde  $k$  es un número entero,  $N$  la longitud de la ventana y  $F_s$  y que su amplitud sea constante)

## Capítulo 5

# TRANSDUCTORES

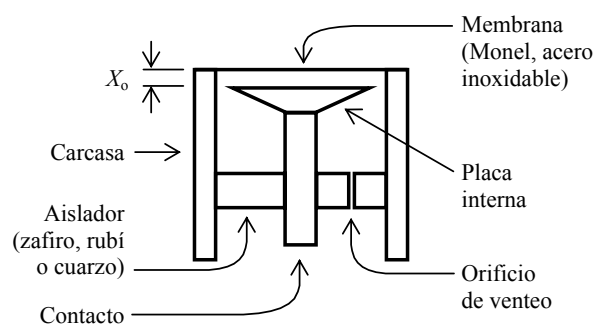
### 5.1 Micrófonos

El transductor por excelencia en mediciones acústicas es el micrófono de instrumentación. Este transductor convierte una presión sonora  $p(t)$  en una tensión eléctrica  $v(t)$  a través de una relación idealmente lineal:

$$v(t) = S_m p(t). \quad (5.1)$$

donde  $S_m$  es la constante de sensibilidad. En un micrófono real,  $S_m$  depende de la frecuencia y de la orientación y tipo de campo sonoro.

Si bien existen diversos principios físicos que permiten implementar micrófonos, los micrófonos de instrumentación son siempre capacitivos, debido a que en teoría tienen una respuesta plana en un rango más amplio de frecuencia.<sup>88</sup> En la figura 72 se ilustra el esquema de un micrófono capacitivo.



**Figura 72.** Diagrama esquemático de un micrófono capacitivo de instrumentación.

El principio de funcionamiento es el siguiente. Hay una placa interna (*back plate*) rígida con algunas perforaciones y un diafragma o membrana tensa, ambas conductoras. Se forma, por lo tanto, un capacitor de capacidad

$$C_o = \epsilon_o \frac{\pi a^2}{X_o} \quad (5.2)$$

<sup>88</sup> Los micrófonos dinámicos (de bobina móvil) tienen, en cambio, limitaciones en baja frecuencia por la ley de Faraday (la tensión generada depende de la velocidad y por lo tanto, de la frecuencia) y en alta frecuencia por la gran inercia de la bobina (la aceleración, y por lo tanto la velocidad, disminuyen con la frecuencia). Para lograr una respuesta plana requieren sofisticados filtros acústicos que compensen la respuesta natural

donde  $\varepsilon_0 = 8,84 \times 10^{-12}$  F/m es la constante dieléctrica del aire,  $a$  el radio de la membrana,  $X_0$  la distancia de reposo entre la membrana y la placa posterior.

Supongamos ahora que en las placas hay cargas fijas y opuestas  $\pm Q_0$ . Entonces entre ambas placas habrá una tensión  $E_0$ , denominada tensión de polarización:

$$E_0 = \frac{Q_0}{C_0}. \quad (5.3)$$

Al incidir una onda sonora sobre el diafragma, éste vibrará con una amplitud proporcional a la amplitud de la presión sonora del lado externo (el lado interno está a presión aproximadamente constante), por lo que la distancia entre el diafragma y la placa interna variará. Entonces, de acuerdo con la ecuación (5.2), variará la capacidad y, en consecuencia, también variará la tensión:

$$E = \frac{Q_0}{C}. \quad (5.4)$$

Se puede demostrar que, con buena aproximación, la componente variable de la tensión está dada por (Miyara, 2007; Miyara, 2013)

$$v(t) = \frac{E_0}{X_0} \frac{1}{\pi a^2 \left( \frac{8\tau}{\pi a^4} + \frac{\gamma P_0}{V_0} \right)} p(t) \quad (5.5)$$

donde  $\tau$  es la tensión lineal del diafragma,  $\gamma = 1,4$  el cociente entre los calores específicos a presión constante y a volumen constante, y  $V_0$  el volumen de la cavidad del micrófono.<sup>89</sup> Si la tensión aplicada al diafragma es alta se puede despreciar el segundo término del denominador con lo que la aproximación se simplifica:

$$v(t) = \frac{E_0}{X_0} \frac{a^2}{8\tau} p(t) \quad (5.6)$$

La sensibilidad es, entonces

$$S_m = \frac{E_0}{X_0} \frac{a^2}{8\tau} \quad (5.7)$$

En un micrófono de instrumentación es deseable tener una elevada sensibilidad con el fin de aumentar la relación señal/ruido. Cuanto mayor sea la relación señal/ruido más bajo será el mínimo nivel medible por el instrumento.

Aunque en apariencia la tensión lineal  $\tau$  de la membrana debería ser baja para obtener una alta sensibilidad, ello repercute negativamente en la respuesta en frecuencia, reduciéndola. Este compromiso resulta inadmisibles en un micrófono de instrumentación. Por ello se utilizan tensiones lineales elevadas (del orden de 1000 N/m), lo cual a

<sup>89</sup> Este volumen  $V_0$  no es igual a  $\pi a^2 X_0$  porque la zona entre la placa interna y la membrana está comunicada mediante perforaciones con el volumen detrás de la placa interna



su vez requiere utilizar materiales de gran límite elástico. Se utiliza Monel, una aleación de níquel (67 %) y cobre (33 %) sumamente resistente a la corrosión y a los esfuerzos mecánicos, o bien aceros inoxidables especiales o titanio.

Para lograr una sensibilidad alta se debe, por consiguiente, aumentar  $E_0$  y reducir  $X_0$ . Se trabaja con tensiones de 200 V y una separación de reposo entre la membrana y la placa interna de unos 20  $\mu\text{m}$  (el espesor de la membrana suele ser menor de 10  $\mu\text{m}$ ). Se puede constatar que con un nivel de presión sonora de 110 dB el desplazamiento medio es del orden de 1  $\mu\text{m}$ . También puede incrementarse el radio de la membrana, aunque salvo en algunas aplicaciones de laboratorio hoy en día se usan casi exclusivamente micrófonos de media pulgada (aproximadamente 13 mm).

## 5.2 Polarización

En la introducción anterior supusimos que de alguna manera existía una carga aplicada a las placas del capacitor que constituye el micrófono. Hay dos formas de hacerlo. La primera es aplicando una tensión de polarización externa a través de una resistencia lo suficientemente alta como para que no atenúe (por efecto de divisor de tensión) la tensión generada por el proceso de transducción. El capacitor se cargará a esa tensión y habrá adquirido la carga necesaria.

La segunda forma es mediante el uso de un material especial conocido como *electret*. Se trata de un polímero que en condiciones normales es un dieléctrico (aislante) pero al calentarlo a elevada temperatura se vuelve conductor. En ese estado se lo puede cargar utilizándolo como una de las placas de un capacitor. Al enfriarse, las cargas eléctricas quedan atrapadas, proveyendo la carga  $Q_0$  que requiere el micrófono para poder funcionar. Estas cargas atraen cargas en la otra placa y crean el campo eléctrico necesario para el funcionamiento.

Si bien los micrófonos electret son más costosos, son una buena elección para su uso en sonómetros, ya que se evita la complicación adicional de utilizar un convertidor conmutado para elevar la tensión pequeña de la batería hasta los 200 V normalmente utilizado en la polarización de micrófonos de instrumentación.

## 5.3 Preamplificador

Circuitalmente, el micrófono es una fuente de tensión con un capacitor en serie. El orden de magnitud de dicho capacitor es de algunas decenas de pF, lo cual significa que la impedancia interna a 20 Hz es del orden de 100 M $\Omega$ . Si se le conectara directamente la resistencia de entrada  $R_i$  del amplificador de entrada del sonómetro, a menos que ésta fuera muy alta, la configuración divisora de tensión (figura 73a) atenuaría demasiado la tensión a la entrada del amplificador. A eso se agregaría la capacidad parásita de los cables,  $C_{\text{cable}}$ , que atenuaría la señal en alta frecuencia. En lugar de ello se utiliza un preamplificador consistente en uno o dos transistores de efecto de campo en configuración como seguidor de tensión (ganancia 1) cuya única finalidad es realizar una adaptación de impedancia, reduciendo la elevada impedancia interna del micrófono a unos pocos cientos de ohm (figura 73b). El preamplificador se sitúa normalmente muy próximo al micrófono, alojándolo en el propio vástago metálico que empalma el micrófono al cuerpo del sonómetro. Éste sirve, además, como blindaje electrostático.

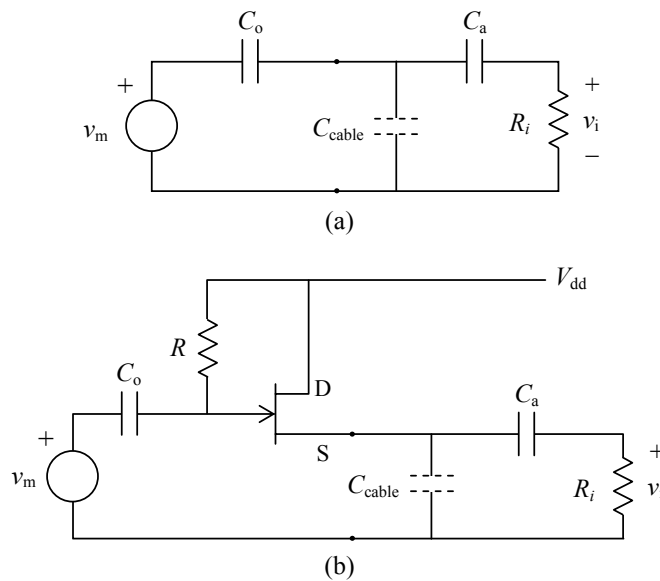
## 5.4 Campos sonoros

La distribución en el tiempo y en el espacio de la presión sonora se suele denominar *campo sonoro*. Si bien la complejidad de cualquier campo sonoro de la realidad es muy grande en casi todos los casos, desde el punto de vista metrológico casi todas las

situaciones pueden aproximarse por cuatro casos límite con propiedades bien definidas: campo libre, campo de presión, campo difuso y campo estacionario.

#### 5.4.1 Campo libre

Se denomina *campo libre* a cualquier campo sonoro en el cual está perfectamente definida la dirección de propagación. Generalmente se dan en ámbitos abiertos donde no haya superficies reflectantes. A su vez pueden ser planos, cilíndricos o esféricos, pero para fuentes relativamente alejadas (a partir de unos pocos metros) en general pueden considerarse planos. Es el caso típico de las mediciones en exteriores, aunque también se da en el interior de una *cámara anecoica* (es decir, una sala recubierta de materiales y/o estructuras absorbentes del sonido de modo de reducir prácticamente a 0 toda la energía reflejada).



**Figura 73.** (a) Modelo eléctrico de la conexión entre un micrófono y su preamplificador. (b) El micrófono con un preamplificador a FET en conexión como seguidor. En ambos casos el capacitor  $C_a$  es un desacoplador de corriente continua y  $R_i$  es la resistencia de entrada del amplificador remoto.  $V_{dd}$  polariza a la vez al micrófono y al FET.

#### 5.4.2 Campo de presión

El *campo de presión* es un campo sonoro que oscila en fase y con la misma amplitud en todos los puntos de una región en el espacio de interés. En general esta situación se da en volúmenes pequeños comparados con la longitud de onda mínima de interés. Una situación típica es en un acoplador acústico, por ejemplo un calibrador, en el cual un generador de sonido y un micrófono a ensayar comparten un pequeño volumen cerrado de unos pocos centímetros cúbicos. Otro caso es el de un oído artificial para prueba de auriculares. Se admite en general que en cualquier recinto cuya máxima dimensión es menor que  $1/36$  de la longitud de onda el campo sonoro es de presión.

### 5.4.3 Campo difuso

Un campo sonoro es *difuso* o *aleatorio* en una zona cuando cualquier dirección de propagación es igualmente probable. Generalmente se tiene un campo difuso en recintos cerrados de grandes dimensiones con respecto a la máxima longitud de onda, especialmente los de forma irregular o que contienen variedad de objetos. En general, el campo difuso posee, además, un espectro continuo.

### 5.4.4 Campo estacionario

Se tiene un *campo estacionario* en una zona siempre que la presión sonora oscile en todos sus puntos en fase o a lo sumo en contrafase. Sucede únicamente con tonos senoidales en pequeños recintos. Un ejemplo es un *tubo de Kundt* o un *tubo de impedancia* utilizado para medir absorción sonora.

## 5.5 Micrófonos y campos sonoros

Un mismo micrófono se comporta de diferente manera cuando se lo coloca en campos sonoros de diferente tipo. Para ver las razones, tengamos en cuenta que salvo que el micrófono fuera de dimensiones muy pequeñas con respecto a la mínima longitud de onda de interés, su sola presencia es capaz de alterar el campo sonoro. En efecto, se producen en el cuerpo del micrófono (y su preamplificador) fenómenos de reflexión y difracción que se suman al campo original. También interviene la impedancia acústica de radiación, ya que al vibrar la membrana por acción de la presión externa, radia una presión adicional que, nuevamente, se superpone al campo existente en ausencia del micrófono.

Reconociendo esta realidad física y la necesidad metrológica de evaluar diferentes campos sonoros, los fabricantes producen micrófonos que, sin dejar de ser micrófonos que detectan y convierten la presión sonora real a la que están sometidos (incluyendo su propio efecto), tienen correcciones en su respuesta que los hacen aptos para los distintos campos sonoros.<sup>90</sup>

Tendremos, en primer lugar, los *micrófonos de campo de presión*, que presentan una respuesta en frecuencia plana en presencia de un campo sonoro de presión. Son ideales para medir campos de presión y campos estacionarios.

Luego están los *micrófonos de campo libre*. En éstos la respuesta ha sido modificada de manera de presentar una respuesta en frecuencia plana ante un campo libre, siempre que se coloque el micrófono con la membrana apuntando hacia la dirección de procedencia del sonido.

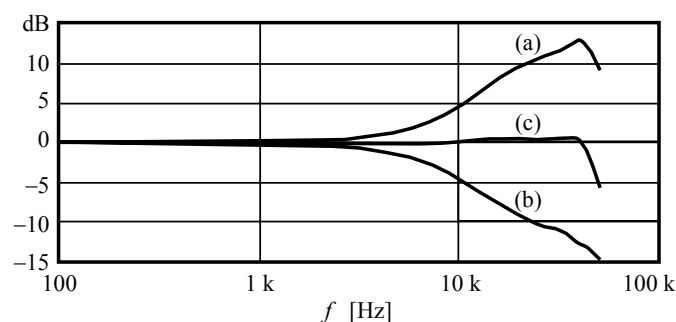
Por último tenemos los *micrófonos de campo difuso*, desarrollados para presentar una respuesta plana cuando están inmersos en un campo difuso.

Algunos sonómetros que utilizan procesamiento digital de señales permiten modificar la respuesta en frecuencia mediante la aplicación de un filtro digital de modo que el sonómetro parezca responder como si tuviera un micrófono con una respuesta diferente a la real.

<sup>90</sup> En sonido profesional existen otras variantes de micrófonos que intrínsecamente *no son* de presión. Son micrófonos que tienen total o parcialmente abierto el fondo de la cápsula, de modo que la membrana está expuesta a la presión sonora de ambos lados. Cuando la abertura es total se denomina *micrófono de gradiente de presión*, ya que su salida es proporcional no a la presión sonora sino al gradiente espacial de la presión. Así, en una dirección en la que el sonido varía mucho será sensible, mientras que en una dirección en la cual no varíe será completamente insensible. Son éstos los denominados *bidireccionales* o *figura de ocho*. Cuando la abertura es parcial, puede considerarse una superposición ponderada de la respuesta de un micrófono de presión y uno de gradiente de presión. Se encuentran entre otros ejemplos, los micrófonos cardioides, supercardioides e hipercardioides.

## 5.6 Respuesta en frecuencia

La respuesta en frecuencia de los diferentes micrófonos está diseñada para ser plana cuando se los utiliza en un campo sonoro del tipo preferencial. Este uso puede implicar también respetar una orientación dada. Por ejemplo, en el caso de campo libre, la respuesta es plana en dirección frontal. En la figura 74 se muestra, para un campo libre procedente de dicha dirección, la presión sonora que aparece en el diafragma en presencia del micrófono (a), la modificación en la respuesta en frecuencia del micrófono (b) para poder lograr una respuesta aproximadamente plana. En particular esto implica que sometido el micrófono a un campo de presión la respuesta tendrá una caída en alta frecuencia. Esto explica por qué al calibrar un sonómetro cuyo micrófono de 1/2" es de campo libre, la indicación es 0,15 dB menor que el valor nominal del calibrador. Así, si el calibrador es de 94 dB, la calibración correcta será 93,85 dB. Cuando ese sonómetro deba medir un campo libre de 94,0 dB, dará entonces la lectura correcta ya que la presión real en el diafragma será 0,15 dB mayor, quedando compensada por la respuesta del micrófono.

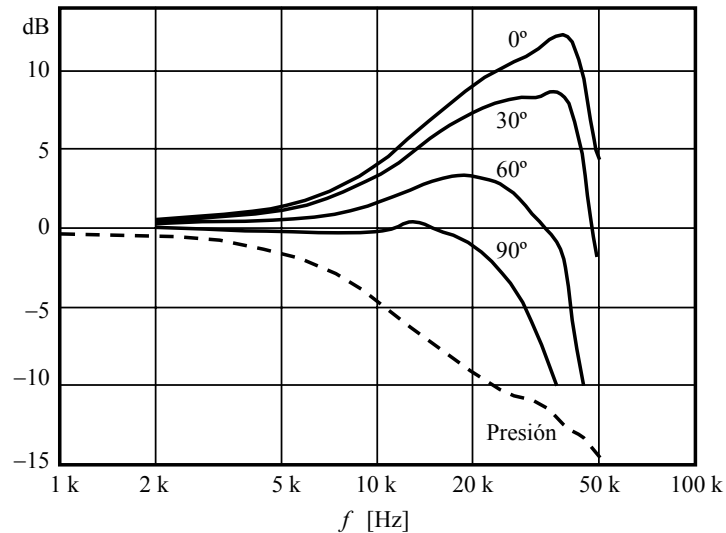


**Figura 74.** (a) Nivel de presión sonora en la membrana de un micrófono de 1/2" con su grilla protectora ante un campo libre de incidencia frontal, relativo al nivel de presión sonora en ausencia del micrófono. (b) Respuesta a la presión de un micrófono diseñado para tener una respuesta plana al campo libre. (c) Respuesta total. La compensación de (b) asociada al incremento de presión de (a) proporciona una respuesta aproximadamente plana (Brüel & Kjaer, 1996)

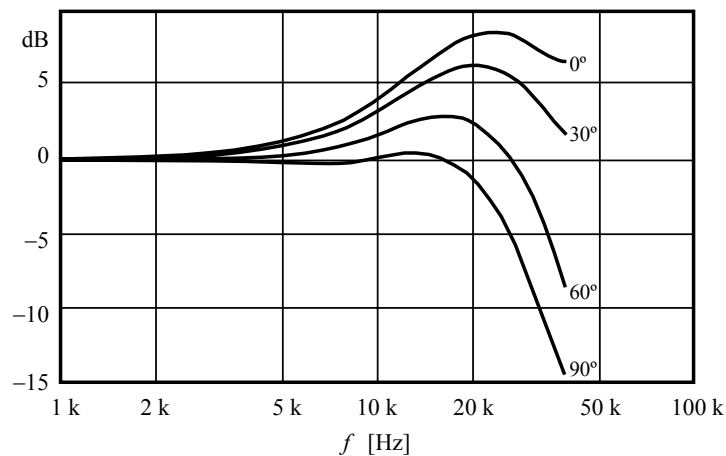
En la figura 75 se muestra la corrección que debe sumarse a la respuesta a la presión (indicada en línea de trazos) para obtener la verdadera respuesta del micrófono ante un campo libre procedente de diversos ángulos con respecto a la dirección axial frontal. Vemos que existe una caída importante en alta frecuencia. En 10 kHz la caída es de 1 dB para 30°, 3 dB para 60° y 5 dB para 90°. Esto muestra que la respuesta para un campo libre puede ser significativamente distinta según la orientación del micrófono respecto a la dirección del campo sonoro.

En la figura 76 se indican las correcciones correspondientes al caso en que se quita la grilla protectora al micrófono. vemos que las curvas son más regulares. La grilla, debido a su forma compleja, con ranuras de diversos tamaños, produce un efecto apreciable, sobre todo en alta frecuencia. No es recomendable retirar la grilla, salvo en situaciones de laboratorio altamente controladas donde se quisiera aprovechar la mayor regularidad de la respuesta.

Una situación interesante se da al medir ruido vehicular, especialmente cuando el tránsito es liviano y está constituido por la circulación de vehículos aislados. En efecto, el ángulo de procedencia del sonido de un vehículo individual cambia durante el paso del mismo, por lo que si se deseara medir realmente el ruido del vehículo debería efectuarse una corrección variable en el tiempo. Dicha corrección es posible si se conoce aproximadamente la velocidad, utilizando filtros FFT.



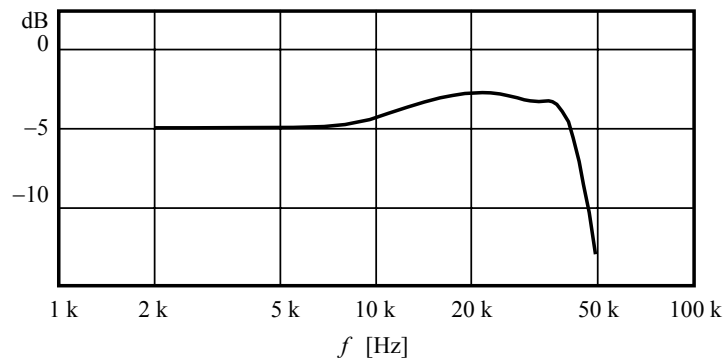
**Figura 75.** Corrección para campo libre con respecto a la respuesta de presión de un micrófono de 1/2" típico, con grilla protectora, para diferentes ángulos de incidencia (Beranek, 1993)



**Figura 76.** Corrección para campo libre con respecto a la respuesta de presión de un micrófono de 1/2" típico, sin grilla protectora, para diferentes ángulos de incidencia (Brüel & Kjaer, 1996)

Si bien es frecuente realizar mediciones en interiores donde puede presumirse que el campo sonoro es aproximadamente difuso, en realidad los micrófonos de campo difuso son raros. La mayoría de los sonómetros viene equipada con un micrófono de campo libre. Afortunadamente la corrección para transformar una respuesta libre en una respuesta difusa se reduce a una única curva de corrección. Esto es así porque justamente el campo difuso no depende de la orientación. Esta corrección en realidad surge de integrar todas las correcciones para los diferentes ángulos mediante un procedimiento que está normalizado por la Norma IEC 61183. En algunos casos se proporciona junto con el micrófono un anillo plástico de una forma apropiada que realiza un filtrado acústico de tipo difractivo, que equivale a la curva de corrección. En otros casos el propio sonómetro posee una configuración que permite, mediante un filtro implementado digitalmente en un DSP, transformar la respuesta del instrumento a una respuesta de campo difuso, es decir que las lecturas obtenidas corresponden al valor del campo difuso si se midiera con un micrófono de presión muy pequeño que no alterara el campo sonoro.

En la figura 77 se muestra la corrección con respecto a la respuesta de presión para obtener una respuesta de campo difuso. Debe tenerse en cuenta que esta la corrección no es la que se debe aplicar a las lecturas obtenidas mediante un micrófono de campo libre, ya que el mismo tiene ya una corrección por campo libre de incidencia frontal que habría que restar antes de proceder a aplicar la corrección de campo difuso.



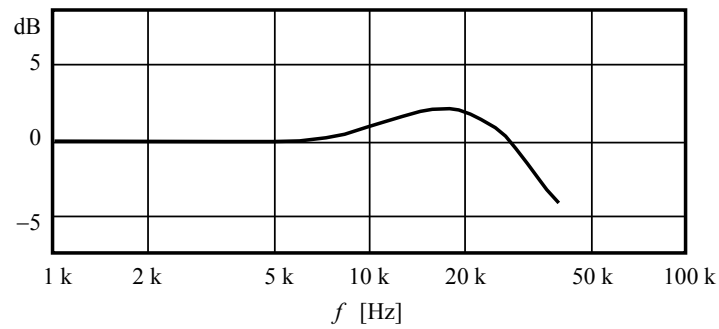
**Figura 77.** Corrección para campo difuso con respecto a la respuesta de presión de un micrófono de 1/2" típico, con grilla protectora (Bera- nek, 1993)

La figura 78 corresponde al caso en que se quita la grilla, aunque debe tenerse en cuenta que las correcciones deberán efectuarse respecto a la respuesta de presión también sin grilla.

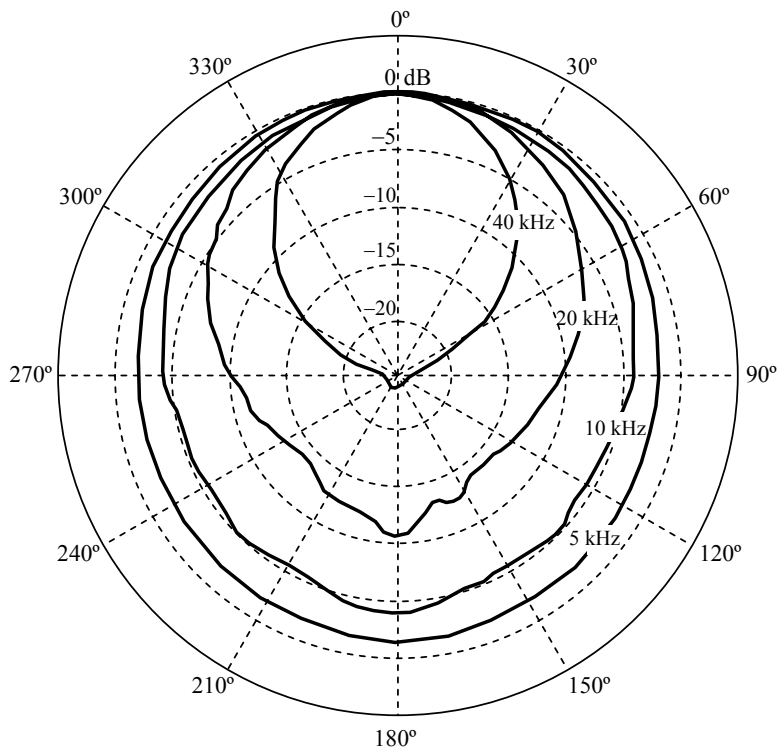
### 5.7 Respuesta direccional

La respuesta en frecuencia y la respuesta direccional están íntimamente relacionadas. Otra forma de presentar esta información es mediante diagramas polares como el que se muestra en la figura 79.

Vemos que en el rango hasta 20 kHz el error es menor de 3 dB dentro de un ángulo de  $\pm 30^\circ$ , pero para desviaciones mayores de la orientación frontal los errores pueden ser mayores.



**Figura 78.** Corrección para campo difuso con respecto a la respuesta de presión de un micrófono de 1/2" típico, sin grilla protectora (Brüel & Kjaer, 1996)



**Figura 79.** Respuesta direccional de un micrófono de campo libre para diferentes frecuencias, normalizada a la respuesta frontal. El ángulo está medido respecto a la dirección frontal (Brüel & Kjaer, 1996).

Lo anterior no tiene en cuenta el cuerpo del sonómetro, que pese a su forma en general aerodinámica (que minimiza el efecto sobre el campo sonoro), puede influir considerablemente. Las especificaciones técnicas de los sonómetros normalmente in-

cluyen estas respuestas direccionales. Para minimizar los efectos del instrumento es recomendable utilizar un cable de extensión de micrófono, aunque en la mayoría de las situaciones normales esta precaución es innecesaria.<sup>91</sup>

## 5.8 Ruido

Hay dos ruidos asociados a un micrófono capacitivo. El primero es el *ruido intrínseco* de la propia cápsula y el segundo el que agrega el preamplificador. El ruido intrínseco se debe a la vibración térmica de la membrana, resultado de pequeñas presiones causadas por el movimiento aleatorio de las moléculas en el aire.<sup>92</sup> La densidad espectral de dicho ruido se comporta en forma similar a la del ruido térmico eléctrico, y su densidad espectral puede calcularse con la misma fórmula reemplazando la resistencia eléctrica por la resistencia acústica del diafragma:

$$\overline{p^2} = 4kTR_a \quad (5.8)$$

donde  $k = 1,38 \times 10^{-23}$  J/K es la constante de Boltzmann,  $T$ , la temperatura absoluta en K y  $R_a$  es la resistencia acústica en  $\text{Ns/m}^5$  del diafragma. La resistencia acústica está definida como la componente real del cociente entre los fasores de la presión sonora y el caudal (volumen desplazado por unidad de tiempo). La resistencia acústica surge principalmente del amortiguamiento que se agrega al diafragma para atenuar su resonancia.

Existen micrófonos de bajo ruido en los que se reduce la resistencia acústica de amortiguamiento al precio de no controlar la resonancia. Para tener una respuesta plana se utiliza un ecualizador electrónico con una respuesta antirresonante. Tal ecualización podría efectuarse también por software.

Dado que a igualdad de amortiguamiento, la resistencia acústica es proporcional al área del micrófono, la densidad espectral de la presión de ruido resultará también proporcional al área. En consecuencia, la presión eficaz de ruido, dada por

$$P_{\text{ef}} = \sqrt{\int_0^{f_{\text{máx}}} \overline{p^2}(f) df} = \sqrt{4kTR_a f_{\text{máx}}} \quad (5.9)$$

es proporcional a la raíz cuadrada del área, es decir, proporcional al radio. Como la sensibilidad, según (5.7), es proporcional al cuadrado del radio, resulta que la relación señal / ruido es directamente proporcional al radio. Así, un micrófono de 1" tendrá el doble de relación señal / ruido que uno de 1/2". Expresado en decibeles, su relación señal / ruido será 6 dB mejor.

El ruido térmico de presión tiene un espectro plano, es decir, es un ruido blanco. Por esta razón, el ruido eléctrico equivalente se comporta de igual forma que la respuesta a la presión del micrófono, más estrictamente, como el cuadrado de su respuesta en frecuencia (ver figura 74a). Sin embargo, la señal aprovechable se obtiene luego del preamplificador, por lo que debe agregarse el ruido de éste, que está formado por una componente de ruido blanco asociada a las resistencias y la juntura del FET (en polari-

<sup>91</sup> Por otra parte los cables de extensión suelen ser extremadamente (y probablemente injustificadamente) costosos si se los compara con un cable de similar calidad

<sup>92</sup> Este movimiento es incorrectamente llamado, a veces, *movimiento Browniano*. El verdadero movimiento browniano no es de un objeto grande como el diafragma de un micrófono ni el de las moléculas de aire, sino el que éstas imprimen a pequeñas partículas de polvo en suspensión visibles con algo de aumento



zación inversa) y una componente de ruido rosa ( $1/f$ ) con énfasis en baja frecuencia. El resultado es un ruido total donde predomina el ruido de baja frecuencia del preamplificador y la componente de frecuencias medias del micrófono. En frecuencias más altas que las de funcionamiento normal, el ruido del micrófono vuelve a caer porque su respuesta comienza a disminuir ya que queda dominada por la inercia del diafragma.

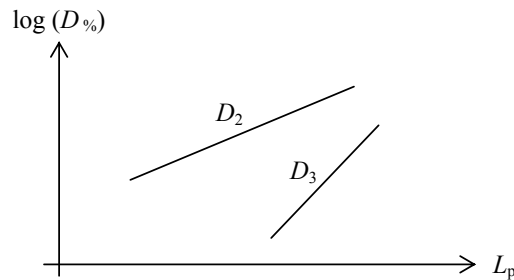
El ruido del micrófono se suele especificar indicando un nivel de presión sonora externo que daría el mismo ruido eléctrico en un micrófono ideal (sin ruido) de igual sensibilidad, tanto con ponderación plana (Z) o con ponderación A. Debido al ruido de baja frecuencia del preamplificador, el ruido acústico equivalente con ponderación A resulta menor. También se lo suele especificar en bandas de tercio de octava.

### 5.9 Distorsión

La distorsión en los micrófonos capacitivos se debe a dos factores. El primero es que la carga  $Q_0$ , que supusimos constante, en realidad varía, y lo hace tanto más cuanto mayor sea la corriente que circula como consecuencia de las variaciones de tensión que constituyen la propia señal. Esta corriente circula por los elementos conectados a la celda básica: la capacidad parásita entre la placa interna y la carcasa (que se encuentra conectada a la membrana), la resistencia de entrada del preamplificador, que es muy alta (cientos de  $M\Omega$ ), y la capacidad de entrada del preamplificador.

El segundo factor es el hecho de que la fórmula (5.5) es sólo una aproximación e ignora términos de orden superior que para pequeñas amplitudes son despreciables pero no para grandes amplitudes. Estos se originan en que la deformación de la membrana deja de ser lineal con la presión a niveles muy grandes.

Ambos factores generan una distorsión de segundo y tercer armónicos que crece como una potencia de la amplitud, de modo que en diagrama lineal en el nivel de presión sonora y logarítmico en la distorsión, tal como se ilustra en la figura 80. Estas distorsiones se acercan al 1 % para niveles de presión sonora del orden de 140 dB. La distorsión disminuye en los micrófonos más pequeños, debido a que la deformación de la membrana es menor a igualdad de todas las otras condiciones.



**Figura 80.** Distorsiones porcentuales de segundo y tercer armónico,  $D_2$  y  $D_3$ , en un micrófono capacitivo (Brüel & Kjaer, 1996).

El preamplificador contribuye muy poco a la distorsión excepto en el extremo superior del rango dinámico, donde incrementa produce un ligero incremento.

### 5.10 Micrófonos micromaquinados

Existe una nueva tecnología que se venía utilizando en los teléfonos celulares pero que está evolucionando hacia la instrumentación. Se trata de los sistemas microelectromecánicos, MEMS (Micro-Electro-Mechanical Systems) y constituye una evolución natural de la tecnología de circuitos integrados por incursión en la mecatrónica (combinación entre la microelectrónica y la mecánica) (Lee, 2009).

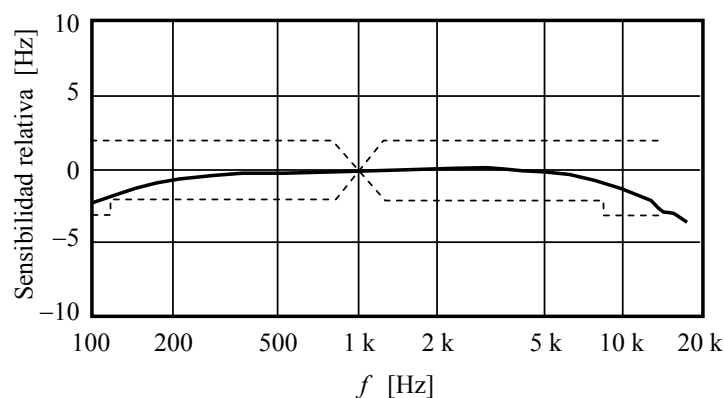
Un ejemplo son los micrófonos capacitivos en miniatura ADMP421 y ADMP521 de Analog Devices (Analog Devices, 2012). Una característica muy interesante es que proveen salida digital de *modulación de densidad de pulsos* (PDM) generada por un modulador sigma-delta con filtro de modelado de ruido de cuarto orden (equivale a cuatro integraciones). La frecuencia del reloj recomendada es de 2,4 MHz, esto es, se trata de un sobremuestreo de 64 veces con respecto a una frecuencia de muestreo de 37,5 kHz, lo que garantiza una frecuencia máxima de 16 kHz. Un inconveniente es que la frecuencia inferior es de 100 Hz.

Este micrófono posee un ruido intrínseco equivalente a del orden de 30 dBA, y una sensibilidad de -26 dBFS. La unidad dBFS es el decibel referido al fondo de escala (full-scale decibel). Esta especificación significa que con un tono de nivel de presión sonora 94 dB el pico de la salida digital estará 26 dB por debajo del máximo valor digital representable (Lewis, 2012). Está especificado como omnidireccional, característica que se debe a su pequeño tamaño (4 mm × 3 mm × 1 mm), lo cual lo convierte en un micrófono de presión que no altera significativamente el campo sonoro en todo el rango de interés.

Un campo de aplicación muy promisorio para estos dispositivos es el de *beam forming*, donde se utiliza un arreglo de micrófonos para generar un sensor de direccionalidad programable, basado en los diferentes retardos que tiene la onda sonora en llegar a los diferentes micrófonos del arreglo según la dirección de procedencia del sonido (Analog Devices, 2013; Lewis, 2012a). Esta técnica se está usando con éxito para realizar mapas de radiación sonora de máquinas y vehículos.

#### 5.10.1 Respuesta en frecuencia

La respuesta en frecuencia especificada de estos dispositivos se muestra en la figura 81, donde la línea llena representa una respuesta medida y la línea de trazos los límites garantizados.



**Figura 81.** Respuesta en frecuencia de los micrófonos micromaquinados ADMP421 relativa a la respuesta a 1 kHz (Lewis, 2012).

### 5.11 Auriculares audiométricos

El micrófono convierte señal acústica (presión sonora) en señal eléctrica (tensión). En muchas mediciones es necesaria la transducción inversa. Un caso de interés es el de las mediciones psicoacústicas, donde interviene el oído humano. En este caso se utilizan auriculares audiométricos. El problema a resolver en este caso es la relación entre la señal eléctrica que se aplica a los terminales del auricular y la presión sonora generada. Sin embargo, esto que parecería una especificación sencilla se complica por varios motivos.

En primer lugar, la relación depende mucho de la conformación del oído individual, de la forma en que se colocan los auriculares, incluyendo la presión que ejercen sobre la cabeza. También depende del tipo de campo sonoro que nos interese utilizar como referencia. Por ejemplo, podría interesarnos generar en el oído un campo sonoro similar al que se generaría ante un campo libre o un campo difuso medido en la posición del sujeto estando éste ausente. Ambos diferirán porque el oído, junto con la cabeza y el torso, igual que un micrófono, tiene diferentes respuestas en frecuencia según la dirección de procedencia de la onda sonora. También podría interesarnos lograr un nivel de presión sonora en el tímpano. Éste diferirá de los anteriores debido a fenómenos de resonancia en el conducto auditivo.

Analizaremos el caso del auricular Sennheiser HDA 200, un auricular compatible con audiometrías de alta frecuencia (es decir que su respuesta está especificada hasta 16 000 Hz), debido a que ha sido posible recopilar información razonablemente completa, si bien la misma se encuentra desperdigada en distintas fuentes. Se trata de un auricular *circumaural*, es decir, que cubre completamente el oído, proveyendo una aislación acústica superior a los 28 dB por encima de los 1000 Hz.

En primer lugar, debido a la diversidad existente en el oído humano según el sujeto, la especificación debe realizarse mediante un acoplador acústico normalizado, siendo la norma utilizada la IEC 61318. En la tabla 5.1 se incluye la información provista por el fabricante, donde  $L_{A,C}$  es el nivel de presión sonora producida por el auricular audiométrico (A) en el acoplador (*coupler*, C).

**Tabla 5.1.** Respuesta en frecuencia del auricular audiométrico HDA 200 según el fabricante, medida en acoplador acústico Brüel & Kjaer Type 4153, que cumple con IEC 60318, y atenuación de los co-bertores (Sennheiser, s/f).

$f_k$ [Hz]	$L_{A,C}$ [dB] @ $V_{ef} = 0,5$ V	$S_{A,C}(f_k)$ [Pa/V]	$Aten$ [dB]
125	112,5	16,9	14,3
250	113,0	17,9	16,9
500	112,0	15,9	22,5
750	111,0	14,2	-
1 000	108,5	10,6	28,6
2 000	104,0	6,3	32,0
3 000	104,0	6,3	-
4 000	104,0	6,3	45,7
5 000	106,5	8,5	-
6 000	107,5	9,5	-
8 000	105,5	7,5	43,8
9 000	105,0	7,1	-
10 000	102,5	5,3	-
11 200	102,0	5,0	-
12 500	103,0	5,7	-
14 000	98,5	3,4	-
16 000	100,0	4,0	-

También se incluyó en la tabla la sensibilidad del auricular referida al acoplador, determinada para cada frecuencia audiométrica  $f_k$  según la fórmula

$$S_{A,C}(f_k) = \frac{P_{A,C}}{v_A} = 10^{\frac{L_{A,C}(f_k)}{20}} \frac{P_{\text{ref}}}{0,5 \text{ V}}. \quad (5.10)$$

donde  $v_A$  es la tensión eficaz aplicada al auricular, que según lo especificado, son 0,5 V.

Para convertir estos valores de sensibilidad referida a un acoplador,  $S_{A,C}(f_k)$  a la sensibilidad referida a campo libre (*free field*, FF),  $S_{A,FF}(f_k)$ , se utiliza el Anexo C de la norma ISO 389-8 que da la diferencia de *niveles de sensibilidad* referidos a campo libre,  $G_F$  y a un acoplador acústico,  $G_C$ :

$$G_F(f_k) - G_C(f_k) = 20 \log \frac{S_{A,FF}}{S_0} - 20 \log \frac{S_{A,C}}{S_0} = 20 \log \frac{S_{A,FF}}{S_{A,C}}, \quad (5.11)$$

donde  $S_0$  es una sensibilidad de referencia arbitraria (normalmente 1 Pa/V, aunque en este caso es irrelevante). Dicha corrección se incluye en la tabla 5.2, para las bandas normalizadas de tercio de octava entre 125 Hz y 8000 Hz.

**Tabla 5.2.** Corrección entre la sensibilidad referida al acoplador y a campo libre para el HDA 200 según ISO 389-8 Anexo C.

$f_k$ [Hz]	$G_F(f_k) - G_C(f_k)$ [dB]
125	-5,0
160	-4,5
200	-4,5
250	-4,5
315	-5,0
400	-5,5
500	-2,5
630	-2,5
800	-3,0
1000	-3,5
1250	-2,0
1600	-5,5
2000	-5,0
2500	-6,0
3150	-7,0
4000	-13,0
5000	-14,5
6300	-11,0
8000	-8,5

Despejando, se obtiene

$$S_{A,FF} = 10^{\frac{G_F(f_k) - G_C(f_k)}{20}} S_{A,C}. \quad (5.12)$$

Para obtener la sensibilidad del auricular audiométrico referida al campo difuso (*diffuse field*, DF),  $S_{A,DF}$ , dado que no hay un equivalente al Anexo C de la ISO 389-8 para campo difuso, debemos recurrir a la Norma ISO 11904-1 que proporciona las correcciones entre los niveles de presión sonora en campo libre,  $\Delta L_{FF}$  y en el tímpano, y entre los niveles en campo difuso y en el tímpano,  $\Delta L_{DF}$ , definidas como

$$\Delta L_{FF}(f_k) = L_T(f_k) - L_{FF}(f_k). \quad (5.13)$$

$$\Delta L_{DF}(f_k) = L_T(f_k) - L_{DF}(f_k). \quad (5.14)$$

Estos valores son los que hay que sumar al valor medido del campo libre o difuso según el caso, con el sujeto ausente, para obtener el nivel de presión sonora que se mediría en el tímpano con una pequeña sonda microfónica colocada cerca del tímpano. Dichas correcciones se transcriben en la tabla 5.3

**Tabla 5.3.** Corrección entre el nivel de presión sonora en campo difuso (DF) y libre (FF) y el nivel de presión sonora correspondiente en el tímpano según ISO 11904-1

$f_k$ [Hz]	$\Delta L_{DF}(f_k)$ [dB]	$\Delta L_{FF}(f_k)$ [dB]
100	0,0	0,0
125	0,2	0,2
160	0,4	0,4
200	0,6	0,6
250	0,8	0,8
315	1,1	1,1
400	1,5	1,5
500	2,1	2,0
630	2,8	2,3
800	3,3	3,1
1 000	4,1	2,7
1 250	5,5	2,9
1 600	7,7	5,8
2 000	11,0	12,4
2 500	15,3	15,7
3 150	15,7	14,9
4 000	12,9	13,2
5 000	10,6	8,9
6 300	9,4	3,1
8 000	9,5	-1,4
10 000	6,8	-3,8
12 500	3,8	-0,1
16 000	0,7	-0,4

Tomando el nivel en el tímpano como la magnitud común a ambas determinaciones, es posible obtener la diferencia entre los niveles de presión sonora en un campo difuso y un campo libre que sean auditivamente equivalentes:

$$L_{DF}(f_k) - L_{FF}(f_k) = \Delta L_{FF}(f_k) - \Delta L_{DF}(f_k). \quad (5.15)$$

De (5.12) y (5.15) resulta la sensibilidad del auricular audiométrico referida al campo difuso:

$$S_{A,DF} = 10^{\frac{\Delta L_{FF}(f_k) - \Delta L_{DF}(f_k)}{20}} 10^{\frac{G_F(f_k) - G_C(f_k)}{20}} S_{A,C}. \quad (5.16)$$

Por último, si estamos interesados en la sensibilidad referida al tímpano,  $S_{A,T}$ , basta aplicar sólo la corrección de la tabla 5.3 para campo libre a (5.12):

$$S_{A,T} = 10^{\frac{\Delta L_{FF}(f_k)}{20}} 10^{\frac{G_F(f_k) - G_C(f_k)}{20}} S_{A,C}. \quad (5.17)$$

Las fórmulas (5.12), (5.16) y (5.17) permiten caracterizar el auricular audiométrico para la mayoría de las situaciones de interés. No lo caracteriza para campo libre de incidencia no frontal (por ejemplo, azimutal o de elevación). Ese problema se resuelve utilizando las *funciones de transferencia relativas a la cabeza*, HRTF (head-related transfer functions) tema en el que no incursionaremos.

EJEMPLO: Supongamos que deseamos administrar a través de los auriculares HDA 200 un tono de 4 kHz que sea equivalente a un tono de 60 dB en campo difuso. Determinaremos la tensión eficaz a aplicar. Para ello apliquemos la fórmula (5.16) para obtener la sensibilidad correspondiente a campo difuso:

$$S_{A,DF} = 10^{\frac{13,2 - 12,9}{20}} \times 10^{\frac{-13}{20}} \times 6,3 \text{ Pa/V} = 1,46 \text{ Pa/V}.$$

El nivel de presión sonora deseado, 60 dB, corresponde a una presión eficaz

$$P_{ef} = 20 \times 10^{-6} 10^{60/20} = 0,02 \text{ Pa},$$

de donde la tensión eficaz requerida será

$$v_A = \frac{P_{ef}}{S_{A,DF}} = \frac{0,02}{1,46} \text{ V} = 13,7 \text{ mV}$$

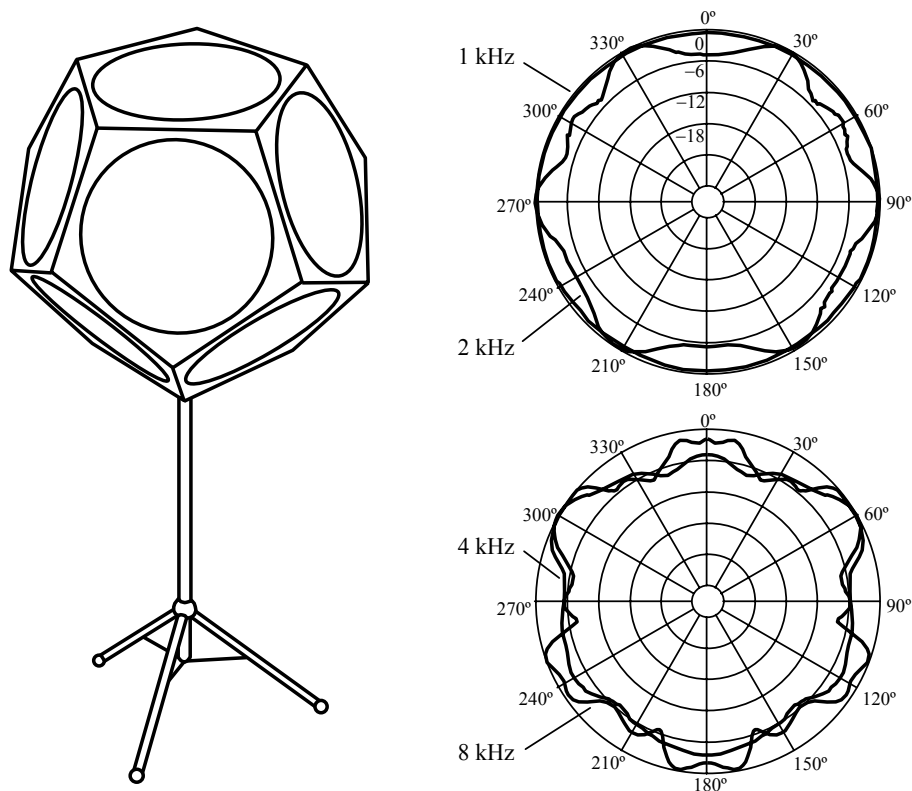
Caracterizando mediante un simple ensayo la constante entre la magnitud normalizada a un máximo de  $\pm 1$  de la señal digital y la tensión de salida de la placa de sonido (medida con un milivoltímetro de verdadero valor eficaz) es posible administrar por software una variable digital que se corresponda con los 13,7 mV deseados.

Puede ser necesario tener en cuenta el efecto de divisor de tensión que pudiera presentar la salida, ya que el auricular tiene una impedancia nominal de 40  $\Omega$ , por lo que una salida de línea normal que está en el orden de 100  $\Omega$  puede implicar una atenuación apreciable que será preciso compensar (también por software).

## 5.12 Fuentes omnidireccionales

Las fuentes omnidireccionales son otro tipo de transductor de gran interés en acústica. Se utilizan en mediciones de parámetros de acústica arquitectónica, tanto en laboratorio como in situ. Generalmente tienen forma de dodecaedro, en cada una de

cuyas caras hay un altavoz. En su versión más habitual, estos altavoces se encuentran alimentados por un mismo amplificador, de manera que todos radian en fase. La distribución regular de las caras asegura la máxima homogeneidad del campo radiado. Si bien se podría también pensar en un icosaedro, el problema es que la proporción entre el área de un círculo inscrito en un triángulo equilátero y el área del triángulo es menor que la correspondiente al dodecaedro, por lo que este último se parece más al ideal de una esfera pulsante en la cual todos sus puntos radian en fase. En la figura 82 se ilustra esquemáticamente una fuente omnidireccional típica.



**Figura 82.** Izquierda, fuente omnidireccional dodecaédrica. Derecha, respuesta polar típica para frecuencias 1 kHz, 2 kHz, 4 kHz y 8 kHz medida a 1 m de distancia.

Según la Norma ISO 3382-1 (equivalente a la Norma IRAM 4109-1), las fuentes omnidireccionales deben cumplir con una máxima variación de directividad dentro de arcos de circunferencia deslizantes de 30° tal como la que se indica en la tabla 5.4. La Norma ISO 140-3 (equivalente a la Norma IRAM 4063-3), por su parte da un criterio para omnidireccionalidad de la fuente basado en el *índice de directividad* DI calculado como:

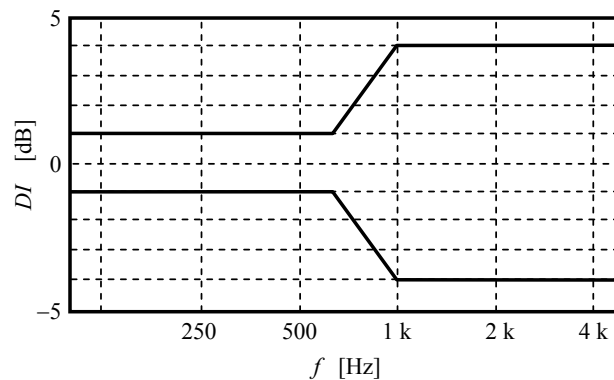
$$DI = L_{360^\circ} - L_{30^\circ}, \quad (5.18)$$

**Tabla 5.4.** Máxima desviación en arcos de circunferencia de 30° para una fuente omnidireccional (ISO 3382-1 / IRAM 4109-1)

$f$ [Hz]	250	500	1000	2000	4000	8000
$\Delta L$ [dB]	$\pm 1$	$\pm 1$	$\pm 1$	$\pm 3$	$\pm 5$	$\pm 6$

donde  $L_{360^\circ}$  es el promedio energético de los niveles en un arco circular de 360°, mientras que  $L_{30^\circ}$  es el promedio energético en un arco de 30°.

Para que la fuente pueda ser considerada omnidireccional el índice de directividad debe hallarse comprendido dentro de la plantilla indicada en la figura 83.



**Figura 83.** Límites aceptables del índice de directividad  $DI$  para que una fuente pueda considerarse omnidireccional para la Norma ISO 140-3 / IRAM 4063-3.

Las desviaciones de los valores aquí sugeridos no pueden corregirse por software, ya que no se trata sólo de una única respuesta en frecuencia.



## Capítulo 6

# PROCESAMIENTO DIGITAL DE SEÑALES

### 6.1 Señales discretas

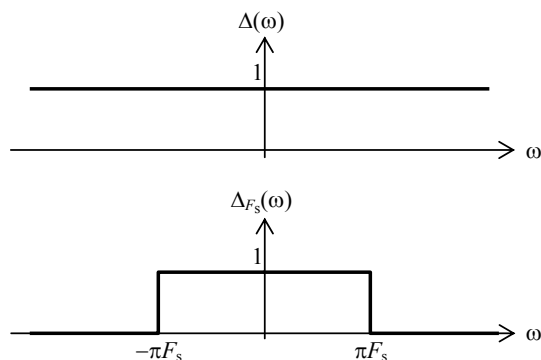
Una vez realizado el muestreo de una señal  $x(t)$  se obtiene una *señal discreta*,<sup>93</sup> es decir una sucesión de valores  $x(kT_s)$ , donde  $T_s$  es el periodo de muestreo. Haciendo abuso de notación, cuando no sea necesario explicitar la relación con el tiempo anotaremos la señal discreta como  $x(k)$ . Salvo que se diga algo en contrario, supondremos que la señal satisface la condición de Nyquist para la tasa de muestreo  $F_s = 1/T_s$  (ver apéndice 10).<sup>94</sup>

### 6.2 Impulso discreto

El *impulso discreto*, también llamado *pulso unitario* es una señal de importancia conceptual, ya que cumple el mismo papel que el impulso o función *delta de Dirac* en los sistemas continuos. Esta señal puede ser considerada como el resultado de muestrear una delta de Dirac  $\delta(t)$  que ha sido previamente pasada por un filtro antialias ideal. Para verlo, tengamos en cuenta que la transformada de Fourier de  $\delta(t)$  es igual a 1 en todo el espectro bilateral:

$$\Delta(\omega) = F\{\delta(t)\}(\omega) = \int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt = 1, \quad (6.1)$$

por lo cual al pasar por un filtro ideal de frecuencia de corte  $F_s/2$  su espectro queda restringido a  $\pm F_s/2$ , donde también vale 1, como se muestra en la figura 84.



**Figura 84.** Espectro de una delta de Dirac  $\delta(t)$  antes y después de pasar por un filtro pasabajos ideal de frecuencia de corte  $F_s/2$ .

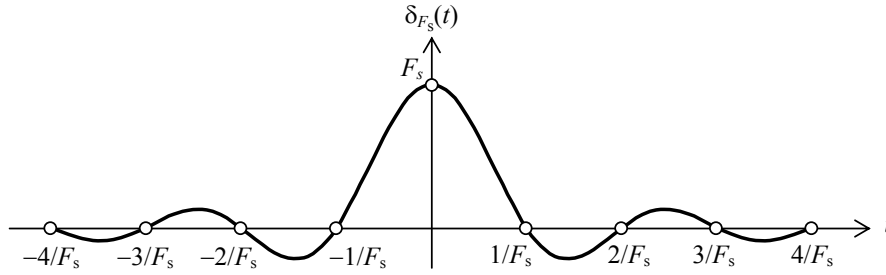
<sup>93</sup> Observemos que una señal discreta no necesariamente es digital. Existen ejemplos de señales discretas analógicas. Un ejemplo son los filtros de capacidades conmutadas (Miyara, 2004).

<sup>94</sup> Este capítulo no pretende dar un tratamiento completo del procesamiento de señales digitales. Es sólo una breve introducción que el lector interesado deberá profundizar en textos específicos, por ejemplo, el de Proakis y Manolakis (Proakis et al., 1998).

Podemos obtener la señal temporal filtrada mediante la transformada inversa de Fourier:

$$\delta_{F_s/2}(t) = \frac{1}{2\pi} \int_{-\pi F_s}^{\pi F_s} e^{j\omega t} d\omega = F_s \frac{\sin \pi F_s t}{\pi F_s t}. \quad (6.2)$$

Esta función se encuentra graficada en la figura 85. Podemos apreciar que la señal resultante es no causal (ya que empieza en  $-\infty$ ), situación característica de cualquier filtro ideal.<sup>95</sup> Sin embargo, en este caso muy particular, si muestreamos con tasa de muestreo  $F_s$ , todas las muestras son nulas excepto la ubicada en  $t = 0$ .



**Figura 84.** Impulso de Dirac filtrado con un filtro pasabajos ideal de frecuencia de corte  $F_s/2$ .

El impulso filtrado tiene un valor en el origen igual a  $F_s$ , que tiene unidades de frecuencia. Esto es porque arbitrariamente supusimos que la integral del impulso era 1 pero en el caso de señales impulsivas digitales, convendrá normalizarlo a 1, quedando definido como

$$u_o(k) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} \quad (6.3)$$

### 6.3 Señal como convolución con un impulso discreto

Dada una señal discreta  $x(k)$  podemos considerarla como la suma de impulsos discretos desplazados y ponderados:

$$x(k) = \sum_{m=-\infty}^{\infty} x(m) u_o(k - m) \quad (6.4)$$

Esta ecuación expresa la señal como una superposición de impulsos discretos ponderados y desplazados. Resultará útil para trabajar con sistemas discretos. El segundo miembro es el *producto de convolución* entre  $x$  y  $u_o$ .

<sup>95</sup> Hay un teorema de Paley-Wiener que implica que una función de espectro acotado (por ejemplo, como resultado de haber pasado por un filtro ideal) no puede ser idénticamente nula en ningún intervalo.

#### 6.4 Sistemas discretos

Un sistema es en general, un dispositivo con una o más entradas y una o más salidas. En nuestro caso estaremos interesados en *sistemas dinámicos discretos* que realizan alguna modificación de las señales de entrada; por ejemplo, filtrarlas.

Consideremos el caso de un *sistema lineal e invariante en el tiempo* (LTI) con una entrada y una salida. Al ser lineal, dadas dos señales de entrada  $x_1(k)$  y  $x_2(k)$  que producen salidas  $y_1(k)$  e  $y_2(k)$  y dos constantes  $a$  y  $b$ , entonces la salida ante  $ax_1(k) + bx_2(k)$  será  $ay_1(k) + by_2(k)$ . Al ser invariante en el tiempo, si ante una entrada  $x(k)$  se tiene una salida  $y(k)$  entonces ante una entrada desplazada en el tiempo  $x(k - n)$  la salida tendrá el mismo desplazamiento:  $y(k - n)$ .

Este sistema puede caracterizarse a través de su *respuesta al impulso* que designaremos como  $h(k)$ . Si aplicamos la linealidad e invariancia en el tiempo a la señal  $x(k)$  expresada como en (6.4), la respuesta  $y(k)$  puede expresarse como

$$y(k) = \sum_{m=-\infty}^{\infty} x(m) h(k - m) \quad (6.5)$$

Se trata de una superposición de respuestas al impulso ponderadas y desplazadas en el tiempo. Es decir, la respuesta se expresa como el producto de convolución entre la entrada y la respuesta al impulso. Con la notación de producto de convolución:

$$y(k) = x * h(k) \quad (6.6)$$

Haciendo un sencillo cambio de variable en (6.5) es fácil ver que el producto de convolución es conmutativo.

$$y(k) = h * x(k) = \sum_{m=-\infty}^{\infty} h(m) x(k - m). \quad (6.7)$$

Por ejemplo, supongamos la siguiente respuesta al impulso

$$h(k) = \begin{cases} 0 & k < 0 \\ Ka^{-k} & k \geq 0 \end{cases} \quad (6.8)$$

Entonces la respuesta a cualquier  $x(k)$  señal será

$$y(k) = \sum_{m=0}^{\infty} Ka^{-m} x(k - m). \quad (6.9)$$

Esta expresión da la respuesta en forma de serie pero tiene el inconveniente de que para obtener la salida en un solo instante se requiere conocer la totalidad de la respuesta al impulso y la totalidad de la señal de entrada. Esto no es conveniente, dado que en general no conocemos de antemano toda la señal. Particularmente no sería aplicable a un sistema de tiempo real con un flujo permanente de señal de entrada.<sup>96</sup>

<sup>96</sup> Sería aplicable, en teoría, para una señal que ha sido grabada

### 6.5 Sistemas con respuesta al impulso finita e infinita

Los sistemas discretos pueden tener una respuesta al impulso infinitamente larga o, por el contrario, dicha respuesta puede tener longitud finita  $M + 1$ . En el primer caso se denominan sistemas de *respuesta al impulso infinita*, IIR (*infinite impulse response*). En el segundo caso, son los sistemas de *respuesta al impulso finita*, FIR (*finite impulse response*). En este último caso, la ecuación (6.7) puede escribirse como

$$y(k) = \sum_{m=0}^M h(m) x(k-m). \quad (6.10)$$

A diferencia del caso general, en este caso es posible el cálculo de la respuesta ya que involucra sólo un número finito de valores de la señal de entrada.

### 6.6 Ecuación diferencia de un sistema discreto

Con un pequeño artificio se puede simplificar el cálculo de la respuesta de un sistema IIR obteniendo una expresión recursiva. Consiste en intentar obtener la respuesta en el tiempo discreto  $k$  a partir de la respuesta en el tiempo  $k - 1$ :

$$y(k-1) = \sum_{m=0}^{\infty} K a^{-m} x(k-1-m) = a \sum_{m=0}^{\infty} K a^{-(m+1)} x(k-(m+1)). \quad (6.11)$$

Llamando  $n = m + 1$ , resulta

$$y(k-1) = \sum_{n=1}^{\infty} K a^{-n} x(k-n) = a \left( \sum_{n=0}^{\infty} K a^{-n} x(k-n) - K x(k) \right). \quad (6.12)$$

es decir

$$y(k-1) = a(y(k) - Kx(k)).$$

de donde

$$y(k) = \frac{1}{a} y(k-1) + Kx(k) \quad (6.13)$$

Esta ecuación permite calcular los sucesivos valores de  $y(k)$  a partir del valor anterior y del valor actual de la entrada. Partiendo de un valor inicial, por ejemplo  $y(0) = 0$  es posible determinar uno por uno todos los sucesivos valores, sin necesidad de conocer de antemano más que un solo valor de entrada por vez. En particular, dada una señal discreta que pueda describirse analíticamente, podemos obtener una respuesta explícita. Por ejemplo, si  $x(k)$  es un escalón unitario en 0, tenemos

$$\begin{aligned} y(1) &= K \\ y(2) &= \frac{K}{a} + K = K(a^{-1} + 1) \end{aligned}$$

$$\begin{aligned}
y(3) &= \frac{K(a^{-1} + 1)}{a} + K = K(a^{-2} + a^{-1} + 1) \\
&\vdots \\
y(k) &= K \sum_{n=0}^{k-1} a^{-n} = K \frac{1 - a^{-k}}{1 - a^{-1}} = \frac{Ka}{a-1} (1 - a^{-k})
\end{aligned} \tag{6.14}$$

que es una constante menos una exponencial decreciente (para  $a > 1$ ).

La ecuación (6.13) se denomina *ecuación diferencia* del sistema discreto, y es análoga a la ecuación diferencial de un sistema continuo. De hecho, podemos ver que en realidad es una aproximación de la ecuación diferencial. Para ello tengamos en cuenta que la derivada puede estimarse como el cociente incremental para incrementos pequeños. Si en (6.13) explicitamos el tiempo, podemos obtener un cociente incremental restando  $y(k-1)$  y dividiendo por el período de muestreo:

$$\frac{y(kT_s) - y((k-1)T_s)}{T_s} = \frac{1}{T_s} \left( \frac{1}{a} - 1 \right) y(k-1) + \frac{K}{T_s} x(k) \tag{6.15}$$

Esta ecuación es una aproximación de ésta:

$$y'(t) + \frac{a-1}{T_s a} y(t) = \frac{K}{T_s} x(t) \tag{6.16}$$

Es una ecuación de primer orden cuya respuesta al impulso es

$$h(t) = \frac{K}{T_s} e^{-\frac{a-1}{aT_s} t} \tag{6.17}$$

Si bien este resultado no se parece a la respuesta al impulso de la que partimos,  $h(t) = K a^{-k}$ , cuando la caída es lenta, es decir  $a \cong 1$ , resulta

$$e^{-\frac{a-1}{a} t} \cong 1 - \frac{a-1}{a} t = \frac{1}{a}$$

$$\frac{t}{T_s} \cong k$$

En cuanto a la constante multiplicativa, el factor  $1/T_s$  es justamente el que eliminamos al normalizar a 1 la amplitud del pulso unitario.

Igualmente, la aproximación de la derivada utilizada en (6.15) no es demasiado buena para señales rápidamente variables. Después veremos que se obtienen mejores aproximaciones a las ecuaciones diferenciales originales si en lugar de aproximar la derivada se expresa la ecuación en su forma integral y se aproxima la integral.

Podemos generalizar la ecuación (6.13) planteando una ecuación donde la salida dependa de  $n$  salidas anteriores y  $m$  entradas anteriores:

$$y(k) = b_0x(k) + b_1x(k-1) + \dots + b_Mx(k-M) - a_1y(k-1) - \dots - a_Ny(k-N) \quad (6.18)$$

Se dice que este sistema es de orden  $N$  ya que depende de  $N$  valores anteriores de la propia salida. Igual que antes, es posible obtener recursivamente la expresión de la respuesta. Sin embargo, hay métodos más eficientes, que veremos más adelante.

Si en (6.18) fuera  $a_n = 0$  para todo  $n = 1, \dots, N$ ,

$$y(k) = b_0x(k) + b_1x(k-1) + \dots + b_Mx(k-M), \quad (6.19)$$

es decir, la salida actual  $y(k)$  no depende de las salidas pasadas sino sólo de  $M$  entradas pasadas, estamos en presencia de un sistema FIR, es decir, con respuesta al impulso de duración finita. Si, por ejemplo,  $b_0 = 1$ ,  $b_1 = 0,5$ ,  $b_2 = 0,25$  y para el resto  $b_m = 0$ , entonces la respuesta al impulso será

$$\begin{aligned} h(0) &= b_01 + b_10 + b_20 = 1 \\ h(1) &= b_00 + b_11 + b_20 = 0,5 \\ h(2) &= b_00 + b_10 + b_21 = 0,25 \\ h(k) &= b_00 + b_10 + b_20 = 0 \quad k \neq 0, 1, 2 \end{aligned}$$

Podemos afirmar que, en general, la respuesta al impulso de un sistema (6.19) está dada por

$$h(k) = \begin{cases} b_k & k = 0, 1, \dots, M \\ 0 & k > M \end{cases} \quad (6.20)$$

## 6.7 Respuesta en frecuencia de un sistema discreto

Supongamos un sistema discreto con respuesta al impulso  $h(k)$ . Utilizaremos como señal de ensayo una señal de la forma

$$x(k) = e^{j k \Omega} \quad (6.21)$$

donde  $\Omega$  es la frecuencia angular normalizada:

$$\Omega = 2\pi \frac{f}{F_s} \quad (6.22)$$

Es decir  $\Omega = \pi$  corresponde a la frecuencia de Nyquist, es decir  $F_s/2$ .<sup>97</sup> La exponencial compleja es equivalente a una señal senoidal, sólo que es analíticamente más simple de trabajar. La respuesta del sistema será la convolución con  $h(k)$ :

<sup>97</sup> Algunos autores usan  $\omega$  para la frecuencia normalizada. Aquí usamos  $\Omega$  para evitar incoherencias de notación.

$$y(k) = h * e^{jk\Omega} = \sum_{m=-\infty}^{\infty} h(m) e^{j(k-m)\Omega} = e^{jk\Omega} \sum_{m=-\infty}^{\infty} h(m) e^{-jm\Omega},$$

es decir,

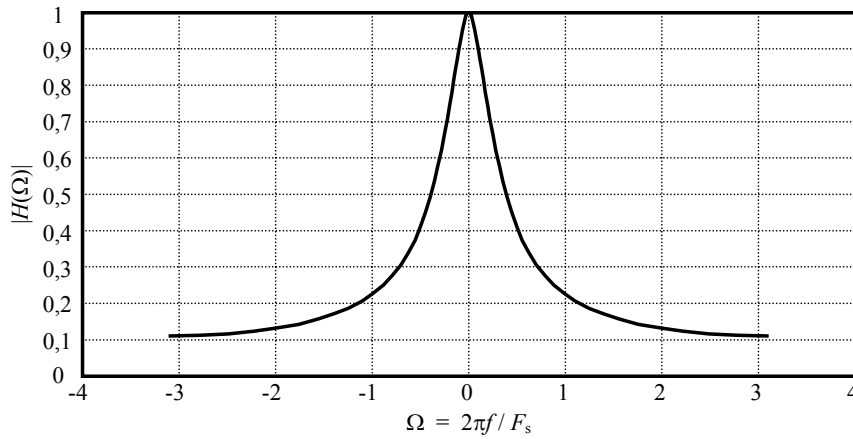
$$y(k) = x(k) \underbrace{\sum_{m=-\infty}^{\infty} h(m) e^{-jm\Omega}}_{\text{Respuesta en frecuencia}}. \quad (6.23)$$

La respuesta en frecuencia coincide con la transformada de Fourier discreta de la respuesta al impulso. Notar que la transformada de Fourier de una señal discreta es una función de variable continua (ya que  $\omega$  varía en forma continua) y es periódica con período  $2\pi$ . Sin embargo las frecuencias superiores a  $\pi$  (que equivale a  $F_s/2$ ) representan alias.

A modo de ejemplo, calculemos la respuesta en frecuencia para el sistema de la ecuación (6.8),

$$\sum_{m=-\infty}^{\infty} h(m) e^{-jm\Omega} = \sum_{m=0}^{\infty} K a^{-m} e^{-jm\Omega} = \frac{K}{1 - a^{-1} e^{-j\Omega}}. \quad (6.24)$$

En la figura 85 se muestra esta respuesta para  $a = 1,25$  y  $K = 0,2$  (el valor de  $K$  se eligió para tener respuesta 1 en continua).

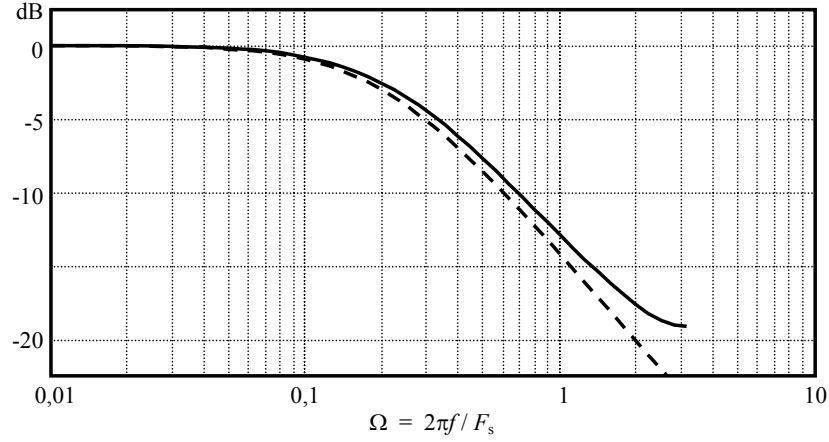


**Figura 85.** Respuesta en frecuencia bilateral del sistema del ejemplo de la ecuación (6.24).

En la figura 86 se representa la respuesta en escala logarítmica de frecuencia y en dB, junto con la de un filtro pasabajos continuo de primer orden de la misma frecuencia de corte, obtenida por comparación con la ecuación diferencial equivalente, (6.16):

$$f_c = \frac{1}{2\pi\tau} = \frac{a-1}{2\pi a} F_s \quad (6.25)$$

Vemos que para baja frecuencia, donde la derivada puede aproximarse por el cociente incremental y por lo tanto la ecuación de diferencia es asimilable a la ecuación diferencial, las respuestas son similares, pero a alta frecuencia, sobre todo en proximidades de la frecuencia de Nyquist  $F_s/2$ , las respuestas difieren bastante.



**Figura 86.** Respuesta en frecuencia en escala logarítmica. En línea de trazos, la respuesta correspondiente al sistema continuo equivalente. .

## 6.8 Transformada $z$ de una señal discreta

Se define la *transformada  $z$*  de una señal discreta  $x(k)$  como

$$X(z) = \sum_{k=-\infty}^{\infty} x(k) z^{-k} \quad (6.26)$$

La transformada  $z$  desempeña el mismo papel para señales discretas que las transformadas de Fourier y Laplace para señales continuas. Mirando la ecuación (6.23) vemos que la respuesta en frecuencia de un sistema con respuesta impulsiva  $h(k)$  está dada por la transformada  $z$  de la respuesta impulsiva,  $H(z)$ , evaluada en  $z = e^{j\Omega}$ .

Como ejemplo podemos obtener algunas transformadas. Por ejemplo, para el pulso unitario, teniendo en cuenta que, salvo  $u_0(0)$ , todos los otros valores son 0,

$$U_0(z) = 1 z^{-0} = 1. \quad (6.27)$$

La transformada de un escalón unitario,  $u_1(k)$ , es

$$U_1(z) = \sum_{k=0}^{\infty} z^{-k} = \frac{1}{1 - z^{-1}} \quad (6.28)$$

La transformada de una exponencial para  $k > 0$ , es decir,  $g(k) = b^k u_0(k)$  resulta



$$G(z) = \sum_{k=0}^{\infty} b^k z^{-k} = \frac{1}{1 - b z^{-1}}. \quad (6.29)$$

### 6.9 Transformada $z$ de una ecuación diferencia

Veamos ahora que se puede resolver una ecuación diferencia en el dominio frecuencial por medio de la transformada  $z$ .

Como primer ejemplo aplicaremos la transformada  $z$  a ambos miembros de la ecuación diferencia (6.13), pero para ello necesitamos saber cómo se transforma la señal retardada  $y(k-1)$ .

$$\begin{aligned} Z\{y(k-1)\}(z) &= \sum_{k=-\infty}^{\infty} y(k-1) z^{-k} = \\ &= \sum_{k=-\infty}^{\infty} y(k-1) z^{-(k-1+1)} = \\ &= z^{-1} \sum_{k=-\infty}^{\infty} y(k-1) z^{-(k-1)} = z^{-1} Y(z) \end{aligned} \quad (6.30)$$

Es decir, un retardo de una muestra se traduce en la multiplicación de la transformada  $z$  por  $z^{-1}$ . La transformada de (6.13) resulta, entonces

$$Y(z) = \frac{1}{a} z^{-1} Y(z) + KX(z). \quad (6.31)$$

Podemos obtener la *función de transferencia en  $z$*  como

$$\frac{Y(z)}{X(z)} = \frac{K}{1 - \frac{z^{-1}}{a}}. \quad (6.32)$$

Si comparamos con la ecuación (6.24) que daba la respuesta en frecuencia del mismo sistema, vemos que equivale a reemplazar  $z = e^{j\Omega}$ . Esto es análogo a lo que sucede en un sistema continuo, donde reemplazando  $s = j\omega$  en la función de transferencia de Laplace se obtiene la respuesta en frecuencia.

Podemos generalizar lo anterior aplicando transformada  $z$  a la ecuación diferencia general (6.18). Resulta

$$\begin{aligned} Y(z) &= (b_0 + b_1 z^{-1} + \dots + b_M z^{-M}) X(z) - \\ &\quad - (a_1 z^{-1} + \dots + a_N z^{-N}) Y(z) \end{aligned} \quad (6.33)$$

es decir

$$Y(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + \dots + a_N z^{-N}} X(z) \quad (6.34)$$

de donde resulta la *función de transferencia*

$$\frac{Y(z)}{X(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + \dots + a_N z^{-N}} \quad (6.35)$$

Un caso particular de la ecuación (6.35) es cuando la entrada es un impulso discreto, en cuyo caso la salida es la respuesta al impulso. En ese caso  $X(z) = 1$  e  $Y(z) = H(z)$ , de donde resulta que la función de transferencia es la transformada de la respuesta al impulso:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{1 + a_1 z^{-1} + \dots + a_N z^{-N}} \quad (6.36)$$

De la ecuación (6.34) resulta, entonces, que la transformada  $z$  de la salida es el producto de las transformadas  $z$  de la entrada y de la respuesta al impulso:

$$Y(z) = H(z) X(z) \quad (6.37)$$

### 6.10 Transformada $z$ de una convolución

La ecuación (6.5) expresa la salida de un sistema como la convolución entre la entrada y la respuesta al impulso y la (6.37) expresa una relación similar en el dominio transformado. Esta propiedad es general: la transformada  $z$  de una convolución  $u*v$  es el producto de las transformadas de  $u$  y  $v$ .<sup>98</sup> Par verlo apliquemos la definición:

$$Z\{u * v\} = \sum_{n=-\infty}^{\infty} \left( \sum_{m=-\infty}^{\infty} u(m)v(n-m) \right) z^{-n}$$

Ahora introducimos  $z^{-n}$  dentro de la segunda suma e invertimos el orden de la sumas

$$Z\{u * v\} = \sum_{m=-\infty}^{\infty} \left( \sum_{n=-\infty}^{\infty} u(m)v(n-m) z^{-n} \right)$$

Ahora sumamos y restamos  $m$  en el exponente de  $z$  y operamos:

$$Z\{u * v\} = \sum_{m=-\infty}^{\infty} \left( u(m) \sum_{n=-\infty}^{\infty} v(n-m) z^{-(n-m)} z^{-m} \right) \quad (6.38)$$

Reordenando y con un cambio de variable,

<sup>98</sup> Esta propiedad también se cumple para transformadas de Fourier y Laplace.

$$\begin{aligned}
Z\{u * v\} &= \sum_{m=-\infty}^{\infty} \left( u(m) z^{-m} \sum_{n=-\infty}^{\infty} v(n-m) z^{-(n-m)} \right) = \\
&= \sum_{m=-\infty}^{\infty} \left( u(m) z^{-m} Z\{v\} \right) = \\
&= \left( \sum_{m=-\infty}^{\infty} u(m) z^{-m} \right) Z\{v\} = Z\{u\} Z\{v\}
\end{aligned} \tag{6.39}$$

### 6.11 Transformada z y respuesta en frecuencia

La sumatoria de la ecuación (6.23) es igual a la transformada z de la respuesta al impulso calculada en  $e^{j\Omega}$ . Esto implica que la respuesta en frecuencia puede calcularse como  $H(e^{j\Omega})$ , donde  $-\pi < \Omega < \pi$  corresponde a  $-F_s/2 < f < F_s/2$ .

### 6.12 Soluciones de una ecuación diferencia

La resolución de una ecuación diferencia sigue los mismos lineamientos de la solución de las ecuaciones diferenciales en tiempo continuo. Para ello se descompone el problema en la obtención de una solución particular  $y_p$  y la obtención de la solución de la ecuación homogénea asociada (es decir, aquella en la que la entrada es nula),  $y_H$ . Por linealidad, vale:

$$y(k) = y_p(k) + y_H(k). \tag{6.40}$$

Para obtener la solución de la ecuación homogénea,

$$y(k) + a_1 y(k-1) + \dots + a_N y(k-N) = 0 \tag{6.41}$$

se supone que es de la forma

$$y_H(k) = \lambda^k. \tag{6.42}$$

Sustituyendo, resulta

$$\lambda^k + a_1 \lambda^{k-1} + \dots + a_N \lambda^{k-N} = 0,$$

que puede reescribirse, sacando  $\lambda^{k-N}$  como factor común,

$$(\lambda^N + a_1 \lambda^{N-1} + \dots + a_N) \lambda^{k-N} = 0.$$

Esta ecuación se cumple para todo  $k$  si y sólo si

$$\lambda^N + a_1 \lambda^{N-1} + \dots + a_N = 0. \tag{6.43}$$

Esta ecuación admite  $N$  soluciones complejas que son las raíces  $\lambda_n$  del polinomio del primer miembro, denominado *polinomio característico* del sistema. Si los coeficientes

son reales, como es generalmente el caso, las raíces son reales o están agrupadas de a pares conjugados. Si no hay raíces múltiples, la solución general de la ecuación homogénea es una combinación lineal de las soluciones (6.42):

$$y_H(k) = c_1 \lambda_1^k + c_2 \lambda_2^k + \dots + c_N \lambda_N^k. \quad (6.44)$$

En cuanto a la solución particular, si la entrada admite formas analíticas simples, el método es suponer que la respuesta tiene una forma específica relacionada con la forma de la entrada y tratar de obtener los parámetros. En la tabla 6.1 se muestran unos pocos casos.

**Tabla 6.1.** Forma de la solución particular de una ecuación diferencia para algunas señales analíticas típicas.

Entrada		Solución particular
Tipo	Expresión	
Constante	$A$	$B$
Exponencial	$Aa^k$	$Ba^k$
Potencia	$Ak^b$	$B_0 k^b + B_1 k^{b-1} + \dots + B_b$
Senoide	$A \sin(\Omega k + \varphi)$	$B \sin(\Omega k + \psi)$

A modo de ejemplo, obtengamos la solución particular para la ecuación diferencia (6.13) ante una rampa

$$x(k) = k.$$

Según la tabla 6.1, la solución particular es de la forma

$$y_P(k) = B_0 k + B_1.$$

Reemplazamos en (6.13) y reordenamos

$$((a - 1)k + 1)B_0 + (a - 1)B_1 = aKk$$

Ahora evaluamos para  $k = 0$  y  $k = 1$ , obteniendo el siguiente sistema de ecuaciones:

$$\begin{cases} B_0 + (a - 1)B_1 = 0 \\ aB_0 + (a - 1)B_1 = aK \end{cases}$$

Resulta, finalmente,

$$y_P(k) = \frac{aK}{a - 1}k - \frac{a^2 K}{a - 1} \quad (6.45)$$

A esta solución se agregará la solución general de la homogénea, que se obtiene resolviendo

$$\lambda - \frac{1}{a} = 0$$

Resulta  $\lambda = a^{-1}$ , por lo que la solución homogénea será

$$y_H(k) = c a^{-k}$$

que coincide con la forma de la respuesta impulsiva.

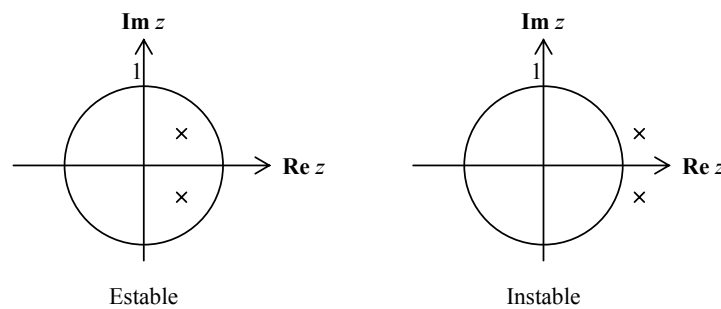
### 6.13 Polos y estabilidad de un sistema discreto

Se dice que un sistema cualquiera es *estable* si, partiendo del reposo, luego de una entrada de duración finita el sistema retorna asintóticamente al estado de reposo. Para que ello suceda basta con que las soluciones homogéneas tiendan a 0. Las soluciones homogéneas son de la forma (6.42) donde  $\lambda$  es cualquiera de las raíces  $\lambda_k$  de la ecuación característica (6.43), que son también las raíces del denominador de la función de transferencia  $z$ , (6.36), es decir, son sus *polos* (valores de  $z$  donde la función se hace  $\infty$ ).

Vemos que para que ello suceda, debe cumplirse, para todas las raíces  $\lambda_k$ , que

$$|\lambda_k| < 1. \quad (6.46)$$

Esto sucede si los polos de la función de transferencia se encuentran dentro del círculo de radio unitario centrado en 0. Esta situación se muestra en la figura 87



**Figura 87.** Ubicación de un par de polos conjugados de la función de transferencia  $z$  de un sistema discreto estable y uno inestable.

### 6.14 Inversión de una transformada $z$ racional

Un método alternativo para obtener la respuesta de un sistema ante una entrada consiste en aplicar la ecuación (6.37), multiplicando la transformada  $z$  de la entrada por la función de transferencia (que coincide con la transformada de la respuesta al impulso). Esto nos proporciona la transformada  $z$  de la respuesta. A partir de ella podemos obtener la respuesta mediante la inversión de la transformada  $z$ . Si bien existe una fórmula teórica para la transformada  $z$  inversa, en general no se aplica y en su lugar se uti-

lizan métodos alternativos aplicables a casos especiales. Uno de ellos es cuando la transformada  $z$  es racional, es decir, de la forma  $P(z)/Q(z)$ , donde  $P$  y  $Q$  son polinomios.

En este caso se puede utilizar una descomposición en fracciones simples. Para el caso frecuente en que las raíces no son coincidentes, la descomposición es

$$\frac{P(z)}{Q(z)} = \sum_{m=1}^N \frac{A_m z}{z - p_m} \quad (6.47)$$

donde  $p_m$  son los polos complejos, es decir los ceros del denominador  $Q(z)$ . Los términos en (6.47) tienen la particularidad de que su transformada inversa es inmediata ya que si recordamos la ecuación (6.29) vemos que la forma es similar a la transformada  $z$  de una exponencial. Podemos reescribir

$$\frac{P(z)}{Q(z)} = \sum_{m=1}^N \frac{A_m}{1 - p_m z^{-1}} \quad (6.48)$$

por lo que su transformada inversa es

$$Z\left\{\frac{P(z)}{Q(z)}\right\}(k) = \sum_{m=1}^N A_m p_m^k u_0(k). \quad (6.49)$$

En muchos casos el principal interés de trabajar con transformadas  $z$  está vinculado con la información sobre las propiedades espectrales de la señal de salida o la respuesta en frecuencia del sistema, más que con la obtención de la forma explícita de la respuesta.

Por otra parte cuando están involucradas señales de la realidad, es imposible una representación analítica que permita conocer la transformada  $z$  de la señal de entrada, por lo cual sólo podemos tener la esperanza de conocer una representación del espectro promedio a largo plazo, el cual nos permite sacar conclusiones interesantes sobre el espectro promedio a largo plazo de la respuesta..

### 6.15 Conversión bilineal de un sistema continuo en discreto

Supongamos que tenemos un sistema continuo con función de transferencia  $H(s)$ , y que deseamos obtener un sistema discreto equivalente. Aunque así planteado, el problema no tiene una solución general, bajo ciertas condiciones existe una transformación de la variable  $s$  de Lapace en la variable  $z$  razonablemente satisfactoria.

Vimos anteriormente que la aproximación de una derivada mediante el cociente incremental puede ser demasiado grosera en frecuencias altas (cercanas a la frecuencia de Nyquist). En su lugar podemos pensar, a la inversa, que una señal es la integral de su derivada. Si  $u(t) = y'(t)$ , entonces se cumple que

$$y(kT_s) = \int_{(k-1)T_s}^{kT_s} u(t) dt + y((k-1)T_s) \quad (6.50)$$

Podemos aproximar la integral en un intervalo tan corto como  $T_s$  usando la regla del trapecio:

$$\int_{(k-1)T_s}^{kT_s} u(t) dt \cong \frac{u(kT_s) + u((k-1)T_s)}{2} T_s, \quad (6.51)$$

de donde, pasando a las señales discretas asociadas,

$$y(k) \cong \frac{u(k) + u(k-1)}{2} T_s + y(k-1).$$

Ahora transformemos  $z$  y reordenemos:

$$Y(z) (1 - z^{-1}) \cong \frac{1 + z^{-1}}{2} T_s U(z).$$

Resulta

$$U(z) \cong \frac{2}{T_s} \frac{1 - z^{-1}}{1 + z^{-1}} Y(z). \quad (6.52)$$

Pero  $U(z)$  es la transformada  $z$  de la derivada de  $y$ , por lo que el operador multiplicativo entre transformadas  $z$

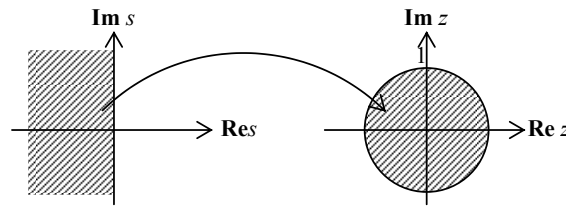
$$\frac{2}{T_s} \frac{1 - z^{-1}}{1 + z^{-1}}$$

es equivalente al operador derivación en el tiempo continuo. Esto sugiere que la transformación

$$s \rightarrow \frac{2}{T_s} \frac{1 - z^{-1}}{1 + z^{-1}} = \frac{2}{T_s} \frac{z - 1}{z + 1} \quad (6.53)$$

permite convertir funciones de transferencia de Laplace en funciones de transferencia  $z$  entre las señales de entrada y salida muestreadas.

La transformación (6.53) es la *transformación bilineal*. Se puede demostrar que transforma el semiplano real negativo, donde se ubican los polos estables de un sistema continuo en el interior del círculo unitario, donde se ubican los polos estables del sistema discreto, como se muestra en la figura 88.



**Figura 88.** Transformación del semiplano real negativo en el interior del círculo unitario mediante la transformación bilineal.

## Capítulo 7

# ALGORITMOS BÁSICOS PARA MEDICIONES ACÚSTICAS

### 7.1 Introducción

En este capítulo trataremos varios de los algoritmos básicos a utilizar en mediciones acústicas basadas en software. Utilizaremos una clase de software matemático de tipo matricial, es decir, donde las operaciones con vectores y matrices con números reales y complejos se realizan con naturalidad y simplicidad como si se tratara de números. Este tipo de software es muy interesante para procesamiento digital de señales, ya que muchos cálculos requieren realizar el mismo proceso sobre cantidades masivas de datos estructurados como vectores o matrices.

Otra característica de interés es la posibilidad de trabajar en forma simbólica, lo que permite trabajar con expresiones algebraicas, reducirlas, desarrollarlas, obtener soluciones explícitas para ecuaciones, derivadas, integrales, etc.

Hay varios programas de este tipo, entre ellos Matlab, Scilab, Octave y Euler. El más difundido es Matlab, y si bien es el que mayor desarrollo posee, es un software de distribución comercial muy costoso. Las otras tres alternativas son compatibles con la Licencia General Pública, GPL (General Public License), de la Free Software Foundation FSF). Si bien la licencia de Scilab no es GPL sino CeCILL, la propia FSF lo declara compatible con la licencia GPL. En todos los casos los códigos fuente en C o C++ están disponibles libremente para su descarga, estudio, modificación e incluso distribución, siempre y cuando se mantenga la compatibilidad con GPL que, entre otras cosas, requiere que cualquier versión derivada deberá distribuirse bajo la misma licencia,

Cualquiera de estos programas podría utilizarse, pero hemos seleccionado Scilab ya que tiene una interfaz muy desarrollada y similar a la de Matlab, por lo cual el trabajo de programación se facilitará para aquellos lectores acostumbrados a Matlab. En el apéndice 13 se incluye una breve introducción a las características de Scilab.

Al igual que en el caso del editor de audio digital Audacity, la decisión de utilizar software libre se basa en la filosofía de promover la mayor difusión posible del conocimiento aplicado y de lograr, a través de un trabajo cooperativo, la mejora permanente. Una de las razones que lo hacen viable es la facilidad y economía del proceso de distribución, realizado a través del uso de redes de datos basadas en Internet o a través de copias de bajo costo en CD, así como la posibilidad de recurrir a una extensa red de usuarios y desarrolladores para efectuar consultas, sugerir mejoras y obtener soporte técnico.

### 7.2 Carga de un archivo .wav

El primer paso es adquirir la señal mediante la función **wavread**. Supongamos que el archivo está ubicado en `c:/sonidos/sonido.wav`. Entonces

```
[x, Fs, Nbits] = wavread('c:/sonidos/sonido.wav');  
N = length(x);
```



copia la señal a una variable  $\mathbf{x}$  que es vector columna en el caso de un archivo monofónico o una matriz de dos columnas en el caso de un archivo estereofónico. La señal está normalizada entre  $-1$  y  $1$ . Con esta sintaxis<sup>99</sup> también se obtiene la cantidad de bits por muestra y la tasa de muestreo en las variables **Nbits** y **Fs**.

En caso de que el archivo sea muy largo, para evitar problemas de saturación de la memoria, convendrá cargar una parte del archivo. Así, por ejemplo

```
N1 = 100000;  
N2 = 1000000;  
[x, Fs, Nbits] = wavread('c:/sonidos/sonido.wav', [N1,N2]);
```

leerá desde la muestra **N1** hasta la **N2**.

### 7.3 Promedio energético y nivel equivalente

El proceso más básico es la obtención del promedio energético o *valor eficaz*. El valor eficaz de una señal continua  $x(t)$  se define como

$$X_{ef} = \sqrt{\frac{1}{T} \int_0^T x^2(t) dt} . \quad (7.1)$$

En el caso de una señal discreta  $x(n)$ , dada por  $N$  muestras, la integración anterior se realiza numéricamente mediante

$$X_{ef} = \sqrt{\frac{1}{NT_s} \sum_{n=1}^N x^2(n) T_s} = \sqrt{\frac{1}{N} \sum_{n=1}^N x^2(n)} . \quad (7.2)$$

Podemos observar que si bien la aproximación numérica de la integral se plantea utilizando el tiempo de muestreo como tiempo incremental, finalmente se simplifica. El cálculo puede realizarse en Scilab mediante la instrucción

```
Xef = sqrt(sum(x.^2)/N);
```

Nótese como la potenciación por componentes,  $\mathbf{x}.\wedge 2$ , y la suma de componentes permitieron un código muy compacto y eficiente.

### 7.4 Calibración

La calibración consiste en relacionar los valores arbitrarios que tiene una magnitud transducida y digitalizada con el valor original, de modo que del conocimiento de los valores digitalizados puedan reconstruirse los valores originales de la magnitud. En nuestro caso existen varias constantes de proporcionalidad: la sensibilidad del micrófono,  $S_m$ , la ganancia del acondicionador de señal del sonómetro,  $G_a$ , la del preamplificador del dispositivo digitalizador,  $G_d$ , la constante del conversor analógico digital,  $K_{A/D} = 2^{N_{\text{bits}}-1}/V_{\text{ref}}$ , y la constante de normalización  $K_n = 2^{-N_{\text{bits}}}$  utilizada por Scilab. En consecuencia, el valor numérico adimensional  $x(n)$  con el que el software representa el valor de presión sonora en el instante  $t_n$  es

<sup>99</sup> Hay otras alternativas que se pueden investigar escribiendo “help wavread”.

$$x(n) = S_m G_a G_d K_{A/D} K_n p(t_n) = K p(t_n). \quad (7.3)$$

Como ya se comentó en la sección 4.2.9, en la práctica no se pueden controlar individualmente las constantes que forman  $K$

La forma más simple de obtener  $K$  es mediante la señal de calibración, cuya magnitud nominal,  $L_{cal}$ , es conocida (típicamente, 94 dB ó 114 dB). Con el procedimiento de la sección 7.2 se obtiene la señal normalizada,  $x_{cal}$ , y se determina su valor eficaz:

$$X_{cal} = \sqrt{\frac{1}{N_{cal}} \sum_{n=1}^{N_{cal}} x_{cal}^2(n)}. \quad (7.4)$$

La presión sonora correspondiente al nivel de presión sonora  $L_{cal}$  es

$$P_{cal} = P_{ref} 10^{L_{cal}/20}, \quad (7.5)$$

donde  $P_{ref} = 20 \mu\text{Pa}$ . Resulta

$$K = \frac{X_{cal}}{P_{ref} 10^{L_{cal}/20}}. \quad (7.6)$$

Si, como es habitual, pretendemos trabajar con niveles de presión sonora en lugar de presiones sonoras, entonces

$$L_p = 20 \log \frac{X_{ef}/K}{P_{ref}} = 20 \log X_{ef} - 20 \log X_{cal} + L_{cal}. \quad (7.7)$$

A efectos de implementar este cálculo hay dos opciones. La primera es disponer de un archivo separado para el tono de calibración, cargarlo como vector mediante la función **wavread** y luego determinar  $X_{cal}$  mediante el procedimiento antes indicado. La segunda es tener el tono de calibración grabado en el mismo archivo, utilizar software de edición de audio digital (Audacity) para localizarlo por simple inspección y determinar el número de muestra inicial y final, **Ncal1** y **Ncal2** para circunscribir el análisis a ese intervalo de tiempo.

Si a su vez **N1** y **N2** son las muestras inicial y final de la porción de señal cuyo nivel de presión sonora deseamos obtener, entonces el siguiente código

```
Lcal = 94;
Ncal = Ncal2 - Ncal1 + 1;
N = N2 - N1 + 1;
Xcal = sqrt(sum(x(Ncal1:Ncal2).^2)/Ncal);
Xef = sqrt(sum(x(N1:N2).^2)/N);
Lp = 20*log10(Xef/Xcal) + Lcal;
```

Obsérvese la sintaxis **x(N1:N2)** para denotar la subsecuencia de  $x$  comprendida entre los índices **N1** y **N2**.

### 7.5 Envoltente energética

La envoltente energética (ver figura 2, capítulo 1) se obtiene aplicando al cuadrado de la señal un filtro pasabajos con una constante de tiempo dada  $\tau$ . La función de transferencia en el dominio de Laplace es

$$H(s) = \frac{1}{1 + \tau s}. \quad (7.8)$$

A efectos del tratamiento digital debe obtenerse un filtro digital (en tiempo discreto) de igual comportamiento. La descripción de los filtros digitales se realiza mejor en el dominio de la transformada  $Z$ , por lo cual es necesario convertir la función transferencia en  $s$  en una función de transferencia en  $z$ . Teniendo en cuenta que las constantes de tiempo que se utilizan ( $\tau_F = 0,125$  s para respuesta rápida y  $\tau_S = 1$  s para respuesta lenta) corresponden a frecuencias de corte mucho menores que la tasa de muestreo, es válido aproximar la derivada por el cociente incremental entre muestras sucesivas. El filtro anterior corresponde a la ecuación diferencial

$$\tau y'(t) + y(t) = x(t)$$

que se aproxima como la siguiente ecuación diferencia:

$$\tau \frac{y(k) - y(k-1)}{T_s} + y(k) = x(k)$$

$$(1 + F_s \tau) y(k) - F_s y(k-1) = x(k)$$

Transformando  $z$  y operando resulta

$$H(z) = \frac{z}{(1 + F_s \tau)z - F_s \tau} \quad (7.9)$$

Esta expresión en la forma de cociente de dos polinomios en  $z$ ,  $B(z)/A(z)$  es apta para aplicar algoritmos de filtrado.

Para su implementación en Scilab primero debe declararse esta función de transferencia como sistema lineal por medio de la función **syslin**. En el código siguiente, **p** y **q** se declaran como los polinomios con raíces 0 y  $F_s \tau / (1 + F_s \tau)$  que integran el numerador y denominador.

```
Fs = 44100;
tau = .125;
p = poly(0, 'z');
q = (1 + Fs*tau) * poly(Fs*tau/(1 + Fs*tau), 'z');
H = syslin('d', p, q);
```

En este caso se ha tomado  $\tau = 0,125$  s, correspondiente a la respuesta rápida. Para respuesta lenta se debe tomar  $\tau = 1$  s. Una forma alternativa es utilizar la función **iir** para sintetizar un filtro pasabajos de Butterworth con frecuencia de corte  $1/(2\pi\tau)$ :

```
H = iir(1, 'lp', 'butt', [1/(2*pi*tau)/Fs, 0], [0,0]);
```

El paso siguiente es aplicar esta función de transferencia a la señal al cuadrado, lo cual se realiza mediante la función `flts`, y luego aplicar la raíz cuadrada:

```
y = sqrt(flts(x.^2, H));
```

Por último, se pueden representar gráficamente la señal  $\mathbf{x}$  y su envolvente energética  $\mathbf{y}$ . En el siguiente código se configura la ventana 1 como predeterminada y se grafican ambas señales mediante el comando `plot2d`:

```
t = [0:length(x)-1] / Fs;
xset('window',1);
plot2d(t, [x', y'])
```

En algunos casos se produce un desborde para las señales de gran duración (más de 15000 puntos). En ese caso se puede hacer un diezmado gráfico tomando una muestra cada  $n$ , donde  $n$  se elegirá convenientemente:

```
sub = 1:n:$;
plot2d(t(sub), [x(sub)', y(sub)'])
```

El diezmado implica un submuestreo que probablemente cause fenómenos de *aliasing* para la señal, ya que de hecho se estará dividiendo por  $n$  la tasa de muestreo efectiva, por lo que la máxima frecuencia admisible también se dividirá por  $n$ . En el caso de la envolvente energética es bastante seguro hacerlo sin riesgo de *aliasing* ya que el filtro utilizado tiene una frecuencia de corte muy baja (en el orden del Hz) por lo cual actúa como un efectivo filtro antialias.

El diezmado también podría llevarse a cabo mediante la función `intdec`, que se estudia más adelante (sección 8.3). Esta función aplica un filtro antialias antes del diezmado

## 7.6 Ponderación A

Las mediciones acústicas suelen hacerse con la interposición de filtros de ponderación, por ejemplo ponderación A o C. Según el modelo de sonómetro la salida de audio puede estar antes o después de dicho filtro. Sin embargo, para mayor versatilidad en el procesamiento posterior es conveniente grabar sin ponderación (ponderación plana). Si bien en teoría es posible eliminar la ponderación A con un filtro A inverso, en la práctica se presenta una dificultad seria, y es que la electrónica de cualquier sistema de grabación agrega inevitablemente ruido de baja frecuencia. Debido a que el filtro A inverso tiene elevada ganancia en baja frecuencia, podría agregar componentes que perturbarían la medición.

Es necesario, por consiguiente, disponer de un algoritmo que permita realizar la ponderación A. Ello puede hacerse implementando un filtro discreto con la misma respuesta en frecuencia. La función de transferencia de la respuesta A, dada por la norma IEC 61672, es

$$A(s) = \frac{1.2588 \cdot 76654,9^2 s^4}{(s + 129,43)^2 (s + 676,7) (s + 4636,36) (s + 76654,9)^2} \quad (7.10)$$

Existen dos enfoques para implementar el filtro. El primero utiliza estructura FIR (finite impulse response, o respuesta finita al impulso) y el segundo, IIR (infinite impulse response, o respuesta infinita al impulso). El primer enfoque utiliza una combinación lineal de  $P$  muestras anteriores de la entrada, donde  $P$  debe ser suficientemente grande como para permitir una respuesta transitoria de una duración comparable a la de la respuesta A. Debido a que dos de los polos son de frecuencia muy baja, siendo la constante de tiempo  $\tau = 1/129,43 = 7,7$  ms (el transitorio es de la forma  $t e^{-t/\tau}$ ), el valor de  $P$  requerido es bastante alto, lo cual hace muy lento el algoritmo.

El segundo enfoque es conceptualmente más complicado, ya que requiere no sólo los valores anteriores de la señal de entrada sino también de la salida (es un filtro recursivo) y podría presentar problemas de inestabilidad. Sin embargo, el tiempo de cómputo es mucho menor, ya que la combinación lineal requiere muchas menos operaciones. Es posible obtener una primera versión del filtro mediante una transformación bilineal de la forma

$$s = 2F_s \frac{1 - 1/z}{1 + 1/z} = 2F_s \frac{z - 1}{z + 1}. \quad (7.11)$$

El siguiente código diseña un filtro A. Para ello primero genera polinomios en  $s$  para el numerador y denominador de A y en  $z$  para el numerador y denominador de la transformación bilineal

```

Fs = 44100;
K = 1.2588*12200^2*(2*pi)^2;
p1 = -20.60*2*pi;
p2 = -107.7*2*pi;
p3 = -737.9*2*pi;
p4 = -12200*2*pi;

Anum = poly([ 0 0 0 0 K], 's', 'coeff' )
Aden = poly([p1 p1 p2 p3 p4 p4], 's')

Bnum = poly(1,'z')
Bden = poly(-1,'z')
B = 2*Fs*Bnum/Bden;

A = horner(Anum/Aden, B)

```

La función resultante es

$$\begin{array}{r}
 \begin{array}{ccccccc}
 & & & & 2 & & 3 & & 4 \\
 0.2572697 & + & 0.0014217z & - & 0.5131128z^2 & - & 0.0014122z^3 & + & 0.2558527z^4
 \end{array} \\
 \hline
 \begin{array}{ccccccc}
 & & & 2 & & 3 & & 4 \\
 0.0043459 & - & 0.1333542z & + & 1.1549654z^2 & - & 2.0246121z^3 & + & z^4
 \end{array}
 \end{array}$$

La transformación bilineal funciona muy bien en la región de las frecuencias centrales, pero introduce alteraciones de la respuesta en baja y alta frecuencia. Se pueden corregir éstas mediante el uso de predistorsión, es decir, modificando la forma de la función de transferencia de manera de compensar tales desviaciones. En este caso se han realizado correcciones empíricas, pero las mismas pueden depender del procesador utilizado, ya que los errores también dependen en alguna medida de la implementación de la aritmética:

```
As = 1.2588*12200^2 * (s/2/%pi)^4 / (s/2/%pi + 107.7);  
As = As / (s/2/%pi + 8)^2;  
As = As / (s/2/%pi + 737.9) / (s/2/%pi + 12200)^2;  
As = As * s / (s + 2*35.3*%pi) * (s/20000/2/%pi + 1);
```

### 7.7 Análisis estadístico

El análisis estadístico se utiliza para determinar la variabilidad del ruido, para aplicación de diversos criterios y para determinar el nivel de ruido de fondo. Consiste en determinar la distribución estadística de los valores de la envolvente energética, expresada en la forma de *niveles estadísticos*  $L_N$ , donde  $L_N$  es el nivel que es excedido un  $N\%$  del tiempo. Estadísticamente,  $L_N$  es igual al percentil  $100 - N$  de los valores de la envolvente energética expresados logarítmicamente en dB. La función **perctl** permite calcular los percentiles de las componentes de un vector indicados por un segundo vector.

Supongamos que en el vector **y** disponemos de la envolvente energética de **x**. Entonces calcularemos los percentiles 1 a 99, es decir los indicados por el vector **n** indicado a continuación:

```
n = [1:99];  
yn = perctl(y, 100 - n);  
yn = [max(y), yn, min(y)];  
Ln = 20*log10(yn/Xcal) + Lcal;
```

donde **Lcal** y **Xcal** son los valores definidos anteriormente (Sección 7.4). Para archivos muy grandes se requiere mucho espacio de memoria debido a que la ejecución de esta instrucción involucra la operación de ordenamiento (**gsort**), por lo que conviene submuestrear **y**. Ello puede hacerse sin problemas de aliasing debido a que el filtro utilizado para extraer la envolvente energética actúa como efectivo filtro antialias. La segunda instrucción se reemplazará, entonces, por

```
yn = perctl(y(1:100:$), 100 - n);
```

Es conveniente, también, antes de la ejecución incrementar el tamaño de la pila (*stack*) de memoria mediante la instrucción **stacksize**, por ejemplo:

```
stacksize(10000000);
```

Se puede modificar permanentemente esta variable editando la entrada correspondiente en el archivo **scilab.star** ubicado en el directorio donde está instalado Scilab:

```
newstacksize = 10000000;
```

El método anterior tiene el inconveniente de que, incluso ampliando la pila de memoria, la misma no alcance. Conviene en ese caso subdividir el problema realizando cálculos sobre bloques con cantidades moderadas de muestras. Sin embargo, no se puede obtener directamente a partir de los percentiles en dos bloques sucesivos los percentiles correspondientes a la suma de ambos bloques. El procedimiento consistirá en trabajar con histogramas, que sí permiten la combinación. Lamentablemente la función **histplot** de Scilab sólo dibuja el histograma, sin que sea posible obtener los datos

numéricos, por lo cual deberemos generar un código propio para obtener el histograma de cada bloque.

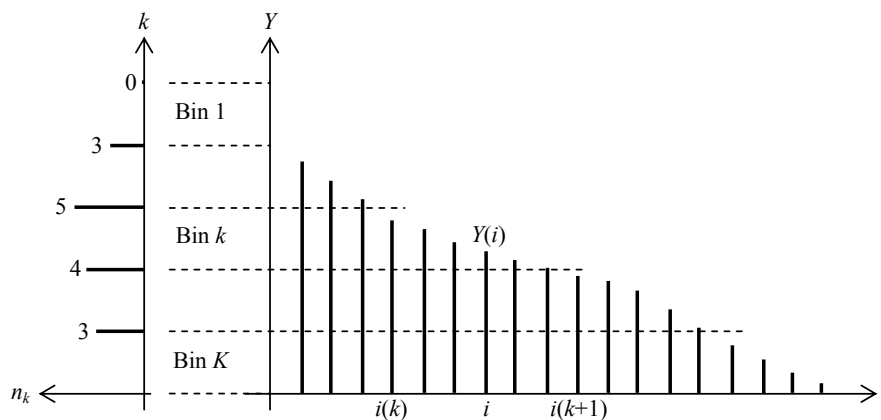
Para ello, primero se ordena de mayor a menor el vector  $\mathbf{y}$  que contiene la envolvente energética submuestreada, mediante la función **gsort**, obteniéndose  $\mathbf{Y}$ . Luego se fija un rango de valores posibles de la variable  $\mathbf{y}$ . Esta elección dependerá del problema particular, por ejemplo en ruido urbano podría comprender entre 20 dB y 120 dB, valores que una vez llevados a la referencia digital supondremos que son -100 dB y 0 dB respectivamente:

```
mini= -100
maxi= 0
```

A continuación se divide dicho rango en una cantidad  $\mathbf{K}$  de casilleros (*bins*), dentro de cada uno de los cuales habrá que contar la cantidad de valores de  $\mathbf{Y}$ .

```
K = 1000;
delta = (maxi - mini)/K;
rango = [maxi:-delta:mini];
```

La cantidad  $\mathbf{K}$  de casilleros, y por lo tanto su ancho **delta**, dependerán de la cantidad de datos disponibles. Cuanto más datos (por ejemplo 1 min de audio, que corresponde a 26460 muestras submuestreadas), más casilleros se pueden tomar por lo que el ancho del casillero será menor. Conviene tomar una cantidad de casilleros lo suficientemente grande como para representar correctamente los rasgos más importantes de la distribución, pero no tanto como para representar hechos irrelevantes.



**Figura 89.** Determinación del histograma de una variable para ser utilizado en el cálculo por bloques de percentiles.

El algoritmo básico para obtener el histograma consiste en encontrar, para cada casillero en orden descendente, la primera muestra que sea menor que el límite superior del casillero. Para ello utilizamos la función **find**, que determina los índices donde se verifica una dada condición lógica. De todos los índices estamos interesados en el me-

nor. Así siguiendo, se arma un vector que esencialmente contiene los índices frontera entre los casilleros. Dado que hay un índice por cada muestra de la envolvente, el último paso es restar los índices de casilleros consecutivos, ya que la resta indica cuántas muestras están dentro del casillero.

El código que realiza esto es el siguiente, donde las dos primeras líneas sólo están para reservar espacio a las variables:<sup>100</sup>

```
i = zeros(1, length(rango));
histo_Y = zeros(1, length(rango));

for k=1: length(rango)
    i(k) = min(find(Y < rango(k))) - 1;
    if k>1
        histo_y(k) = i(k) - i(k-1);
    end
end
```

Es posible lograr mejoras en la eficiencia computacional si en lugar de hacer la búsqueda extendida a todo el vector **Y** se realiza sólo en lo que resta. También se puede sacar la resta de índices fuera del bucle **for** y ejecutarla en bloque mediante la función **diff**. Queda para el lector investigar estas posibilidades, recordando que los comandos **tic** y **toc** colocados antes y después respectivamente de una porción de código informan el tiempo que demoró la ejecución de dicha porción.

Una vez calculado el histograma de dos o más bloques (cada bloque abarcará, por ejemplo, un minuto de audio), podemos obtener el histograma general simplemente sumando los histogramas componente por componente (es decir, se añaden a cada casillero la cantidad de muestras correspondiente al casillero homólogo del nuevo bloque). De hecho se puede tener una variable vectorial que contenga el histograma acumulado a la que se vayan agregando los nuevos espectrogramas que surjan..

Finalmente dispondremos del histograma correspondiente a una gran cantidad de audio (por ejemplo, 30 minutos). El siguiente paso consiste en estimar la distribución de probabilidad acumulada. La función **cumsum** permite obtener la distribución acumulada de frecuencias de aparición de muestras en cada casillero. Normalizando por la suma total de frecuencias, que es la última componente de la distribución de frecuencias acumulada,<sup>101</sup> se obtiene la función de densidad de probabilidad acumulada:

```
f_y = cumsum(histo_y);
f_y = f/f($);
```

Luego de inicializar la variable que va a contener los percentiles,

```
perc_y = zeros(1, 101);
```

---

<sup>100</sup> La p reasignación de espacio para una variable cuando se sabe el tamaño máximo que va a tener es conveniente, ya que se evita el riesgo de que el software tenga que mover toda la variable a otra parte de la memoria cuando la variable ha crecido tanto que ya no queda espacio contiguo diusable. Esto mejora la eficiencia de los algoritmos.

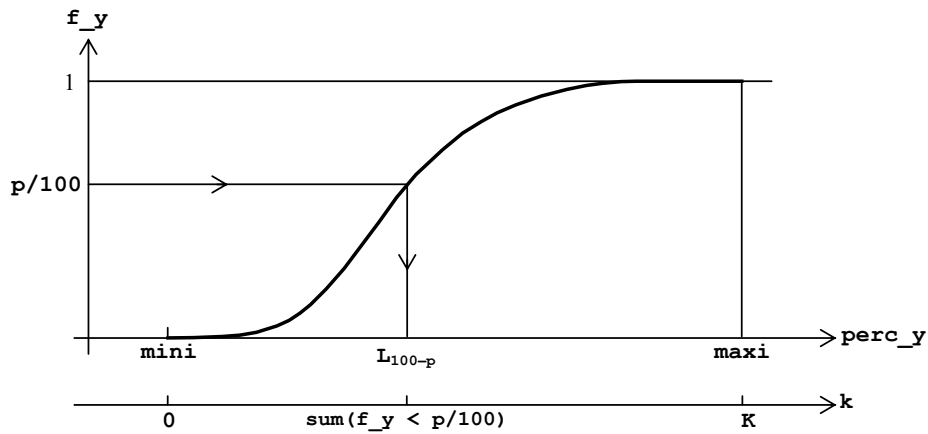
<sup>101</sup> Scilab cuenta con el comodín **\$** para indicar la última componente de un vector o de una dimensión de una matriz



sigue el fragmento de código que realiza el cálculo propiamente dicho. Se trata aquí básicamente de obtener la función inversa de la probabilidad acumulada  $f_y$ . Hay varios enfoques posibles, por ejemplo interpolar linealmente o con funciones *spline*. El método aquí usado es simplemente localizar para qué valores del índice del cual depende  $f_y$  esta función es menor que  $p/100$ . La cantidad de índices para los que ello sucede equivale al índice del valor de  $f_y$  más próximo a  $p/100$ . El índice debe convertirse en un valor de la variable que este índice representa, es decir, los valores de  $y$  correspondientes a cada casillero.

```
for p=0:100
    perc_y(p+1) = mini + delta * sum(f_y < p/100);
end
```

En la figura 90 se muestra esquemáticamente este procedimiento. Observemos que el nivel estadístico  $N$  corresponde al percentil  $p = 100 - N$ . Ello se debe a que el nivel estadístico se define como aquél valor que es *superado* un  $N\%$  del tiempo, mientras que el percentil  $p$  es el que *supera* a un  $p\%$  de los valores de la variable.



**Figura 90.** Determinación del percentil  $p$  a partir de la función de densidad de probabilidad acumulada. El eje paralelo es el eje de los índices. Cada índice corresponde a un incremento **delta** de la variable.

## Capítulo 8

# ANÁLISIS DE ESPECTRO

### 8.1 Introducción

En la actualidad existen diversos instrumentos de medición disponibles comercialmente capaces de realizar la medición del espectro sonoro en bandas de octava y tercio de octava de acuerdo con lo estipulado en la norma IEC 61260. Los mismos son, todavía, excesivamente costosos pese a que el costo de la electrónica digital de alta fiabilidad es cada vez menor. La tendencia moderna es fabricar instrumentos con un sistema de cómputo programable embebido<sup>102</sup> para el cual se proveen módulos de software que permiten no sólo extender las prestaciones y los tipos de análisis disponibles, sino además una actualización de la funcionalidad y prestaciones sin cambiar el hardware. Esta facilidad, que permite amortizar los costos de diseño del hardware, contrariamente a lo que podría presuponerse no ha traído aparejada una reducción de precios sino, por el contrario, un incremento a través de la elevada cotización de cada módulo.

Una alternativa viable es utilizar un instrumento de costo y prestaciones mínimas capaz de proveer una señal calibrada a un grabador digital en formato PCM lineal (no comprimido). La señal será analizada ulteriormente por computadora (o, eventualmente, con un sistema DSP dedicado de bajo costo) admitiendo una enorme variedad de tipos de análisis.

### 8.2 Paradigmas para el análisis de espectro

El análisis espectral se realizaba, inicialmente, mediante un filtro analógico de frecuencia variable en forma continua y ancho de banda constante (aunque configurable) o mediante un banco de filtros analógicos de frecuencias discretas distribuidas uniformemente en una escala logarítmica y ancho de banda proporcional a la frecuencia. Los primeros, denominados analizadores de *ancho de banda constante*, se aplican al análisis de señales con tonos puros. Los segundos, designados como de *porcentaje constante*, tienen mayor aplicación en el caso de señales de espectro continuo o mixto.<sup>103</sup>

Estos filtros se implementaban mediante circuitos eléctricos, al principio pasivos (con capacitores e inductancias) y luego activos (con amplificadores, resistencias y capacitores). Eran de gran tamaño, peso y costo, lo cual dificultaba su traslado para las mediciones de campo, razón por la que su uso estaba mayormente restringido al laboratorio. Más adelante el tamaño se fue reduciendo de la mano de la miniaturización de la electrónica, lo cual supuso una generalización de su uso, con impacto positivo en la investigación de fenómenos acústicos dependientes de la frecuencia.

También se desarrollaron sistemas de alta calidad para grabación en cinta magnética, lo cual posibilitaba el registro de señales calibradas para su análisis ulterior en el laboratorio con los analizadores mencionados.

---

<sup>102</sup> Un sistema embebido es un sistema de cómputo de cierta complejidad integrado a un equipo de propósito específico. Sin embargo, a medida que la tecnología evoluciona, la frontera entre una computadora de propósito general y un sistema embebido se va desdibujando

<sup>103</sup> Se recomienda repasar la sección 1.7 sobre conceptos generales de análisis de espectro.

Con el advenimiento de los sistemas de cómputo digital fue posible implementar algoritmos numéricos para el análisis de espectro. Surgieron así tres enfoques para el análisis de espectro de señales discretas. El primero consiste en la utilización de filtros de banda digitales que emulan los filtros de banda analógicos tradicionales como medio para separar las diferentes componentes espectrales de la señal a analizar. El segundo consiste en implementar el análisis de Fourier mediante la aplicación de una transformación del dominio del tiempo al dominio de la frecuencia. El tercero es un enfoque mixto que, utilizando técnicas de transformación directa e inversa, implementa filtros digitales que permiten un análisis espectral muy detallado.

### 8.3 Filtros digitales

El primer paradigma utiliza algoritmos basados en ecuaciones en diferencias del siguiente tipo,

$$y(k) = - \sum_{n=1}^N a_n y(k-n) + \sum_{m=0}^M b_m x(k-m) \quad (8.1)$$

que expresa el valor actual  $y(k)$  de la salida como combinación lineal de  $N$  valores pasados de la salida y del valor actual y  $M$  valores pasados de la entrada. Esta ecuación corresponde a una ecuación diferencial en el caso continuo. Si  $a_k = 0$  para  $k = 1, \dots, N$ , se tiene un filtro cuya respuesta al impulso es de duración finita (igual a  $M+1$  intervalos de muestreo). Se designa a este tipo de filtros como FIR (*finite impulse response*). Si, en cambio, para algún  $k$  es  $a_k \neq 0$  la respuesta al impulso es de duración infinita. Se los conoce como IIR (*infinite impulse response*). La función de transferencia  $z$  es

$$H(z) = \frac{b_0 z^M + b_1 z^{M-1} + \dots + b_{M-1} z + b_M}{z^N + a_1 z^{N-1} + \dots + a_{N-1} z + a_N} z^{N-M} \quad (8.2)$$

En los filtros obtenidos a partir de la función de transferencia de un filtro continuo en la variable  $s$  de Laplace mediante la transformación bilineal (6.53), resulta siempre  $M=N$ , es decir,

$$H(z) = \frac{b_0 z^N + b_1 z^{N-1} + \dots + b_{N-1} z + b_N}{z^N + a_1 z^{N-1} + \dots + a_{N-1} z + a_N} \quad (8.3)$$

Para el análisis de espectro estos filtros se diseñan de modo de emular las respuestas en frecuencia de los filtros analógicos prototipo, es decir respuestas pasabanda cuyas frecuencias de corte inferior y superior están dadas por

$$f_i = f_o 2^{-\frac{1}{2b}}, \quad f_s = f_o 2^{\frac{1}{2b}}, \quad (8.4)$$

donde  $f_o$  es la frecuencia central de la banda y  $b$  es el denominador de la fracción de octava ( $b = 1$  para banda de octava y  $b = 3$  para banda de tercio de octava).

Para diseñar un filtro de banda IIR se puede utilizar la función **iir** de Scilab. Esta función aplica una transformación bilineal (ver sección 6.15) al correspondiente prototipo en tiempo continuo. Admite dos sintaxis:

```
hz = iir(n, ftype, fdesign, frq, delta)
[p, z, g] = iir(n, ftype, fdesign, frq, delta)
```

donde:

**hz** es una estructura de función de transferencia discreta. Esta estructura tiene 4 componentes:

**hz(1)** contiene la descripción de la estructura: **!r num den dt !**

**hz(2)** contiene el polinomio numerador

**hz(3)** contiene el polinomio denominador

**hz(3)** indica que se trata de un filtro discreto (“c”: continuo; “d”: discreto)

**p** es un vector columna que contiene los polos de la transferencia discreta

**z** es un vector columna que contiene los ceros de la transferencia discreta

**g** es la ganancia requerida para tener ganancia 1 en la(s) banda(s) de paso

**n** es el orden del filtro;

**ftype** es el tipo de plantilla del filtro:

‘lp’: Pasabajos

‘hp’: Pasaaltos

‘bp’: Pasabanda

‘sp’: Rechazabanda

**fdesign** es el tipo de aproximación:

‘butt’: Butterworth (máximamente plano)

‘cheb1’: Chebychev tipo I (ripple constante en la banda de paso)

‘cheb2’: Chebychev tipo II (ripple constante en la banda de atenuación)

‘ellip’: Elíptico o Cauer (ripples constantes en ambas bandas)

**frq** es un vector de 2 componentes: frecuencia inferior de corte y frecuencia superior de corte, ambas normalizadas:  $0 \leq F \leq 0,5$ , donde 0,5 es equivalente a  $F_s/2$ .

NOTA: Esta normalización es necesaria dado que el filtro discreto no está asociado a ninguna tasa de muestreo en particular. Para un filtro pasabanda de frecuencias inferior y superior  $f_i$  y  $f_s$  este argumento será **[fi/Fs, fs/Fs]**;

**delta** es un vector de 2 componentes,  $\delta_1$  y  $\delta_2$ , aplicables a los filtros de Chebychev y Cauer, según corresponda, donde  $\delta_1$  es el ripple máximo en las bandas de paso y  $\delta_2$  es el ripple máximo en las bandas de atenuación. Notar que no están en decibels, sino en escala lineal.

Una vez obtenido el filtro se puede proceder a filtrar la señal mediante **filter**, utilizando la sintaxis

```
[y, zf] = filter(num, den, x [,zi])
```

donde:

**y** es la salida filtrada

**zf** es el estado o condición final, consistente en los últimos  $\max\{N, M\}$  valores de la respuesta. Es útil cuando se va a realizar un filtrado por bloques, por ejemplo para evitar problemas de saturación de memoria

- num** es el numerador del filtro, expresado como coeficientes en orden decreciente de potencia o bien como polinomio. En esta última variante puede utilizarse la variable **hz (2)** obtenida con la primera sintaxis de la función **iir**
- den** es el denominador del filtro, expresado como coeficientes en orden decreciente de potencia o bien como polinomio. En esta última variante puede utilizarse la variable **hz (3)** obtenida con la primera sintaxis de la función **iir**
- x** es la señal de entrada, expresada como vector fila
- zi** es el estado o condición inicial, consistente en los  $\max(N, M)$  valores de la respuesta previos al primer valor. Puede utilizarse haciéndolo igual a **zf** del filtrado de un bloque inmediatamente anterior.

Es necesario tener en cuenta que para frecuencias centrales mucho menores que  $F_s$  este tipo de filtros se vuelve inestable. La razón es que en un filtro IIR, por ser recursivo, se arrastran los errores de truncado de la aritmética. Por más que la representación tenga 16 dígitos decimales, luego de algunos millones de operaciones estos errores tienden a acumularse. Los filtros con  $f_0 \ll F_s$  tienen sus polos muy cerca de la circunferencia unitaria. Por ejemplo, un filtro de Butterworth de banda de tercio de octava centrado en 100 Hz, con una tasa de muestreo de 44100 Hz, tiene los siguientes polos:

$$\begin{aligned} p_{1,2} &= 0,9994220 \pm 0,0129037i \\ p_{3,4} &= 0,9986432 \pm 0,0135627i \\ p_{5,6} &= 0,9992838 \pm 0,0155332i \\ p_{5,6} &= 0,9985264 \pm 0,0146502i \end{aligned}$$

cuyos módulos son, respectivamente, 0,9995053, 0,9995053, 0,9994045 y 0,9986338. Si bien estos polos son teóricamente estables, ya que se encuentran dentro del círculo unitario (región de estabilidad de los filtros discretos), la acumulación de errores lleva a una respuesta divergente.

La solución a este problema consiste en submuestrear la señal de modo que los polos no estén marginalmente cerca del límite de estabilidad. Ello puede lograrse con la función **intdec**. Tengamos en cuenta que antes de submuestrear es necesario realizar un filtrado antialias. Esta función lleva a cabo dicha operación mediante un filtro digital antialias muy eficiente.

La sintaxis es la siguiente:

**xsub = intdec (x,lom)**

donde

- xsub** es la versión submuestreada de **x**
- x** es el vector que contiene la señal a submuestrear
- lom** es el factor por el que se multiplica la tasa de muestreo original para obtener la nueva, es decir:  $F_{s2}/F_{s1}$

Un criterio para evitar la inestabilidad es que la tasa de muestreo  $F_s$  resulte a lo sumo 7 veces la frecuencia central  $f_0$  del filtro. Debe tenerse en cuenta que cuando se submuestra para cumplir con esto, las frecuencias mayores que  $3,5f_0$  se eliminan. Aunque parecería que ello no es importante, podría serlo en algún caso particular, ya que la salida de un filtro de banda continuo sigue teniendo respuesta no nula (aunque

pequeña) más allá de su frecuencia de corte, que con el proceso de submuestreo desaparecería. Desde el punto de vista normativo ello no es importante porque fuera de la banda de paso la Norma IEC 61260 da un límite inferior de atenuación de  $-\infty$  dB.<sup>104</sup> Pero si nos interesara investigar la señal residual para señales fuera de la banda de paso estaríamos en problemas.

El siguiente código realiza el submuestreo, pero sólo si es necesario. Obsérvese que se realiza un submuestreo por una potencia de 2 ya que es más eficiente.

```
if fo < Fs/7
    p = 1;
    q = 2^ceil(log2(Fs/fo/7));
    xx = intdec(x, p/q);
end
Fs2 = Fs/q;
```

Las siguientes líneas definen las frecuencias normalizadas inferior y superior de corte de la banda y obtienen el diseño de un filtro pasabanda de orden 3 de Butterworth

```
Finf = 2^(-1/6) * fo/Fs2;
Fsup = 2^( 1/6) * fo/Fs2;
hz = iir(n, 'bp', 'butt', [Finf Fsup], []);
```

A continuación se aplica el filtro a la señal submuestreada. En este caso no hacen falta los argumentos opcionales **zi** y **zf**:

```
yy = filter(num, den, xx);
```

Finalmente, si fuera importante contar con la señal filtrada, se puede sobremuestrear:

```
if fo > Fs/7
    y = intdec(x, q/p);
end
```

Nótese que al sobremuestrear la longitud final de la señal filtrada podría diferir de la original. Si esto fuera importante se pueden agregar ceros (*zero-padding*) para

Si el único objetivo es obtener el nivel de la banda de tercio de octava, este último paso no será necesario, pudiéndose aplicar a la señal filtrada lo visto en las secciones 7.3 y 7.4.

La aplicación reiterada de este procedimiento a todas las bandas permite obtener el espectro de bandas. Si, una vez obtenidas las señales se las segmenta y se analizan porciones menores, se puede obtener la evolución del espectro de bandas en el tiempo.

Es posible también diseñar filtros FIR pasabanda mediante la instrucción **wfir** que obtengan la señal dentro de una banda. Sin embargo, particularmente en baja frecuencia la longitud de la respuesta puede ser considerable, por lo que computacionalmente es menos eficiente.

<sup>104</sup> La Norma IEC 61260 especifica las características de los filtros de banda, en particular la tolerancia de la respuesta.

#### 8.4 Transformada discreta de Fourier (DFT)

Supongamos una señal periódica  $x(t)$ . Podemos representarla mediante una serie trigonométrica o *serie de Fourier* compleja:

$$x(t) = \sum_{k=-\infty}^{\infty} c_n e^{jn\omega t} \quad (8.5)$$

donde  $\omega = 2\pi/T$ , y los *coeficientes de Fourier*,  $c_n$ , están dados por

$$c_n = \frac{1}{T} \int_0^T x(t) e^{-jn\omega t} dt \quad (8.6)$$

Si la señal está acotada en banda, es decir no tiene contenido espectral más allá de cierto valor de frecuencia  $f_{\text{máx}}$ , entonces (8.5) puede reescribirse con una cantidad finita  $N$  de términos

$$x(t) = \sum_{n=-N/2}^{N/2-1} c_n e^{jn\omega t} \quad (8.7)$$

donde  $N = 2f_{\text{máx}} T$ .<sup>105</sup> Se verifica que la anterior puede expresarse también como

$$x(t) = \sum_{n=0}^{N-1} c_n e^{jn\omega t} \quad (8.8)$$

donde, para  $n = N/2, \dots, N-1$ ,

$$c_n = c_{n-N} \quad (8.9)$$

Es un caso de aliasing temporal, donde debido la periodicidad de la exponencial compleja, diferentes espectros dan origen a la misma señal temporal.

Supongamos ahora que partimos de una señal muestreada en los instantes  $kT/N$  para  $k = 0, \dots, N-1$ . Dado que está acotada en banda, la señal puede representarse mediante la ecuación anterior particularizada a los instantes  $kT/N$ :

$$x\left(k \frac{T}{N}\right) = \sum_{n=0}^{N-1} c_n e^{jn \frac{2\pi}{T} k \frac{T}{N}} = \sum_{n=0}^{N-1} c_n e^{j2\pi \frac{nk}{N}} \quad (8.10)$$

Dado que esto vale para  $k = 0, \dots, N-1$ , éste es un sistema de  $N$  ecuaciones cuyas  $N$  incógnitas son los coeficientes de Fourier  $c_1, \dots, c_N$ . Para resolverlo procedemos a multiplicar por una exponencial similar

<sup>105</sup> Notemos que si  $x(t)$  es periódica de período  $T$ , la máxima frecuencia tiene que ser un múltiplo exacto de  $1/T$

$$x\left(k \frac{T}{N}\right) e^{-j2\pi \frac{mk}{N}} = \sum_{n=0}^{N-1} c_n e^{j2\pi \frac{(n-m)k}{N}}$$

y sumamos respecto a  $k$

$$\begin{aligned} \sum_{k=0}^{N-1} x\left(k \frac{T}{N}\right) e^{-j2\pi \frac{mk}{N}} &= \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} c_n e^{j2\pi \frac{(n-m)k}{N}} = \\ &= \sum_{n=0}^{N-1} c_n \sum_{k=0}^{N-1} e^{j2\pi \frac{(n-m)k}{N}} \end{aligned} \quad (8.11)$$

Si  $n = m$  las exponenciales son todas 1 y por lo tanto la sumatoria interna resulta igual a  $N$ . Si  $n \neq m$ , la sumatoria interna es de la forma  $\sum a^n$ , cuyo valor es, para  $a \neq 1$ ,

$$\sum_{k=0}^{N-1} a^k = \frac{1 - a^N}{1 - a} = \frac{1 - e^{j2\pi(n-m)k}}{1 - e^{j2\pi \frac{(n-m)k}{N}}}$$

Observemos que en la exponencial de arriba se canceló el  $N$  que dividía al exponente, por lo tanto la exponencial es 1 y la sumatoria es 0.

En consecuencia, de todos los términos de (8.11) el único diferente de 0 es aquel para el que  $n = m$ :

$$\sum_{k=0}^{N-1} x\left(k \frac{T}{N}\right) e^{-j2\pi \frac{mk}{N}} = N c_m, \quad (8.12)$$

es decir

$$c_m = \frac{1}{N} \sum_{k=0}^{N-1} x\left(k \frac{T}{N}\right) e^{-j2\pi \frac{mk}{N}}, \quad (8.13)$$

Es interesante observar que, reordenando un poco, (8.13) puede verse como una aproximación numérica de la integral de (8.6):

$$c_m = \frac{1}{T} \sum_{k=0}^{N-1} x\left(k \frac{T}{N}\right) e^{-j2\pi \frac{mk}{N}} \frac{T}{N}, \quad (8.14)$$

donde  $T/N$  representa el ancho de un pequeño intervalo  $\Delta t$ . Lo interesante, y en cierta forma sorprendente, es que esta “aproximación” resulte exacta.

Dado que en (8.13) la escala temporal (dada por el factor  $T/N$ ) no es relevante, podemos omitirla, preservando sólo el valor de las sucesivas muestras.



Podemos, entonces, llamar *transformada discreta de Fourier* (DFT) de la sucesión  $x_1, \dots, x_N$  a la sucesión  $X_1, \dots, X_N$  dada por

$$X(m) = \sum_{k=0}^{N-1} x(k) e^{-j2\pi \frac{mk}{N}}. \quad (8.15)$$

La señal es, así, la *transformada discreta de Fourier inversa* (IDFT) de  $X_1, \dots, X_N$ ,

$$x(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{j2\pi \frac{mk}{N}}. \quad (8.16)$$

Nótese que los coeficientes de Fourier  $c_n$  no coinciden con la transformada discreta. Si se desea obtener dichos coeficientes, es necesario dividir por  $N$ :

$$c_m = \frac{1}{N} X(m), \quad (8.17)$$

Algunos autores definen la transformada discreta de Fourier incluyendo el factor  $1/N$  (Randall et al., 1978) Aunque esto parece tener más sentido, ya que proporciona directamente los coeficientes de Fourier de la señal continua que le da origen, aquí utilizaremos esta definición debido a que los programas matemáticos de cálculo matricial como Octave, Scilab o Matlab usan la definición (8.15) para la transformada discreta de Fourier y la (8.16) para su inversa.

### 8.5 Transformada rápida de Fourier (FFT)

La transformada discreta de Fourier resulta una herramienta muy útil no sólo para el análisis espectral sino para muchas otras aplicaciones. Sin embargo, especialmente para  $N$  grande se vuelve computacionalmente prohibitiva. En efecto, el cálculo de cada componente de la DFT requiere el cálculo de  $N$  multiplicaciones y  $N$  sumas, por lo tanto el cálculo de todas las componentes insumirá  $N^2$  multiplicaciones y  $N^2$  sumas. Para una transformada de  $N=4096$ , ello implica más de 33 millones de operaciones de punto flotante. Un examen atento de (8.15) muestra que hay muchos cálculos repetidos y que, por lo tanto podrían omitirse. En efecto, si bien el exponente de  $e^{-j2\pi/N}$  varía entre 1 y  $N^2$ , en realidad sólo asume  $N$  valores diferentes.

En 1965, Cooley y Tuckey encontraron una forma de reorganizar los cálculos para aprovechar esta circunstancia cuando la cantidad de muestras es una potencia de 2, es decir.  $N=2^n$ . Su algoritmo, denominado *transformada rápida de Fourier*, *FFT*, requiere tan sólo  $\frac{3}{2}N \log_2 N$  operaciones de punto flotante, lo cual para el caso anterior supone reducir en un factor de 455 la cantidad de operaciones. En el Apéndice 14 se da una introducción a este algoritmo.

La introducción de la FFT tuvo un enorme impacto en el desarrollo de muchas técnicas de procesamiento y análisis de señales digitales. Destaquemos, no obstante, que en realidad se trata de aplicaciones de la DFT, sólo que hechas posibles gracias a la eficiencia computacional de la FFT. En lo que sigue la utilizaremos para el análisis de espectro.

En Scilab, la transformada rápida de Fourier se realiza mediante el código

```
y = fft(x, -1);
```

donde  $\mathbf{x}$  es una variable previamente creada. El valor  $-1$  indica el signo del exponente de las exponenciales que figuran en la definición de la DFT. La transformada inversa se logra con

$$\mathbf{x} = \text{fft}(\mathbf{y}, 1);$$

Ambas funciones se realizan más eficientemente cuando el número de muestras es una potencia de 2. Si  $n$  o  $l$  no lo es puede tomarse una porción de señal donde  $N$  sea potencia de 2:

$$\mathbf{y} = \text{fft}(\mathbf{x}(\mathbf{k}:\mathbf{k}+\mathbf{N}-1), -1);$$

donde además se ejemplifica la toma de muestras que comienzan en la muestra  $\mathbf{k}$ , no necesariamente al comienzo de la señal.

También es posible, y muchas veces muy útil, realizar varias transformadas al mismo tiempo, por ejemplo en varias porciones de señal ubicadas como columnas de una matriz  $\mathbf{x}$ :

$$\mathbf{y} = \text{fft}(\mathbf{x}, -1, 1);$$

o bien a lo largo de sus filas:

$$\mathbf{y} = \text{fft}(\mathbf{x}, -1, 2);$$

La función tiene otras opciones que el lector podrá investigar en la página de ayuda.

## 8.6 Análisis de espectro mediante la FFT

Antes de introducir las técnicas específicas basadas en la FFT discutiremos algunas de sus limitaciones. En primer lugar, al decidir usar la FFT para el análisis de espectro de una porción de la señal estamos aceptando implícitamente un modelo periódico de período  $N$  para esa señal, ya que de la ecuación (8.16) se desprende que  $x(k)$  es periódica. Sin embargo, la señal podría no ser periódica, como se muestra en la figura 91, o podría ser periódica con un período diferente. Incluso podría darse el caso de que la señal continua que le dio origen fuera periódica con un período incommensurable con el período de muestreo por lo cual su versión discreta no sería periódica.

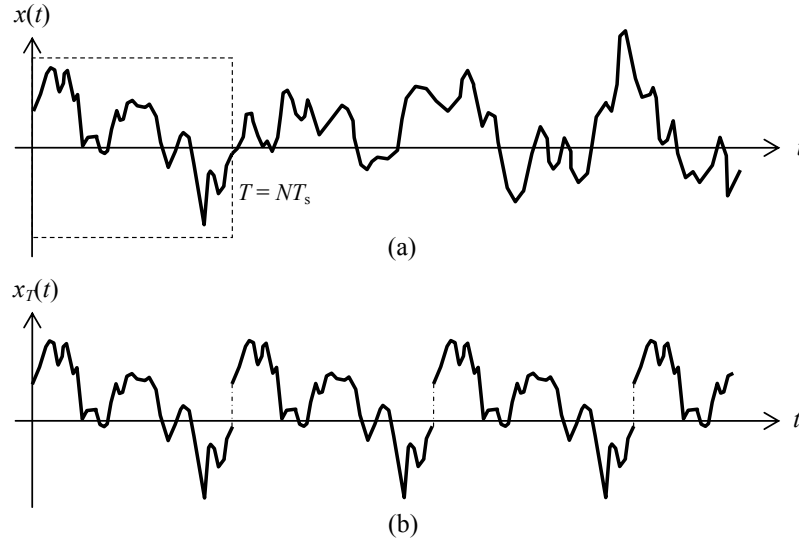
En cualquiera de los casos mencionados estamos suponiendo una estructura armónica determinada para  $x(k)$  que no concuerda con la real.

Otro problema se crea por las discontinuidades que implica la extensión periódica. Las señales periódicas representables con una serie de Fourier finita como la de la ecuación (8.8) tienen espectro acotado, por lo tanto son continuas. La discontinuidad crea frecuencias alias. Este problema se reducirá mediante el uso de ventanas.

### 8.6.1 Análisis de espectro de líneas

Las componentes espectrales que muestra la FFT según la ecuación (8.16) pueden interpretarse en términos del tiempo físico si en el exponente se multiplica y divide por  $T_s$ , se reordena y se tiene en cuenta que  $F_s = 1/T_s$ :

$$x(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{j2\pi \frac{mk}{N} \frac{T_s}{T_s}} = \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{j2\pi m \frac{F_s}{N} k T_s}. \quad (8.18)$$



**Figura 91.** (a) Evolución temporal de una señal. (b) Extensión periódica de la porción recuadrada en (a).

Resulta una serie de armónicos de frecuencias  $mF_s/N$ , de modo que la frecuencia fundamental es  $F_s/N$ , valor que también coincide con la separación entre las frecuencias de los armónicos,  $\Delta f$ , o *resolución espectral*:

$$\Delta f = \frac{F_s}{N}. \quad (8.19)$$

El período correspondiente a la frecuencia fundamental es

$$T = \frac{N}{F_s} = NT_s \quad (8.20)$$

Vemos, así, que para tener mejor resolución (reducir  $\Delta f$ ) es necesario aumentar la ventana de análisis (aumentar  $T$ ). Esta es una de tantas instancias del principio de indeterminación tiempo-frecuencia.

Ahora estamos interesados en determinar la energía de cada armónico. Tengamos en cuenta que para una señal física en la que  $x(k)$  es real, según se muestra en el apéndice 15, existe la siguiente relación de simetría conjugada entre los valores de la DFT:

$$X(N - n) = \overline{X(n)}. \quad (8.21)$$

Esto implica que las correspondientes exponenciales se suman para dar un coseno. Las amplitudes de las componentes del espectro real son:

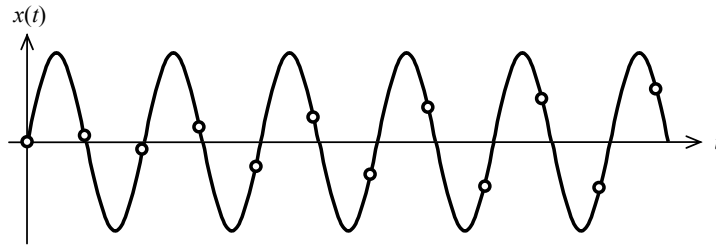
$$X_r(n) = \begin{cases} \frac{X(0)}{N} & n = 0 \\ \frac{2|X(n)|}{N} & n = 1, \dots, N/2 - 1 \\ \frac{X(N/2)}{N} & n = N/2 \end{cases} \quad (8.22)$$

Para el cálculo energético la primera y la última aportan valores eficaces iguales a su amplitud (el primero por ser un valor de continua, el segundo por ser una señal que adopta sólo dos valores opuestos), en tanto que los  $N/2 - 1$  valores intermedios se consideran como tonos y aportan un valor eficaz  $\sqrt{2}$  veces menor.

Es importante resaltar una vez más que éstas no son las componentes presentes en la señal sino en el modelo periódico que se ha adoptado para la señal.

### 8.6.2 El problema de las frecuencias próximas a $F_s/2$

Cuando una señal tiene contenido de frecuencias muy cercanas a la frecuencia de Nyquist  $F_s/2$ , la cantidad de muestras que se recogen por ciclo es muy pequeña. Esto implica que si se muestrea durante un tiempo insuficiente dichas componentes de alta frecuencia podrían no quedar correctamente representadas. En la figura 92 se muestra un ejemplo. Como se puede observar, si se utilizaran las muestras indicadas, el valor eficaz de dicha componente se vería groseramente subestimado.



**Figura 92.** Tono puro de frecuencia ligeramente menor a la tasa de Nyquist,  $F_s/2$ .

Sin embargo, la señal cumple con los requisitos del teorema de muestreo y, a largo plazo, la señal se podría reconstruir perfectamente. Podemos establecer un criterio que permita garantizar que el muestreo no es sólo suficiente desde el punto de vista teórico sino también práctico. Para ello necesitaríamos contar con una cantidad suficiente de muestras como para que algunas de ellas alcancen valores cercanos a los de pico y, en lo posible, para que los efectos de tomar una cantidad no entera de semiperiodos no afecte demasiado deberíamos alcanzar varias veces el pico.

Observando la gráfica anterior es fácil presuponer que si seguimos tomando muestras en algún momento se alcanzarán valores cercanos al pico. Para que ello suceda deberemos alcanzar algún valor de tiempo discreto  $k_0$  tal que

$$\operatorname{sen}\left(2\pi f \frac{k_o}{F_s}\right) \cong 1. \quad (8.23)$$

Obsérvese que no es suficiente con pedir que el argumento del seno sea próximo a  $\pi/2$ , ya que incluso para  $k_o = 1$  es bastante mayor que  $\pi/2$ . La aproximación requerida se dará si logramos encontrar unos valores enteros  $k_o$  y  $m$  tales que

$$2\pi f \frac{k_o}{F_s} \cong 2\pi m + \frac{\pi}{2},$$

es decir,

$$k_o \frac{f}{F_s/2} \cong 2m + \frac{1}{2}.$$

La idea es separar en el primer miembro una parte entera que sea asimilable a  $2m$  y una parte fraccionaria que asimilaremos a  $1/2$ . Dado que  $f \cong F_s/2$  podemos escribir

$$k_o - k_o \left(1 - \frac{f}{F_s/2}\right) \cong 2m + \frac{1}{2} = 2m + 1 - \frac{1}{2}.$$

Bastará adoptar

$$\begin{cases} k_o \cong \frac{1}{2} \frac{F_s}{F_s - 2f} \\ 2m + 1 = k_o \end{cases} \quad (8.24)$$

En otras palabras, habrá que seleccionar el impar más próximo a la aproximación de arriba. Por ejemplo, si  $F_s = 44100$  Hz y  $f = 22049$  Hz,  $k_o$  deberá ser el impar más próximo a 11025, es decir,  $k_o = 11025$ . Se puede constatar que en este caso el seno alcanza la unidad en forma exacta.

A menos que pudiéramos garantizar que se cubre *exactamente* un cuarto de período (lo indispensable para lograr un cálculo preciso del valor eficaz), convendrá en realidad tomar

$$k \gg k_o. \quad (8.25)$$

En la práctica tal garantía no existe, ya que en primer lugar no conocemos a priori la frecuencia del tono.

Una forma alternativa de resolver el problema es no permitir que se alcancen frecuencias demasiado próximas a  $F_s/2$ . De (8.24), si pretendemos

$$k < k_{\max} \quad (8.26)$$

podemos adoptar

$$k_o = k_{\max}/10$$

con lo cual deberá ser

$$f < \frac{F_s}{2} \left( 1 - \frac{1}{2k_0} \right) = \frac{F_s}{2} \left( 1 - \frac{5}{k_{\text{máx}}} \right) \quad (8.27)$$

Por ejemplo, si pretendemos un cuadro de a lo sumo 4096 muestras, la frecuencia deberá ser menor de 22023 Hz. Como vemos, es lo bastante próxima a  $F_s$  como para que cualquier filtro antialias real la elimine casi totalmente.

Por otra parte, en la mayoría de las situaciones prácticas el ruido de alta frecuencia queda atenuado por multitud de fenómenos de tipo disipativo, por lo que rara vez este problema cobrará importancia.

### 8.6.3 El problema de las frecuencias subarmónicas

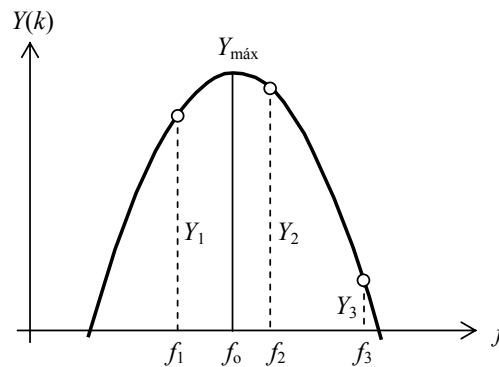
Un problema similar al anterior se da en el extremo de las bajas frecuencias, particularmente en las frecuencias subarmónicas, es decir, aquellas que están por debajo de la frecuencia fundamental, ya que si la ventana de análisis es demasiado corta, estaremos muestreando porciones de la onda poco representativas de la verdadera contribución al espectro.

El mismo tipo de criterio al seguido en la sección previa supone que la mínima frecuencia para la que se puede obtener un error razonable sería  $10F_s/N$ . Para una tasa de muestreo de 44100 Hz y una ventana de 4096 muestras corresponde a 107 Hz, que es una frecuencia bastante alta como para ignorarla. Sin embargo si admitimos un error de 0,5 dB, la ecuación (1.6) permite concluir que con dos ciclos alcanza, por lo cual la frecuencia puede reducirse a  $2F_s/N$ , esto es, unos 22 Hz.

### 8.6.4 Detección precisa de tonos puros

Cuando en una señal existen tonos puros, resulta de interés poder determinar las frecuencias y amplitudes correspondientes. Si un determinado tono coincidiera con una línea espectral de la transformada discreta de Fourier, es decir, si su frecuencia fuera un múltiplo de  $F_s/N$ , entonces salvo por algún otro contenido espectral, el espectro contendría una única y nítida línea asociada a esa frecuencia.

Si, en cambio, la frecuencia no coincidiera con ninguna línea espectral, aparecerán múltiples líneas asociadas, fenómeno conocido como *derrame espectral*. Es posible, sin embargo, estimar con alguna precisión la frecuencia correspondiente. El método consiste en seleccionar las tres líneas de mayor amplitud asociadas a un pico espectral y aplicar una interpolación parabólica, tal como se muestra en la figura 93.



**Figura 93.** Aproximación parabólica de la distribución de líneas espectrales correspondientes a un tono puro.

El pico se identifica por la presencia de un claro máximo relativo, en este caso  $Y_2$ . No debe confundirse con el valor máximo de la aproximación,  $Y_{\text{máx}}$ , que es inicialmente desconocido.

Planteamos el siguiente sistema de ecuaciones, cuyos datos conocidos son las frecuencias y las amplitudes de la FFT (en módulo) y cuyas incógnitas son los parámetros de la parábola, cuya forma hemos elegido para que  $f_0$  sea uno de ellos (de hecho, el único que nos interesa).

$$\begin{cases} Y_{\text{máx}} + K(f_1 - f_0)^2 = Y_1 \\ Y_{\text{máx}} + K(f_2 - f_0)^2 = Y_2 \\ Y_{\text{máx}} + K(f_3 - f_0)^2 = Y_3 \end{cases} \quad (8.28)$$

Restando de a pares eliminamos  $Y_{\text{máx}}$ , y dividiendo las dos ecuaciones resultantes eliminamos  $K$ :

$$\frac{K((f_2 - f_0)^2 - (f_1 - f_0)^2)}{K((f_3 - f_0)^2 - (f_1 - f_0)^2)} = \frac{Y_2 - Y_1}{Y_3 - Y_1}$$

Aunque parece una ecuación de segundo grado en  $f_0$ , en realidad los cuadrados se eliminan:

$$\begin{aligned} \frac{f_2^2 - 2f_2f_0 + f_0^2 - (f_1^2 - 2f_1f_0 + f_0^2)}{f_3^2 - 2f_3f_0 + f_0^2 - (f_1^2 - 2f_1f_0 + f_0^2)} &= \frac{Y_2 - Y_1}{Y_3 - Y_1} \\ \frac{f_2^2 - 2f_2f_0 - f_1^2 + 2f_1f_0}{f_3^2 - 2f_3f_0 - f_1^2 + 2f_1f_0} &= \frac{Y_2 - Y_1}{Y_3 - Y_1} \\ \frac{f_2^2 - f_1^2 - 2(f_2 - f_1)f_0}{f_3^2 - f_1^2 - 2(f_3 - f_1)f_0} &= \frac{Y_2 - Y_1}{Y_3 - Y_1} \end{aligned}$$

Eliminando denominadores

$$\begin{aligned} (f_2^2 - f_1^2)(Y_3 - Y_1) - (f_3^2 - f_1^2)(Y_2 - Y_1) &= \\ = 2((f_2 - f_1)(Y_3 - Y_1) - (f_3 - f_1)(Y_2 - Y_1))f_0 \end{aligned}$$

Por último

$$f_0 = \frac{(f_2^2 - f_1^2)(Y_3 - Y_1) - (f_3^2 - f_1^2)(Y_2 - Y_1)}{2((f_2 - f_1)(Y_3 - Y_1) - (f_3 - f_1)(Y_2 - Y_1))} \quad (8.29)$$

En cuanto a la amplitud, puede estimarse superponiendo las energías de las tres líneas más importantes:

$$Y_0 = \sqrt{Y_1^2 + Y_2^2 + Y_3^2} \quad (8.30)$$

### 8.6.5 Ventanas

Una cuestión importante en el uso de la FFT es que debido a discontinuidades de extremo (valores diferentes en el comienzo y el final del intervalo de análisis de  $N$  puntos) el espectro obtenido aparece contaminado por un énfasis en las altas frecuencias. En efecto, en el análisis de Fourier clásico de funciones periódicas se demuestra que los coeficientes tienden a cero tanto más rápidamente cuantas más derivadas continuas tenga la función. Así, si la propia función presenta discontinuidades, como en una onda cuadrada o una onda diente de sierra ideales, los coeficientes tienden a 0 con  $1/n$  cuando  $n \rightarrow \infty$ . Cuando la función es continua pero la derivada tiene discontinuidades (por ejemplo una onda triangular), los coeficientes tienden a 0 según  $1/n^2$ . Y, en general, si la derivada de orden  $m-1$  es continua pero la  $m$  es discontinua, tienden a 0 con  $1/n^{m+1}$ . Al seccionar una porción de la onda en forma arbitraria podemos estar creando discontinuidades que introducirán una alteración del espectro de alta frecuencia (ver figura 91).

La forma de resolverlo es multiplicar la señal por alguna de varias posibles *funciones ventana*,  $w(n)$ :

$$y(n) = w(n)x(n). \quad (8.31)$$

Éstas son funciones que son 0 (o muy cercanas a 0) en los extremos, y máximas en el centro. Por ejemplo, la ventana de Hann<sup>106</sup> está definida por

$$w_{\text{hann}}(n) = \frac{1}{2} \left( 1 - \cos \frac{2\pi n}{N-1} \right), \quad (8.32)$$

la de Blackman por

$$w_{\text{blackman}}(n) = 0,42 - 0,5 \cos \frac{2\pi n}{N-1} + 0,08 \cos \frac{4\pi n}{N-1}. \quad (8.33)$$

y la de Blackman-Harris por

$$w_{\text{blackman-harris}}(n) = 1 - 1,36 \cos \frac{2\pi n}{N-1} + 0,39 \cos \frac{4\pi n}{N-1} - 0,032 \cos \frac{6\pi n}{N-1}. \quad (8.34)$$

Por otra parte, la ventana de Hamming, definida por

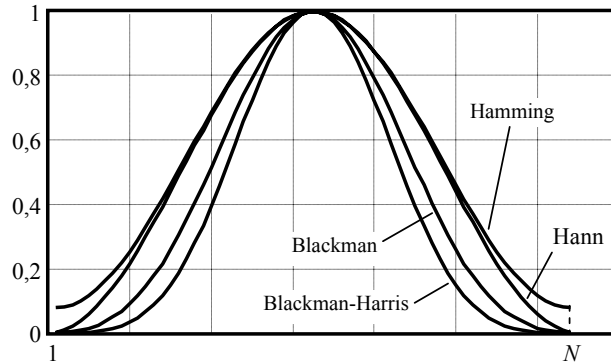
<sup>106</sup> Con frecuencia se la denomina, incorrectamente, ventana de *Hanning* cuando su autor realmente se llama Julius von Hann. Esto se debe probablemente a que existe otra ventana, la de Hamming cuyo nombre responde a su autor, Richard Wesley Hamming. La ortografía muy similar puede llevar a confusiones, por lo que es importante evitar la designación incorrecta.



$$w_{\text{hamming}}(n) = 0,54 - 0,46 \cos \frac{2\pi n}{N-1}. \quad (8.35)$$

es la única (salvo la ventana cuadrada que equivale a no usar ninguna ventana) que no anula completamente la señal en los extremos.

En la figura 94 se muestra la forma temporal que presenta cada una de estas ventanas.



**Figura 94.** Respuesta temporal de diversas funciones ventana para uso en el análisis de espectro por FFT.

El uso de estas ventanas tiene dos consecuencias. La primera es que la energía de la señal se ve alterada, ya que sólo en el punto medio del intervalo la ventana es igual a 1. Esto implica la necesidad de una corrección que puede depender del tipo de señal. La segunda es que cada componente senoidal pura de la señal original es reemplazada por varias componentes, es decir que la energía de un armónico tiene una tendencia a repartirse en varias frecuencias cercanas. Esto limita la capacidad de discriminar frecuencias muy próximas, de modo que la resolución de análisis empeora con respecto a su valor teórico  $F_s/N$ . Sin embargo, salvo en casos poco frecuentes, el interés se centra en espectros continuos en los que no aparecen tonos puros muy próximos que sea interesante separar. Es el caso de la mayoría de los sonidos y ruidos del ambiente.

La ventana de Hann permite una buena discriminación pero a costa de una considerable contaminación (ruido) en alta frecuencia. La ventana de Blackman ofrece un excelente compromiso entre pérdida de resolución y ruido residual en alta frecuencia, que es muy pequeño. La de Blackman-Harris empeora un poco la resolución a cambio de obtener una contaminación muy baja en alta frecuencia. La ventana de Hamming tiene la propiedad de lograr una buena resolución y al mismo tiempo repartir la contaminación en las altas frecuencias de manera bastante uniforme.

La ventana de Hann puede implementarse en Scilab mediante la función **window**, que permite varios tipos de ventanas (rectangular, triangular, Hann, Hamming y Kaiser). La transformada rápida incluyendo la multiplicación por la ventana es

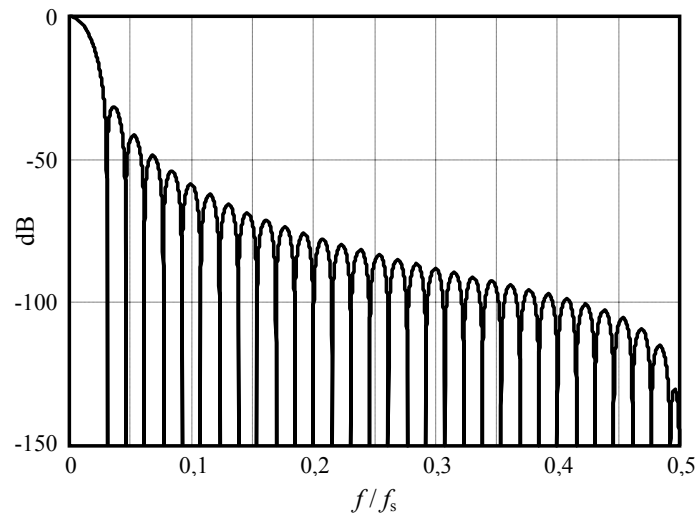
```
y = fft(x(k:k+N-1) .* window('hn',N), -1);
```

La ventana de Blackman, en cambio, no aparece en la versión 5.4.1 de Scilab (vigente al escribir este texto), por lo que será preciso calcularla o agregarla a la función **window**.

En el primer caso, se calcula un vector **black** con los valores de la ventana calculados según (8.33):

```
black = 0.42 + 0.5*cos(2*pi*[0:N]/(N-1)) ...
        + 0.08*cos(4*pi*[0:N]/(N-1));
y = fft(x(k:k+N-1).* black, -1);
```

El espectro de las ventanas (analizado en más detalle en el apéndice 16) tiene gran importancia, dado que la aplicación de la ventana implica que el espectro de la señal aparece convolucionado con el espectro de la ventana. La ventana que menos efecto produciría sería la que tuviera por espectro una delta de Dirac ya que es la única función que convolucionada con otra no la altera. Por consiguiente, cuanto más se parezca el espectro de una ventana a una delta de Dirac, menores serán los efectos negativos. Esto implica que lo deseable es tener un lóbulo de baja frecuencia (principal o central) alto y angosto y lóbulos laterales mucho menores. Como ejemplo, en la figura 95 se muestra el espectro (unilateral) de la ventana de Hann. Se observa un lóbulo central relativamente ancho, pero el primer lóbulo lateral es de  $-33$  dB relativos al lóbulo central. De allí en más se opera una reducción ulterior hacia las altas frecuencias (en el apéndice 16 se analizan otras ventanas).



**Figura 95.** Espectro unilateral de la ventana de Hann para  $N=64$  donde la escala del eje de frecuencia es relativa a la tasa de muestreo.

El ancho del lóbulo central determina la mínima separación que puede haber entre dos líneas espectrales (o entre dos tonos puros, cada uno de los cuales tiene dos pares de líneas en el espectro bilateral) para que sean discriminables por el analizador. En el caso de la ventana rectangular el lóbulo central tiene un ancho igual a

$$\Delta f = \frac{2}{T} = \frac{2F_s}{N}. \quad (8.36)$$

donde  $T$  es la duración de la ventana. En el caso de la ventana de Hann este valor se duplica, y en la de Blackman se triplica.

En la elección de una ventana hay un triple compromiso. Por un lado el poder de discriminación dado por el ancho del lóbulo principal. Por otro lado, el grado de derrame espectral, dado por el pico del primer lóbulo lateral. Finalmente, el grado de localización en el tiempo del espectro, dado por la longitud de la ventana. Así, por ejemplo, el uso de una ventana de Blackman permite un buen desempeño en cuanto a derrame espectral, pero para lograr buena discriminación de frecuencia, se requiere, según (8.36) aumentar la longitud  $N$ , perdiendo discriminación en el tiempo (principio de incertidumbre tiempo-frecuencia). En la tabla 8.1 se indican el máximo del lóbulo lateral y el ancho de banda del lóbulo principal para varias ventanas.

**Tabla 8.1.** Propiedades relevantes de algunas ventanas para análisis espectral

Ventana	Error máximo para ruido blanco [dB]	Ancho del lóbulo principal [Hz]	Máximo lóbulo lateral [dB]
Rectangular	-	$2f_s/N$	-13
Triangular	2,5	$4f_s/N$	-27
Hann	3	$4f_s/N$	-32
Hamming	3,4	$4f_s/N$	-43
Blackman	2	$6f_s/N$	-58
Blackman-Harris	1,5	$8f_s/N$	-92

Otro problema de las ventanas es la acumulación de derrames espectrales en el caso de señales espectralmente complejas como el ruido blanco. Esto sucede porque por cada línea espectral del ruido se agrega un ruido residual, y la superposición de todos ellos de hecho incrementa el ruido, provocando un error. El máximo error se detalla también en la tabla 8.1.

Un riesgo potencial se presenta con señales no estacionarias o con transitorios rápidos o de corta duración con respecto a la longitud de la ventana. Así, por ejemplo, un pulso mucho más corto que la ventana sufrirá una ponderación muy baja si el pulso aparece al principio o al final de la ventana comparado con cuando aparece en el centro. Este problema se puede resolver mediante técnicas de solapamiento. La ventana de Hann, por ejemplo admite un solapamiento del 50 % que permite una reconstrucción perfecta ya que una serie de ventanas solapadas dan una suma idénticamente igual a 1.

### 8.6.6 Análisis de espectro de bandas con FFT

El segundo tipo de análisis es el espectro de bandas de porcentaje constante, es decir, que cada banda es un determinado porcentaje mayor que la anterior. Así, en el caso de las bandas de octava cada banda es 100 % mayor que la banda anterior. La manera de lograr este análisis consiste en superponer energéticamente (es decir, cuadráticamente) las líneas espectrales comprendidas dentro de cada banda. Dado que el tamaño de las bandas crece exponencialmente con el índice de la banda, las bandas inferiores contienen muy pocas líneas y las superiores muchas. Según la selección del tamaño  $N$  de la ventana de análisis podría ocurrir, inclusive, que en la banda inferior no hubiera

ninguna línea espectral. En la tabla 8.2 se indican los límites de frecuencia, los anchos de banda y la cantidad de líneas espectrales para varios valores de  $N$  correspondientes a las diversas bandas de octava para  $F_s = 44100$  Hz. Como se aprecia, conviene adoptar por lo menos  $N = 8192$ , para asegurar que la banda inferior involucre varias líneas. Cuantas más líneas incluya cada banda, más preciso será el resultado, aunque, por el principio de incertidumbre, también menos localizado en el tiempo.

**Tabla 8.2.** Anchos de banda y cantidad de líneas espectrales de una FFT según el tamaño  $N$  de la ventana de análisis

$f_o$ [Hz]	$f_i$ [Hz]	$f_s$ [Hz]	$B$ [Hz]	$N = 4096$	$N = 8192$	$N = 16384$	$N = 32768$
31,5	22,4	4,50	22,6	2	4	8	17
63	45,0	90,0	45,0	4	8	17	33
125	90,0	180	90,0	8	17	33	67
250	180	355	175	16	33	65	130
500	355	710	355	33	66	132	264
1000	710	1400	690	64	128	256	513
2000	1400	2800	1400	130	260	520	1040
4000	2800	5600	2800	260	520	1040	2081
8000	5600	11200	5600	520	1040	2081	4161
16000	11200	22400	11200	1040	2081	4161	8322

Para el caso de bandas de tercio de octava la longitud de la ventana deberá ser todavía mayor dado que las primeras bandas son extremadamente angostas. Por ejemplo, la banda de tercio de octava centrada en 20 Hz tiene un ancho de banda de 4,63 Hz, mientras que la banda de octava centrada en 31,5 Hz tiene un ancho de banda de 22,3 Hz. Esto implica que la separación entre líneas,  $F_s/N$  debe ser pequeña, lo cual obliga a utilizar para la FFT una ventana muy larga.

La cantidad  $m_1$  de líneas en la primera banda de fracción  $b$  de octava  $f_{o1}$  puede calcularse como

$$m_1 = N \frac{f_{o1}}{F_s} \left( 2^{\frac{1}{2b}} - 2^{-\frac{1}{2b}} \right). \quad (8.37)$$

Para poder representar la energía sonora presente en dicha banda, en principio debe haber como mínimo una línea espectral en dicha ventana. Sin embargo, para tener una aceptable precisión teniendo en cuenta el derrame espectral para una ventana de Hann, que corresponde a 4 líneas, debe ser  $m_1 \geq 4$ , lo cual se logra si

$$N \geq 4 \frac{F_s}{f_{o1}} \frac{1}{2^{\frac{1}{2b}} - 2^{-\frac{1}{2b}}} \quad (8.38)$$

Para bandas de octava ( $b = 1$ ) resulta  $N \geq 7983$ , y para bandas de tercio de octava ( $b = 3$ ),  $N \geq 38696$ , de donde, adoptando las potencias de 2 más cercanas, resultan, respectivamente,  $N = 8192$  y  $N = 65536$  (en ambos casos, para  $F_s = 44\,100$  Hz).

Para calcular el valor de la banda  $h$ -ésima se procede así:

$$X_{\text{oct}, h} = \sqrt{\sum_{k=\lfloor 2^{-1/(2b)} N f_{oh}/F_s \rfloor + 1}^{\lfloor 2^{1/(2b)} N f_{oh}/F_s \rfloor} 2 \left| \frac{X(k)}{N} \right|^2} \quad (8.39)$$

donde  $X(k)$  es la transformada FFT de  $x(n)$ ,  $f_{oh}$  es la frecuencia central de la banda  $h$  y  $\lfloor \cdot \rfloor$  representa la parte entera. Si se aplicó una ventana temporal  $w(n)$ , debe corregirse por la atenuación que produce la ventana, dividiendo por

$$C_w = \sqrt{\frac{1}{N} \sum_{n=1}^N w^2(n)}. \quad (8.40)$$

ya que ése será el factor por el que se afecta el valor eficaz de las componentes frecuenciales. Esta corrección supone que la señal es estacionaria en una escala temporal correspondiente a la ventana. En caso de no serlo, habría que tomar ventanas solapadas para evitar que se elimine información cercana a los bordes. El caso más favorable es cuando las ventanas solapadas suman 1, por ejemplo, en el caso de la ventana de Hann con un solapamiento del 50 %. En este caso no hace falta efectuar ninguna corrección.

El algoritmo descrito a continuación supone que el tamaño de la señal es 8192. El primer paso es aplicar la ventana de Blackman:

```
N = 8192;
black = 0.42 + 0.5*cos(2*pi*[0:N]/(N-1)) ...
        + 0.08*cos(4*pi*[0:N]/(N-1));
xb = x .* black;
```

Observar que se utiliza aquí el producto componente por componente (es decir, `.*`) entre los vectores **x** y **black**. A continuación se calcula el módulo del espectro de líneas, teniendo en cuenta el factor de corrección mencionado anteriormente (igual a  $1/N$  para la primera línea y  $2/N$  para el resto):

```
e = 2/N * abs(fft(xb, N));
e(1) = e(1)/2;
```

Este espectro consta de 8192 líneas, de las cuales sólo son significativas las 4096 primeras. El código que sigue determina, para cada banda, los índices mínimo y máximo de las líneas comprendidas en esa banda. Para ello se tiene en cuenta que la banda de partida es la de  $f=1000$  Hz. A partir de ella se obtienen las frecuencias centrales exactas dividiendo y multiplicando por 2. Aunque éstas no coinciden con las frecuencias normalizadas para las bandas de 31,5 Hz y 63 Hz, la diferencia es pequeña y, en cambio, el código se simplifica mucho. Los extremos de cada banda se obtienen dividiendo y multiplicando por  $\sqrt{2}$  respectivamente.

Finalmente, dado que las líneas espectrales están uniformemente repartidas a una distancia de  $F_s/N$ , dividiendo por dicho valor se obtienen los extremos aproximados, que se convierten en enteros mediante las funciones **ceil** (mínimo entero superior) y **floor** (máximo entero inferior):

```

for h = 1:10
    imin(h) = ceil(1000 * 2^(h - 6.5) * N / Fs);
    imax(h) = floor(1000 * 2^(h - 5.5) * N / Fs);
    if imax(h) > N/2 - 1
        imax(h) = N/2 - 1;
    end
end

```

El condicional evita la aparición de líneas más allá de la frecuencia de Nyquist, lo cual puede suceder en la banda superior, que por definición podría contener hasta aproximadamente 22 400 Hz pero está limitada a 22 050 Hz.

A continuación se calcula el contenido de las bandas sumando los cuadrados de las líneas espectrales de cada banda (superposición energética) y luego aplicando raíz cuadrada:

```

for h = 1:10
    y(h) = sqrt(sum(e(imin(h):imax(h)).^2));
end

```

Si bien para mayor claridad este paso se ha separado del lazo anterior, podría incluirse dentro del mismo.

Por último, es necesario corregir los valores anteriores por la reducción de energía que impone la ventana de Blackman. Para ello se obtiene un factor  $Q$  igual a la recíproca del valor eficaz de la ventana:

```

Q = 1 / sqrt(mean(black.^2));
Yef = y * Q / sqrt(2);

```

La división por  $\sqrt{2}$  es al solo efecto de obtener el valor eficaz de las bandas, al que se podrá aplicar la ecuación de calibración (7.7)).

### 8.6.7 Densidad espectral y promediación de espectros

Hay muchos casos en los que interesa medir la densidad espectral, por ejemplo para la caracterización de ruidos de banda ancha relativamente estacionarios, para la aplicación de modelos dinámicos basados en funciones de transferencia, para identificación de fuentes de señal o incluso para la realización de diagnóstico preventivo de fallas. Un problema con la transformada discreta de Fourier es que al ser de corta duración no alcanza a representar un espectro promedio a largo plazo.

Podríamos pensar en tomar una FFT con una cantidad muy grande de puntos pero surgen aquí dos problemas. El primero es que la cantidad de cálculos a realizar y, sobre todo, la cantidad de memoria requerida, hacen imposible la aplicación eficiente de una FFT, por ejemplo, de decenas de millones de puntos. El segundo es que, si bien la resolución del espectro mejora proporcionalmente a la cantidad de puntos, no se consigue que ninguna de las numerosas líneas espectrales se estabilice ya que cada una está siendo calculada una única vez.

La manera de solucionar este problema es tomar numerosos espectros FFT correspondientes a cuadros sucesivos y promediar las líneas homólogas (es decir, correspondientes a la misma frecuencia). En este caso sí estaremos determinando numerosas veces la misma magnitud, lo que tiende a estabilizar el promedio (véase la sección 1.7.4). Por otra parte, desde el punto de vista de la eficiencia computacional, requiere menos

cálculos efectuar varios espectros de corta duración que uno solo que abarque el mismo tiempo total. En efecto, si llamamos  $N$  al tamaño de cada FFT pequeña y  $M$  a la cantidad de ventanas de análisis, la cantidad de operaciones para una FFT es  $K N \log_2 N$ , donde  $K$  es una constante. La realización de  $M$  cálculos análogos insume una cantidad de operaciones  $MKN \log_2 N$ .

El cálculo de una FFT gigantesca de  $NM$  puntos requiere  $MKN \log_2(MN)$  cálculos, es decir

$$MKN \log_2(MN) = MKN \log_2 N + MKN \log_2 M \quad (8.41)$$

Vemos que en este último caso se realizan  $MKN \log_2 M$  operaciones más. Por ejemplo, para analizar un minuto de audio a 44100 Hz se requerirían una FFT de 2646000 muestras que requieren 56 millones de operaciones. La realización de  $M$  cálculos de FFT de 4096 puntos requiere  $M = 646$  cuadros. La cantidad total de operaciones será entonces 32 millones, es decir se ahorra casi un 50 % de operaciones. El ahorro es tanto mayor cuanto más muestras en total se deben analizar.

### 8.7 Filtros FFT

Los filtros IIR permiten realizar funciones filtrantes con un orden (cantidad de retardos) bajo, lo cual los hace computacionalmente muy eficientes. En contraposición, pueden volverse fácilmente inestables por la acumulación de errores de truncado. En efecto, por la recursividad, el valor de cada muestra de la salida interviene en todas las muestras subsiguientes. Además es imposible lograr una respuesta de fase lineal con la frecuencia, característica deseable para mejorar la respuesta transitoria.

Los filtros FIR son inherentemente estables, dado que cada muestra repercute en sólo un número fijo de muestras futuras, y posibilitan un diseño de fase lineal con la frecuencia (Proakis et al., 1998). Presentan el inconveniente de que requieren un orden elevado para conseguir una respuesta dada. Ello se debe a que un filtro selectivo tiene una respuesta al impulso prolongada, lo que exige gran número de términos. El tiempo de cómputo crece, así, considerablemente.

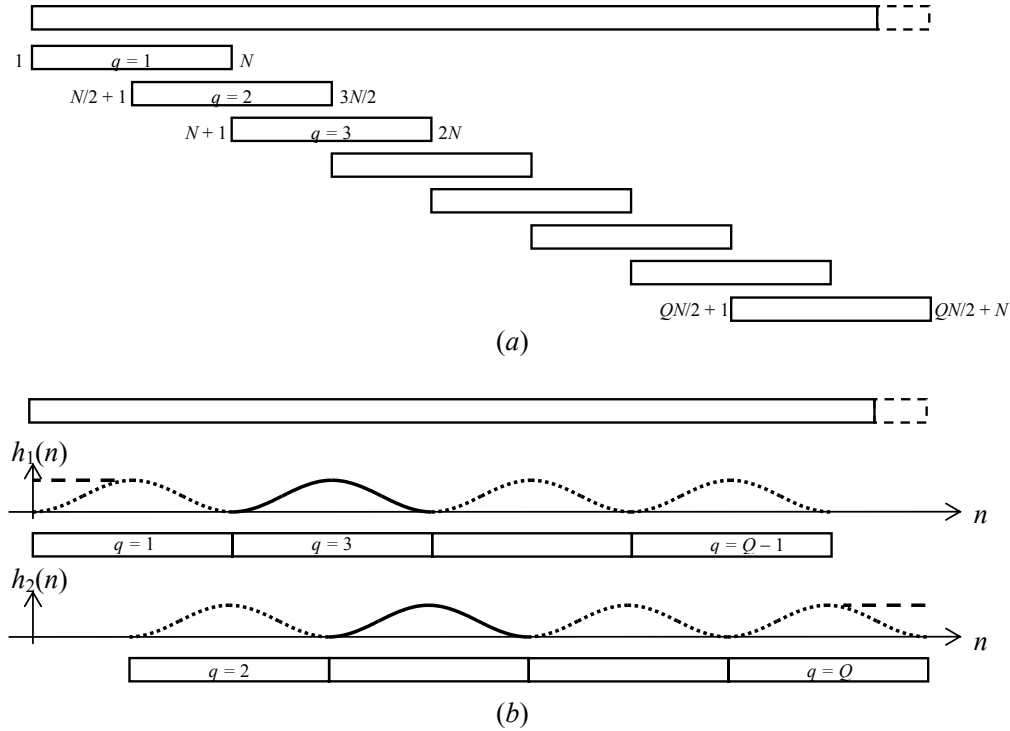
Una alternativa consiste en realizar el filtrado en el dominio de la frecuencia, lo cual implica una simple multiplicación de la transformada de Fourier de la señal,  $X(k)$ , por la ventana filtrante:

$$Y(k) = X(k) W(k), \quad (8.42)$$

donde  $W(k)$  contiene la respuesta en frecuencia para la frecuencia  $kF_s/N$ . La señal filtrada se recupera, entonces, mediante la transformada inversa de Fourier. Este procedimiento permite reducir el tiempo de cómputo considerablemente. En efecto, para calcular  $N$  muestras de un filtro FIR de orden  $N$  se requieren  $N^2$  sumas y productos, mientras que con el procedimiento indicado sólo se requerirían del orden de  $N \log_2 N + N$ . Para  $N = 4096$ , por ejemplo, se reduce en un factor 315.

Tal como está planteado, dicho filtrado por FFT se limita a una ventana de análisis de longitud  $N$ . Si bien es posible incrementar  $N$ , por un lado el algoritmo se vuelve menos eficiente y, por otro, implica aumentar innecesariamente la resolución del espectro. Para extender el método a señales arbitrariamente largas se subdivide la señal en cuadros o ventanas de longitud  $N$  (es decir,  $N$  muestras cada uno) con un solapamiento  $N/2$ . A cada cuadro se le aplica la FFT obteniéndose el correspondiente espectro, se multiplica cada espectro por la ventana filtrante deseada, y se le aplica la FFT inversa. La señal temporal

resultante en cada cuadro es la versión filtrada del cuadro original. Las partes solapadas en cuadros sucesivos son muy parecidas excepto en los extremos debido a comportamientos transitorios causados por el fenómeno de Gibbs.<sup>107</sup> Estas diferencias se pueden reducir mucho aplicando una suma ponderada de ambas partes solapadas de modo que cerca de los extremos el peso tienda a 0 y en cambio en la región central del cuadro tienda a 1. La ventana de Hann resulta ideal dado que en la zona solapada la suma de las dos ventanas es idénticamente igual a 1. En la figura 96 se muestra la estructura del método.



**Figura 96.** Filtrado FFT de longitud arbitraria. (a) Descomposición en cuadros de longitud  $N$  y solapamiento  $N/2$ . (b) Ventanas de Hann para los cuadros pares e impares. Se resaltaron dos ventanas individuales. En línea horizontal de trazos se muestran las modificaciones en los cuadros inicial y final.

Si la ventana correspondiente al primer cuadro es

$$h_1(n) = 0,5 \left( 1 - \cos \left( \frac{2\pi(n-1)}{N} \right) \right) \quad (8.43)$$

y la correspondiente al segundo es

<sup>107</sup> El fenómeno de Gibbs consiste en la aparición de grandes oscilaciones en una aproximación de Fourier limitada en banda (por ejemplo una serie de Fourier truncada a un número finito de términos) en las proximidades de una discontinuidad (por ejemplo, en los flancos ascendentes y descendentes de una onda cuadrada).



$$h_2(n) = 0,5 \left( 1 - \cos \left( \frac{2\pi(n - N/2 - 1)}{N} \right) \right) = 0,5 \left( 1 + \cos \left( \frac{2\pi(n - 1)}{N} \right) \right), \quad (8.44)$$

se verifica inmediatamente que

$$h_1(n) + h_2(n) = 1. \quad (8.45)$$

Este tipo de método se conoce como solapado y suma (overlap-add) y fue introducido por Crochiere (1980). Sobre el proceso descrito anteriormente, debemos destacar dos hechos. El primero es que a efectos de tener una cantidad entera de cuadros, previamente se prolonga la señal con la cantidad necesaria de ceros (*zero-padding*). El segundo es que en los cuadros extremos las ventanas aplicadas se modifican para obtener una ponderación plana en todo el intervalo y no sólo donde se da la superposición.

Los filtros FFT constituyen una excelente solución al problema de filtrado con respuesta arbitraria ya que filtran la señal dejando muy pocos residuos indeseados. Tienen incluso, fase nula, que es un caso particular de fase lineal con la frecuencia. Que un filtro tenga fase lineal con la frecuencia significa que todas las componentes espectrales que atraviesan el filtro se retardan en una misma cantidad, ya que el retardo es

$$\Delta T(f) = \frac{\varphi(f)}{2\pi} T = \frac{Kf}{2\pi} \frac{1}{f} = \frac{K}{2\pi} \quad (8.46)$$

En el caso del filtro FFT la constante  $K$  es nula. En particular esto implica que se trata de un filtro *no causal*, ya que para una señal que empieza en algún instante intermedio de la primera ventana, la respuesta comienza antes, es decir, tiene un pre-transitorio, a veces llamado *pre-eco*, si bien el efecto en general no es significativo. Esto se debe a la naturaleza de la DFT, ya que está representando una señal a lo largo de una ventana mediante una serie de componentes de amplitud constante a lo largo de la ventana. Una señal que empiece con unas cuantas muestras nulas es correctamente representada por  $N$  términos cuyas amplitudes son la solución de un sistema de ecuaciones (ver ecuación (8.10)), pero si se eliminan o modifican algunos de dichos términos, no hay garantía de que la suma de los que queden siga siendo nula en los instantes en que lo era la señal original.

Como caso extremo que ilustra el problema, supongamos que representamos un pulso unitario ubicado en el centro de la ventana. La transformada discreta del pulso unitario tiene amplitud constante. Si aplicamos un filtro que elimine todas las frecuencias excepto una, quedarán sólo dos líneas espectrales conjugadas, cuya suma tiene amplitud constante a lo largo de la totalidad de la ventana. Este fenómeno es similar al derrame espectral y podríamos denominarlo *derrame temporal*.<sup>108</sup>

Con todo, la no causalidad se extiende a lo sumo a la longitud de una ventana, por lo que si ésta no es excesivamente larga, para la mayor parte de los propósitos metrológicos no tiene mayores consecuencias.

Otra pequeña dificultad del filtro FFT es que aparece una ligera modulación de amplitud de una frecuencia igual a  $2/T$  donde  $T$  es la duración de la ventana de análisis. Esta modulación puede reducirse aumentando el solapamiento, por ejemplo, haciendo un solapamiento al 25 %. En ese caso también es posible utilizar una ventana de Black-

<sup>108</sup> Existe una marcada dualidad entre los fenómenos espectrales y temporales debido a la similitud entre las transformadas discretas directa e inversa de Fourier (sólo difieren en el signo de los exponentes)

man, ya que la misma se puede descomponer en dos ventanas de Hann de diferente amplitud y frecuencia:

$$w_{blackman}(n) = 0,5 \left( 1 - \cos \frac{2\pi n}{N-1} \right) + 0,08 \left( 1 - \cos \frac{4\pi n}{N-1} \right) \quad (8.47)$$

Los filtros FFT se pueden utilizar para realizar filtros prácticamente ideales, tanto pasabajos como pasaaltos, pasabanda o rechazabanda. Podrían servir, por ejemplo, para eliminar un tono parásito que contamina una medición con mínimo efecto sobre el resto del espectro. En este caso es recomendable utilizar una resolución alta ( $N$  grande) y seleccionar una ventana que deje pasar algunas líneas alrededor del tono a eliminar.

Al diseñar una ventana filtrante para un filtro FFT es necesario tener en cuenta las condiciones de simetría de la FFT dadas por (8.21), ya que de lo contrario la respuesta pasará a tener componentes imaginarias. Con respecto a la ecuación (8.21), nos indica qué componentes deben aparecer con la misma amplitud en la ventana filtrante, pero dado que en Scilab y otros entornos matemáticos similares los índices comienzan en 1 y no en 0, la definición de la FFT tal como está implementada es

$$X(n) = \sum_{k=1}^N x(k) e^{-j2\pi \frac{(n-1)(k-1)}{N}} \quad (8.48)$$

de donde la relación de simetría pasa a ser, para  $n = 1, \dots, N$ ,

$$X(N - n + 2) = \overline{X(n)}. \quad (8.49)$$

Entonces la ventana filtrante deberá cumplir

$$W(N - n + 2) = \overline{W(n)}. \quad (8.50)$$

Nótese que si bien en la mayoría de los casos la ventana será real, una ventana con parte imaginaria permite también incorporar una información de fase.

Normalmente se diseña la ventana para  $n = 1, \dots, N/2$  y luego se aplica (8.50). Un fragmento de código que implementa esto partiendo de una ventana **w** definida como vector columna para las frecuencias menores que  $F_s/2$ , es

```
w = [w; zeros(N/2-length(w), 1)];  
w = [w; 0; conj(w($:-1:2))];
```

La realización de la FFT puede aprovechar la posibilidad que ofrece la función **fft** de realizar varias transformadas con un solo comando. Supongamos que en la matriz **xx** de dimensión  $N \times m$ , donde  $N$  es el tamaño de la ventana tenemos la señal a transformar descompuesta en cuadros en forma de vector columna. Entonces el código

```
xx = fft(xx, -1, 1);
```

realiza  $m$  transformadas, una para cada columna. El último argumento indica que las transformadas se calculan a lo largo de la dimensión 1, es decir, las filas (esto significa que en cada caso la variable a transformar son todas las filas de cada columna)

Los filtros FFT se pueden utilizar para realizar filtros de banda de octava y tercio de octava. En ese caso, la ventana deberá diseñarse primero para las frecuencias menores que  $F_s/2$  de acuerdo a la siguiente expresión:

$$W(n) = \begin{cases} 0 & \text{si } 1 \leq n < \frac{f_o}{F_s} 2^{-1/2b} \\ 1 & \text{si } \frac{f_o}{F_s} 2^{-1/2b} \leq n < \frac{f_o}{F_s} 2^{1/2b} \\ 0 & \text{si } \frac{f_o}{F_s} 2^{1/2b} \leq n \leq N/2 \end{cases} \quad (8.51)$$

donde  $b = 1$  para banda de octava y  $b = 3$  para banda de tercio de octava. El siguiente fragmento de código permite implementar este diseño:

```
finf = fo / 2^(1/tipo/2);
fsup = fo * 2^(1/tipo/2);
w = zeros(1, floor(finf/Fs*N));
w = [w, ones(1, floor(fsup/Fs*N) - floor(finf/Fs*N))];
w = [w, zeros(1, N/2 - floor(fsup/Fs*N))];
```

En <http://www.fceia.unr.edu.ar/acustica/mediciones> podrá obtenerse una realización completa del filtro FFT y su aplicación al filtrado de bandas.

Es interesante notar que en el caso en que se desee obtener una descomposición de la señal en un conjunto completo de bandas, el algoritmo se puede hacer computacionalmente más eficiente si en lugar de realizar para cada banda la FFT y la IFFT, se hace una sola FFT y luego una IFFT para cada banda. Esto permite reducir la carga computacional prácticamente a la mitad.

También conviene tener en cuenta que, a diferencia de lo que sucede con los filtros IIR, donde la obtención del valor eficaz necesariamente requiere obtener la respuesta temporal del filtro para, a partir de ella, realizar el cálculo, en el caso de utilizar FFT ello no es necesario. En efecto, una vez obtenida la representación espectral la energía en cada banda puede obtenerse sumando las energías de las líneas espectrales comprendidas en la banda, tal como se indica en la ecuación (8.39). Conceptualmente se está aplicando la igualdad de Parseval, según la cual la energía se puede obtener tanto en el dominio temporal como en el espectral.

Por este motivo, el uso de un filtro FFT sólo está indicado para los casos en que sea de interés disponer de la señal filtrada..

## 8.8 Algunas aplicaciones de la FFT

### 8.8.1 Determinación de carácter tonal

En algunos casos, sobre todo en la evaluación de ruido urbano o en algunos estudios psicoacústicos, es necesario determinar el carácter tonal de un ruido. Esta cualidad se refiere al hecho de que un oyente medio perciba una sensación de tonalidad en el ruido a evaluar. La presencia de tonos puros perceptibles generalmente acentúa los efectos negativos del ruido. Así, diversos criterios tanto de molestia como de efectos sobre

la capacidad auditiva penalizan los ruidos tonales sumando una constante al nivel de evaluación.<sup>109</sup>

La Norma ISO 1996-2 y su equivalente argentino, la IRAM 4113-2, establecen un procedimiento para determinar la audibilidad de tonos. El mismo se basa en criterios psicoacústicos de enmascaramiento.

Supongamos que realizamos una grabación digital con una tasa de muestreo  $f_s = 44100$  Hz de un ruido que contiene un tono de 500 Hz. Queremos determinar si el mismo resultará audible aplicando los métodos de la norma ISO 1996-2 (o su equivalente argentino, IRAM 4113-2).

En la frecuencia de 500 Hz el ancho de la banda crítica es  $B_{\text{crit}} = 100$  Hz. El ancho de banda efectivo del método de análisis de espectro de banda angosta debe ser un 5 % del ancho de banda crítico, es decir:

$$B_{\text{ef}} < 0,05 B_{\text{crit}} = 5 \text{ Hz} . \quad (8.52)$$

Si efectuamos el análisis mediante FFT de longitud  $N$  muestras, el ancho de banda efectivo del método con una ventana de Hann es (apéndice 16)

$$B_{\text{ef}} = 1,5 \frac{F_s}{N} , \quad (8.53)$$

de donde

$$N > \frac{1,5 F_s}{B_{\text{ef}}} = \frac{1,5 \times 44100 \text{ Hz}}{5 \text{ Hz}} = 13230 .$$

Elegimos la siguiente potencia de 2, es decir,  $N = 16384$ . El ancho de banda efectivo resulta entonces igual a 4,04 Hz, que satisface el requisito. En la banda crítica de 500 Hz, es decir el intervalo [450 Hz; 550 Hz] hay una cantidad  $M$  de líneas espectrales dada por

$$M = \frac{B_{\text{crit}}}{F_s / N} \cong 37 .$$

Supongamos que el tono en cuestión posee un nivel de presión sonora  $L_{\text{pt}} = 65$  dB, definido como la superposición energética de aquellas líneas hasta 6 dB menores que el pico local que caracteriza al tono:

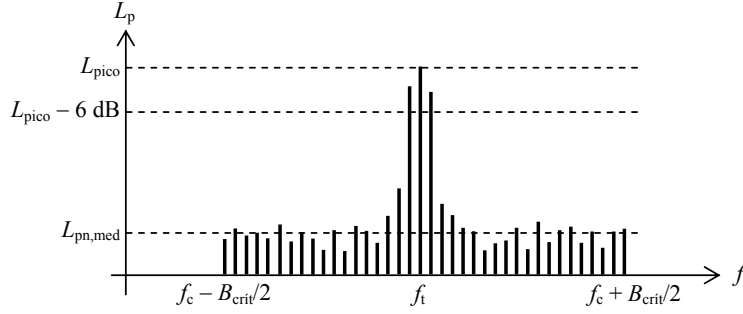
$$L_{\text{pt}} = 10 \log \left( \sum_{L_{p,k} > L_{\text{pico}} - 6 \text{ dB}} 10^{L_{p,k}/10} \right) \quad (8.54)$$

Supongamos también que el ruido enmascarante tiene un nivel medio en la banda crítica  $L_{\text{pn,med}} = 44$  dB, definido como

<sup>109</sup> En la valoración de ruidos se suele llamar *nivel de evaluación* a algún parámetro obtenido a partir de una medición objetiva agregando correcciones que tienen en cuenta diversos aspectos contextuales o la tipología del ruido o su fuente.

$$L_{pn,med} = 10 \log \left( \frac{1}{H} \sum_{L_{p,k} < L_{pico} - 6 \text{ dB}} 10^{L_{p,k}/10} \right) \quad (8.55)$$

donde  $H$  es la cantidad de líneas espectrales que satisfacen la condición  $L_{p,k} < L_{pico} - 6 \text{ dB}$ . En la figura 97 se muestra un ejemplo.



**Figura 97.** Espectro FFT de un ruido con una componente tonal en la banda crítica centrada en  $f_c$ , representada por varias líneas espectrales.

El valor total del ruido enmascarante se obtiene, según la norma, mediante la siguiente expresión:

$$L_{pn} = L_{pn,med} + 10 \log \frac{B_{crit}}{B_{ef}} = 44 \text{ dB} + 14 \text{ dB} = 58 \text{ dB} \quad (8.56)$$

El hecho de que la corrección sea  $10 \log (B_{crit}/B_{ef})$  y no  $10 \log (B_{crit}/(F_s/N))$ , donde  $F_s/N$  es la separación entre líneas espectrales, obedece a que un ruido de banda ancha se ve sobredimensionado debido al derrame espectral (ver apéndice 16).

La diferencia de niveles entre el tono y el ruido enmascarante es

$$Y = L_{pt} - L_{pn} = 65 \text{ dB} - 58 \text{ dB} = 7 \text{ dB}$$

La *audibilidad tonal* se define como

$$\Delta L_{ta} = L_{pt} - L_{pn} - MT$$

donde  $MT$  es el *umbral de enmascaramiento*, un parámetro psicoacústico que admite un modelo simplificado dado por

$$MT = \log \frac{1}{1 + (f_c / 502 \text{ Hz})^{2,5}} - 2 \text{ dB} \quad (8.57)$$

Resulta:

$$\Delta L_{ta} = 65 \text{ dB} - 58 \text{ dB} - (-2,3 \text{ dB}) = 9,3 \text{ dB}$$

Esto significa que el tono se encuentra 9,3 dB por encima del umbral de enmascaramiento.

El ajuste por tonalidad  $K_t$  es el valor a ser agregado al nivel equivalente ponderado  $A$  obtenido por medición directa (o cálculo) y está dado por:

$$K_t = \begin{cases} 6 \text{ dB} & 10 \text{ dB} < \Delta L_{ta} \\ \Delta L_{ta} - 4 \text{ dB} & 4 \text{ dB} \leq \Delta L_{ta} \leq 10 \text{ dB} \\ 0 \text{ dB} & \Delta L_{ta} < 4 \text{ dB} \end{cases} \quad (8.58)$$

En nuestro caso corresponde a la segunda opción, por lo cual

$$K_t = 9,3 \text{ dB} - 4 \text{ dB} = 5,3 \text{ dB}$$

### 8.8.2 Convolución mediante FFT

La convolución con una respuesta al impulso es una operación que posibilita la realización de filtros y otros procesamientos de señal en el dominio tiempo. El problema es que la realización de una convolución de larga duración requiere una cantidad prohibitiva de operaciones.

Puede llevarse a cabo una convolución recordando que la transformada de una convolución es el producto de las transformadas. Sin embargo, para el caso de señales de duración acotada o periódica representables por una DFT, deberemos usar la *convolución circular*, definida como

$$x * y(k) = \sum_{h=0}^{N-1} x(h) y((k-h) \bmod N) \quad (8.59)$$

donde  $a \bmod b$  ( $a$  módulo  $b$ ) es el resto de dividir  $a$  por  $b$ . Podemos plantear

$$x * y(k) = \sum_{h=0}^{N-1} \left( \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{j \frac{2\pi m h}{N}} \right) \left( \frac{1}{N} \sum_{n=0}^{N-1} Y(n) e^{j \frac{2\pi (k-h)n}{N}} \right)$$

Observar que la periodicidad de la exponencial compleja permitió eliminar el operador *módulo*. Operando,

$$\begin{aligned} x * y(k) &= \frac{1}{N^2} \sum_{h=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} X(m) Y(n) e^{j \frac{2\pi m h}{N}} e^{j \frac{2\pi (k-h)n}{N}} = \\ &= \frac{1}{N^2} \sum_{h=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} X(m) Y(n) e^{j \frac{2\pi n k}{N}} e^{j \frac{2\pi (m-n)h}{N}} \end{aligned}$$

Conmutando el orden de las sumatorias y sacando factores comunes,

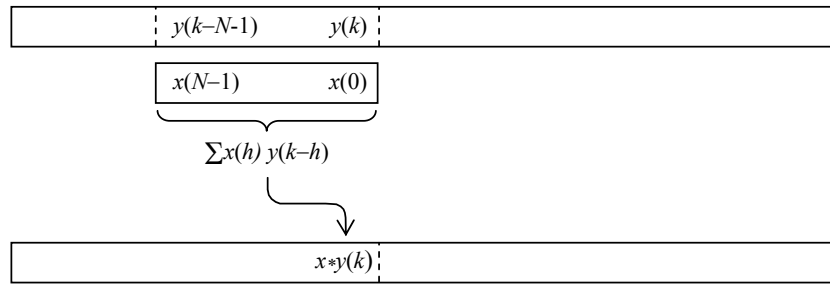
$$x * y(k) = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} X(m) Y(n) e^{j \frac{2\pi n k}{N}} \sum_{h=0}^{N-1} e^{j \frac{2\pi (m-n)h}{N}}$$

La suma más interna es  $N$  para  $m = n$  y 0 para  $m \neq n$ . Resulta, entonces

$$x^* y(k) = \frac{1}{N^2} \sum_{n=0}^{N-1} X(n)Y(n) e^{j \frac{2\pi nk}{N}} \quad (8.60)$$

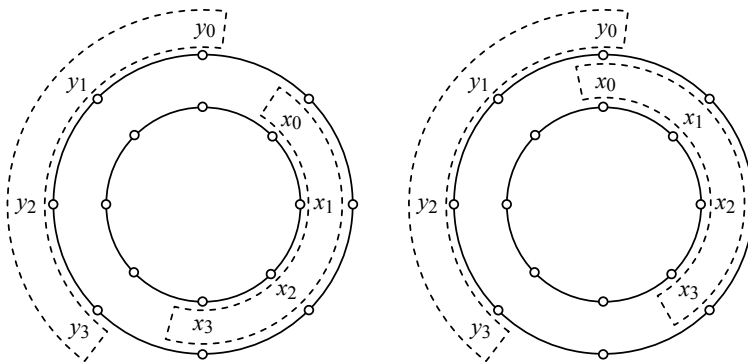
es decir, la convolución circular es la transformada inversa del producto de las transformadas.

Supongamos que queremos obtener la convolución entre una señal arbitraria  $x(k)$  y una señal corta  $y(k)$  (una respuesta al impulso finita o aproximada por una respuesta finita). En la figura 98 se muestra la construcción del valor correspondiente al instante discreto  $k$ . Al variar  $k$  la versión retrógrada de  $x$  se va deslizando.



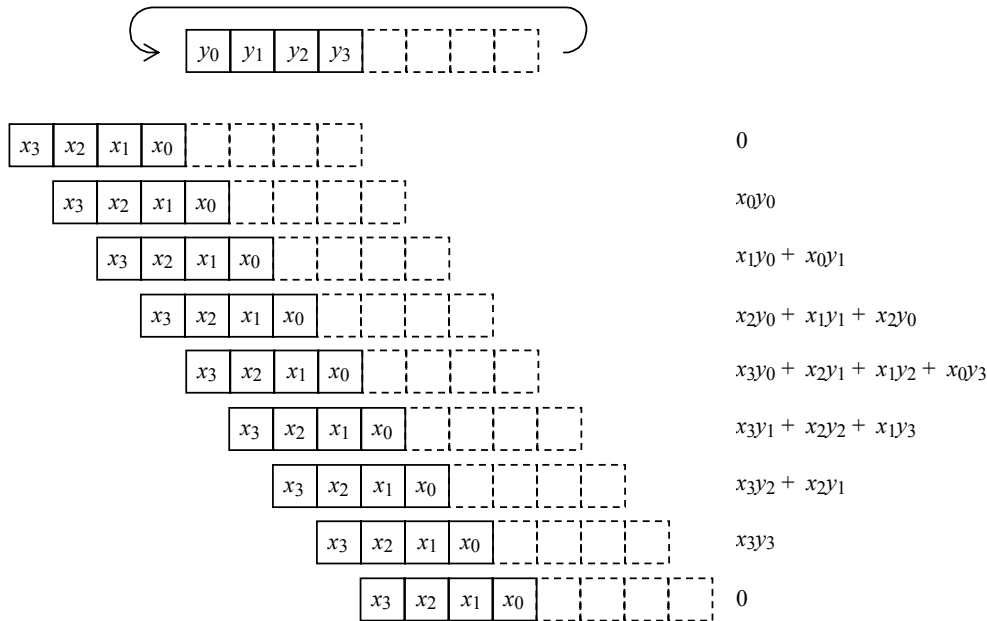
**Figura 98.** Convolución de una señal arbitrariamente larga  $y(k)$  con una señal corta  $x(k)$ . Se muestra la obtención de la muestra  $k$ .

En el caso de una convolución circular el deslizamiento es a lo largo de un camino circular. La manera de poder utilizar la convolución circular (lo que permitiría aplicar la FFT y la IFFT) es realizar la convolución por fragmentos y a cada fragmento completarlo con una cantidad igual de ceros (*zero-padding*). En la figura 99 se muestra un ejemplo para una convolución de  $N=4$  muestras. Se construyen dos señales expandidas a  $2N$ , completando con ceros (no indicados en la figura para mayor claridad). En este caso el deslizamiento es circular. Se muestra la posición para el cálculo de dos valores



**Figura 99.** Convolución circular de dos señales  $y_k$  y  $x_k$  de longitud  $N=4$  que han sido completadas con  $N$  ceros (no indicados). Se muestran las posiciones de la señal retrógrada  $x_k$  para dos instantes. El cálculo requiere multiplicar los valores internos y externos y sumar.

En la figura 100 se muestra la misma convolución desplegada en forma lineal para todas las posibles posiciones de la señal deslizante. Esta convolución se realiza transformando ambas señales mediante la FFT, multiplicando, y antitransformando con IFFT.



**Figura 100.** Convolución lineal de dos señales  $y_k$  y  $x_k$  de longitud  $N=4$ , realizada mediante convolución circular luego de completar con  $N$  ceros. La convolución resultante en realidad tiene periodo  $2N$ , pero se conservan sólo  $2N$  valores

Una vez lograda esta forma de convolucionar dos señales usando FFT, la misma se aplica a  $M$  sucesivos cuadros de  $N$  muestras tomados de una señal arbitrariamente larga, cada uno de los cuales es expandido con 0's. Cada convolución tiene un tamaño  $2N$ , pero se encuentran solapadas en  $N$  muestras. Sumando los resultados de todas ellas se obtiene la convolución total. En la figura 101 se ilustra este procedimiento.

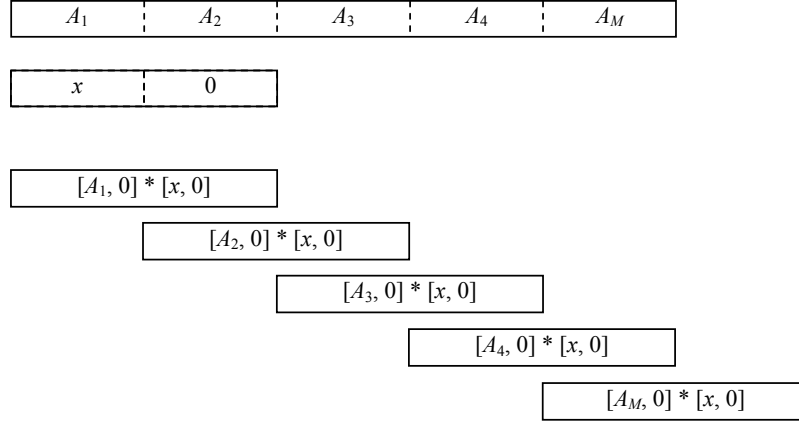
La carga computacional pasa de ser  $MN^2$  como sería con el cálculo directo, a ser  $2MN \log_2 2N = 2MN (1 + \log_2 N)$ . Las ventajas se observan cuando  $N$  es alto, lo cual corresponde a respuestas al impulso largas.

### 8.8.3 Correlación mediante FFT

La *correlación* entre dos señales, también llamada *correlación cruzada*, es una operación de interés porque permite detectar rasgos comunes a ambas señales. Para tiempo largo se la define como:

$$\varphi_{xy}(k) = \sum_{h=-\infty}^{\infty} x(h) y(h-k) \quad (8.61)$$





**Figura 101.** Convolución lineal de una señal formada por cuadros de longitud  $N$  designados como  $A_1, \dots, A_5$  y una señal  $x$  de longitud  $N$ . Se realizan convoluciones parciales luego de completar con  $N$  ceros. Las convoluciones resultantes, solapadas al 50 %, se suman

Para tiempo corto, es

$$\Phi_{xy,N}(n_o, k) = \sum_{h=n_o-N}^{n_o+N-1} x(h) y(h-k) \quad (8.62)$$

Para su cálculo eficiente también se puede aplicar la FFT. Para ello notemos que la forma general se parece a la de una convolución sólo que el argumento de  $y$  tiene el signo cambiado. Podemos asimilar esta expresión a la (8.59) mediante un adecuado cambio de variable:

$$\Phi_{xy,N}(n_o, k) = \sum_{h=-N}^{N-1} x(n_o + h) y(n_o + h - k) \quad (8.63)$$

Llamando

$$\begin{aligned} x_o(h) &= x(n_o + h) \\ y_o(h) &= y(n_o - h) \end{aligned} \quad (8.64)$$

resulta

$$\Phi_{xy,N}(n_o, k) = \sum_{h=-N}^{N-1} x_o(h) y_o(k - h) = x_o * y_o(k). \quad (8.65)$$

Ahora es calculable con el método de la sección anterior.

#### 8.8.4 Filtros de bandas críticas

En experimentos psicoacústicos muchas veces se requiere trabajar con bandas críticas. Los filtros de banda expuestos anteriormente también se pueden implementar para bandas no uniformes, como son las bandas críticas. Para ello basta reemplazar, en las ecuaciones (51) los límites de las bandas críticas según las aproximaciones de Zwicker (Zwicker et al., 1999)<sup>110</sup>. Dada la frecuencia  $f_0$  (en Hz) se puede obtener el ancho de banda crítica como

$$\Delta f_{BC} = 25 + 75 \left( 1 + 1,4 \left( \frac{f}{1000} \right)^2 \right)^{0,69}, \quad (8.66)$$

Aquí hay dos posibilidades. La primera es suponer que la banda está repartida mitad hacia cada lado de la frecuencia en cuyo caso la ventana queda definida como

$$W(n) = \begin{cases} 0 & \text{si } 1 \leq n < \frac{f_0 - \Delta f_{BC}}{F_s} \\ 1 & \text{si } \frac{f_0 - \Delta f_{BC}}{F_s} \leq n < \frac{f_0 + \Delta f_{BC}}{F_s} \\ 0 & \text{si } \frac{f_0 + \Delta f_{BC}}{F_s} \leq n \leq N/2 \end{cases} \quad (8.67)$$

Este método no garantiza que la banda crítica se encuentre centrada verdaderamente en  $f_0$ . Otra posibilidad es partir de la fórmula de Zwicker que da el índice de banda crítica correspondiente  $z_0$  (en bark),

$$z_0 = 13 \operatorname{arctg} \frac{f}{1316} + 3,5 \operatorname{arctg} \left( \left( \frac{f_0}{7500} \right)^2 \right), \quad (8.68)$$

a partir de éste obtener los índices  $z_0 \pm 0,5$  bark y luego invertir la (8.68) numéricamente (por ejemplo usando el método iterativo de Newton-Raphson) para obtener las frecuencias inferior y superior a utilizar en (8.67).

#### 8.8.5 Determinación de funciones transferencia

Un último ejemplo de aplicaciones de la FFT es en la determinación de funciones de transferencia. Si se dispone de dos señales  $x(t)$  e  $y(t)$  que son, respectivamente, la entrada y la salida de un sistema, obtenidas simultáneamente, entonces es posible obtener la función de transferencia simplemente dividiendo las FFT de ambas señales:

$$H(n) = \frac{Y(n)}{X(n)} \quad (8.69)$$

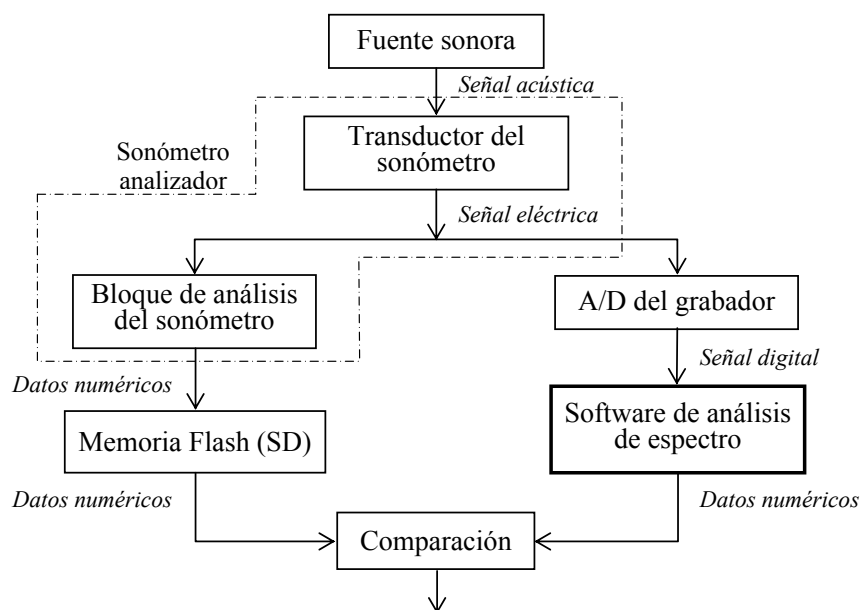
<sup>110</sup> También existen otras aproximaciones, como la de Buus (Buus, 1997), que si bien tiene cierto error con respecto a la de Zwicker tiene la ventaja de que se obtiene fácilmente la inversa.

### 8.9 Contraste de algoritmos para uso en mediciones

Las mediciones acústicas, como sucede con toda medición, descansan en el uso de instrumental y equipamiento que cumple con determinadas normas técnicas. Ello, juntamente con la verificación del estado de calibración es lo que garantiza que los resultados de la medición tendrán una incertidumbre conocida, compatible, además, con el uso al que están destinados.

Cuando en lugar de utilizar un instrumento diseñado específicamente para un tipo de medición (por ejemplo, el análisis de espectro) se integra un sistema conformado por varias piezas originalmente no previstas para cumplir esa función, será imprescindible realizar un proceso de validación para garantizar que el sistema cumple con la misma normativa. Éste es el caso cuando se integra al sistema una componente de software.

En la figura 102 se comparan ambos procesos de medición. Se distinguen las señales físicas (acústica y eléctrica), de la señal digital y de los datos numéricos abstractos que representan el resultado de la medición. La señal a grabar grabada se obtiene de la salida del propio analizador.

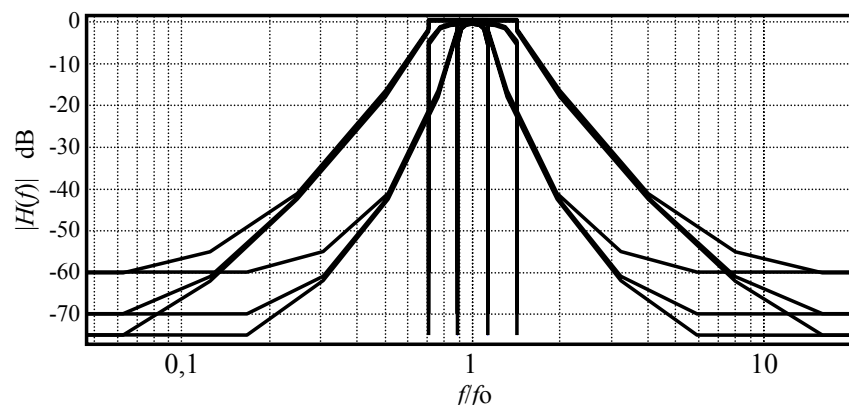


**Figura 102.** Diagrama de bloques del proceso de medición utilizando (a) un sonómetro analizador (izquierda) y (b) una cadena de medición formada por el transductor, el grabador digital y el software de análisis (derecha).

#### 8.9.1 Norma IEC 61260

Los filtros de banda de fracción de octava destinados a mediciones acústicas se encuentran normalizados por la norma IEC 61260. En la misma se dan, entre otras cosas, las plantillas que debe satisfacer la respuesta en frecuencia de uno de tales filtros, es decir, los límites inferior y superior de la respuesta en frecuencia. Existen tres clases, descritas como 0, 1 y 2 en orden decreciente de exigencias. En las figuras 13 a 16 del capítulo 1 se vieron las plantillas para los filtros de octava y tercio de octava en distintas

escalas. En la figura 103 se comparan las plantillas de filtros de tercio de octava con las de octava. Todas las plantillas son verificadas teóricamente por filtros pasabanda continuos de Butterworth de tercer orden.



**Figura 103.** Plantillas para filtros de tercio de octava y de octava según IEC 61260. Vista global.

### 8.9.2 Contraste de algoritmos

La idea general es comparar los algoritmos de análisis de espectro en sus diversas variantes con los resultados obtenidos para las mismas mediciones realizadas con un analizador de espectro comercial, el Brüel & Kjaer 2250 con el módulo de software BZ-7223 de análisis de bandas de octava y tercio de octava instalado. Este instrumento posee tanto una salida como una entrada de señal eléctrica. El instrumento se puede configurar para que analice la señal acústica o bien la eléctrica.

Si bien el instrumento posee una cadena electroacústica y un procesador digital de señal, se excluye del análisis el micrófono y su preamplificador. Sólo se analiza la cadena desde la entrada de señal eléctrica hasta el indicador numérico. Esto es porque el algoritmo a ensayar pretende sustituir sólo al bloque de análisis del sonómetro y no al sonómetro completo. En otras palabras, tanto el bloque de análisis como el software analizan la señal ya transducida, preamplificada y digitalizada. En el caso del software se agrega la conversión analógica-digital realizada por el grabador. Aunque ciertamente podría influir en el resultado de la medición, su análisis no es objeto de esta sección.<sup>111</sup> En la figura 104 se muestra el diagrama de bloques del proceso de comparación.

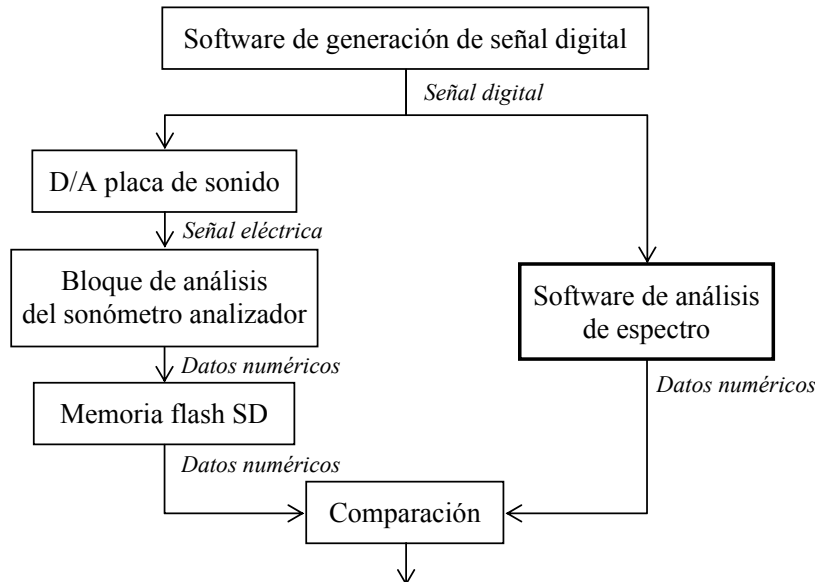
### 8.9.3 Procedimiento

El procedimiento básico consiste en obtener la respuesta ante tonos puros de diversas frecuencias y amplitud constante. Debido a que el ensayo a realizar debe poder aplicarse tanto a un instrumento que recibe señal analógica como a un algoritmo que recibe señal digital, es necesario partir de una misma señal digital (generada por computadora) y, en el primer caso, convertirla en analógica de la manera más transparente posible. Ello se logra mediante la placa de sonido de la computadora. Si la placa fuera ideal, sólo contribuiría con una constante de conversión independiente de la frecuencia:

<sup>111</sup> No obstante, se ha constatado que los efectos de la no linealidad y la respuesta en frecuencia de un buen grabador digital son despreciables (Miyara et al., 2010).

$$x(kT) = K_{D/A, \text{ placa}} x_d(k), \quad (8.70)$$

donde  $x_d(k)$  es la señal digital y  $K_{D/A, \text{ placa}}$ , la constante del conversor D/A. Pero dado que no es ideal,  $K_{D/A, \text{ placa}}$  depende de la frecuencia y, por consiguiente, es necesario determinar su respuesta en frecuencia para luego compensarla.



**Figura 104.** Diagrama de bloques del proceso de comparación entre la medición de espectro mediante el sonómetro analizador y la realizada por software.

#### 8.9.4 Calibración de la placa de sonido

La respuesta de la placa se determina utilizando el sonómetro con ponderación Z (plana). Se parte de las siguientes premisas:

- La placa de sonido es lineal en el rango de niveles en el que se va a utilizar. Esto implica desactivar cualquier tipo de compresión o control automático de ganancia.
- La respuesta en frecuencia de la placa no necesariamente es plana.
- La respuesta en frecuencia eléctrica del sonómetro en ponderación Z es plana.

Se genera una señal digital senoidal con amplitud independiente de la frecuencia y con ella se alimenta la placa. Para ello hay tres opciones: generar un archivo .wav y reproducirlo con un cualquier software de reproducción de sonido (deshabilitando cualquier ecualización), generar y reproducir el archivo de sonido con un editor de audio digital (por ejemplo Audacity, Cool Edit, Adobe Audition, Sound Forge, Goldwave) o generarlo y reproducirlo desde un software matemático con función de reproducción (por ejemplo Octave, Scilab, Matlab). La salida de línea de la placa, con máxima ganancia, se aplica a la entrada eléctrica del sonómetro en ponderación Z (plana) y se mide  $L_{p, \text{ placa}}(f_k)$  para  $K$  valores de frecuencia que cubran la totalidad del espectro audible. Si

bien este valor está expresado como nivel de presión sonora, la referencia eléctrica no tiene una interpretación definida.

Se obtiene la respuesta en frecuencia de la placa, normalizada con respecto a  $f_{\text{ref}} = 1000 \text{ Hz}$ :

$$C_{\text{placa}}(f_k) = L_{\text{p, placa}}(f_k) - L_{\text{p, placa}}(f_{\text{ref}}). \quad (8.71)$$

Esto permite referir las respuestas a la respuesta en 1000 Hz. Esta respuesta, medida en un conjunto discreto de frecuencias, puede interpolarse mediante spline cúbica para obtener la respuesta en otras frecuencias que serán necesarias en las pruebas siguientes.

### 8.9.5 Constatación de la respuesta del analizador

Se configura el analizador en el modo de análisis de espectro por bandas de octava (o de tercio de octava). Se selecciona una banda centrada en  $f_{on}$  y se mide la respuesta  $L_p(f_{nh})$  para varias frecuencias  $f_{nh}$  que abarquen desde al menos 4 octavas antes hasta 4 octavas después de  $f_{on}$ . La respuesta en frecuencia de dicha banda será

$$H_n(f_{nh}) = L_p(f_{nh}) - C_{\text{placa}}(f_{nh}). \quad (8.72)$$

Dado que los juegos de frecuencias  $f_{nh}$  se seleccionan para cada banda densificando los valores en la banda de paso y en la zona de transición, el valor de la corrección  $C_{\text{placa}}(f_{nh})$  debe obtenerse por interpolación:

$$C_{\text{placa}}(f_{nh}) = \text{spline}((f_k, C_{\text{placa}}(f_k))_{k=1, \dots, K}, f_{nh}), \quad (8.73)$$

donde  $(f_k, C_{\text{placa}}(f_k))_{k=1, \dots, K}$  son los  $K$  pares (frecuencia, respuesta de la placa) medidos.

### 8.9.6 Contraste de diversos algoritmos

Una vez obtenida la respuesta del analizador normalizado, se procede a ensayar los diversos algoritmos. En este caso las pruebas experimentales se realizan enteramente por software. Para ello se generan tonos puros de diversas frecuencias y amplitud fija y se determina la respuesta de los diversos algoritmos.

El primer algoritmo ensayado fue el filtrado FFT, según detalles vistos anteriormente. El segundo algoritmo fue el filtrado IIR con aproximación de Butterworth digital de tercer y cuarto orden. Ésta se obtuvo mediante la técnica de *pre-warping* (predistorsión del eje de frecuencia) y transformación bilineal. Las funciones de transferencia de tercer orden en transformada  $z$  son, para banda de octava y de tercio de octava, respectivamente,

$$H(z) = \frac{0,0338 - 0,1013z^{-2} + 0,1013z^{-4} - 0,0338z^{-6}}{1 - 1,7009z^{-1} + 2,415z^{-2} - 1,9683z^{-3} + 1,4337z^{-4} - 0,5565z^{-5} + 0,1890z^{-6}}, \quad (8.74)$$

$$H(z) = \frac{0,0018 - 0,0054z^{-2} + 0,0054z^{-4} - 0,0018z^{-6}}{1 - 2,2685z^{-1} + 4,1953z^{-2} - 4,2089z^{-3} + 3,5187z^{-4} - 1,593z^{-5} + 0,5889z^{-6}}, \quad (8.75)$$

mientras que las de cuarto orden son, respectivamente

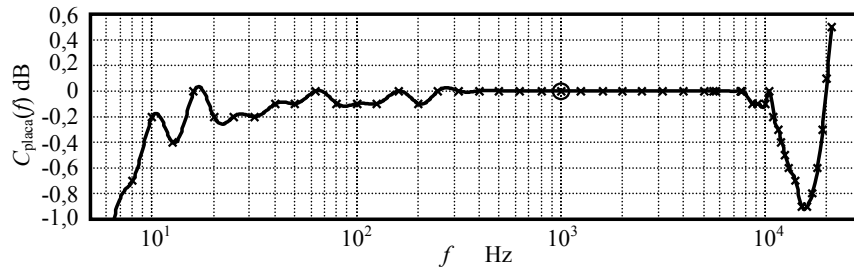
$$H(z) = \frac{0,011121 - 0,044485z^{-2} + 0,066728z^{-4} - 0,044485z^{-6} + 0,011121z^{-8}}{1 - 2,2767z^{-1} + 3,9032z^{-2} - 4,2818z^{-3} + 3,9331z^{-4} - 2,5225z^{-5} + 1,3424z^{-6} - 0,4396z^{-7} + 0,1135z^{-8}} \quad (8.76)$$

$$H(z) = \frac{0,00021965 - 0,0008786z^{-2} + 0,0013179z^{-4} - 0,0008786z^{-6} + 0,00021965z^{-8}}{1 - 3,0301z^{-1} + 6,7628z^{-2} - 9,3175z^{-3} + 10,2696z^{-4} - 7,8390z^{-5} + 1,7868z^{-6} - 1,8025z^{-7} + 0,5005z^{-8}} \quad (8.77)$$

El último algoritmo es la obtención del espectro por clasificación de líneas espectrales a partir de la FFT (Sección 2.2). El filtrado FFT (Accolti et al., 2009 y 2009b) y la clasificación de líneas espectrales (Accolti et al., 2008) se han implementado en trabajos anteriores.

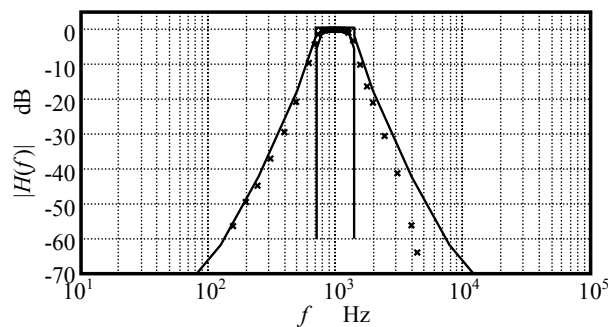
### 8.9.7 Resultados

Se trabajó con el sonómetro con analizador de espectro Brüel & Kjaer 2250. La placa de sonido fue una Creative Audigy SE. En la figura 105 se muestra la respuesta de la placa obtenida con el procedimiento descrito. Las fluctuaciones en alta frecuencia se deben al filtro de reconstrucción del convertor D/A y se minimizan utilizando una tasa de muestreo de 96000 Hz.

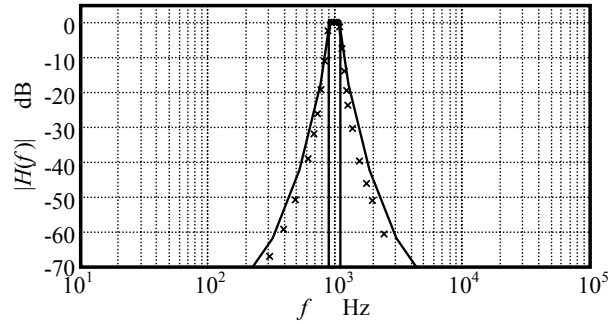


**Figura 105.** Respuesta en frecuencia de una placa Creative Audigy SE con  $f_s = 96000$  Hz. Las cruces indican los valores medidos y, la curva, la respuesta obtenida por interpolación spline.

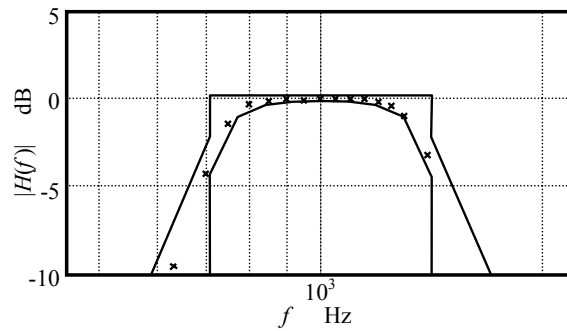
En las figuras 106 a 109 se muestran las respuestas de los filtros de banda de octava y de tercio de octava de 1 kHz para el B&K 2250



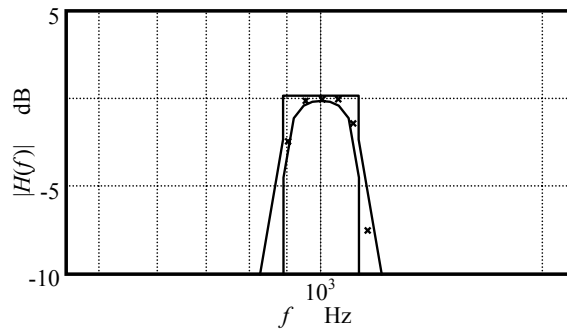
**Figura 16.** Comparación de la respuesta del analizador B&K 2250 con la plantilla de banda de octava de 1 kHz Clase 0 según IEC 61260.



**Figura 107.** Comparación de la respuesta del analizador B&K 2250 con la plantilla de banda de tercio de octava de 1 kHz Clase 0 según IEC 61260.



**Figura 108.** Comparación de la respuesta del analizador B&K 2250 con la plantilla de banda de octava de 1 kHz para Clase 0 según IEC 61260. Detalle de la banda de paso.

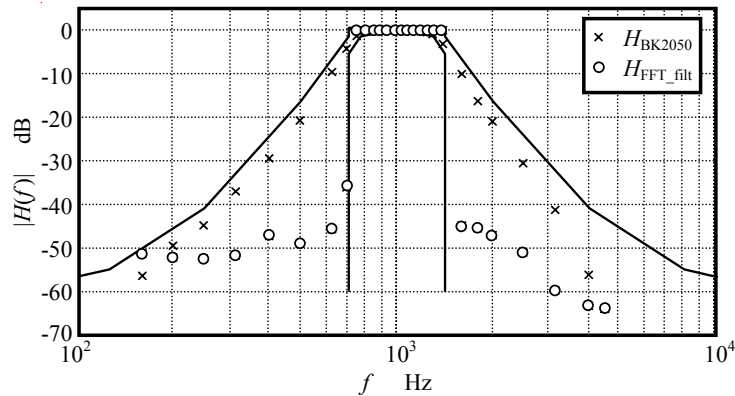


**Figura 109.** Comparación de la respuesta del analizador B&K 2250 con la plantilla de banda de tercio de octava de 1 kHz para Clase 0 según IEC 61260. Detalle de la banda de paso

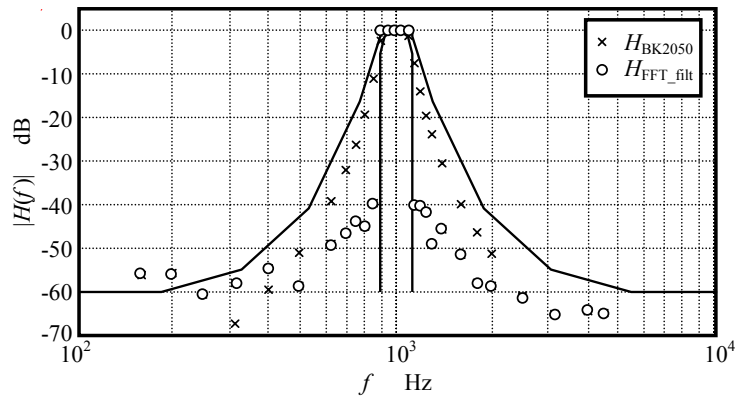
En la figura 110 se muestra la respuesta de un filtro FFT de banda de octava centrado en 1 kHz, con  $N = 8192$  y ventana filtrante ideal (es decir, vale 1 dentro de la banda de paso y 0 fuera de ella). Se compara la respuesta con la del analizador normalizado y con las plantillas correspondientes a la norma IEC 61260 clase 2. En la figura 111 se



muestra la respuesta del filtro FFT de banda de tercio de octava, en este caso con  $N = 65536$ . En ambos casos la respuesta es muy buena dentro de la banda de paso pero empeora hacia las bajas frecuencias.



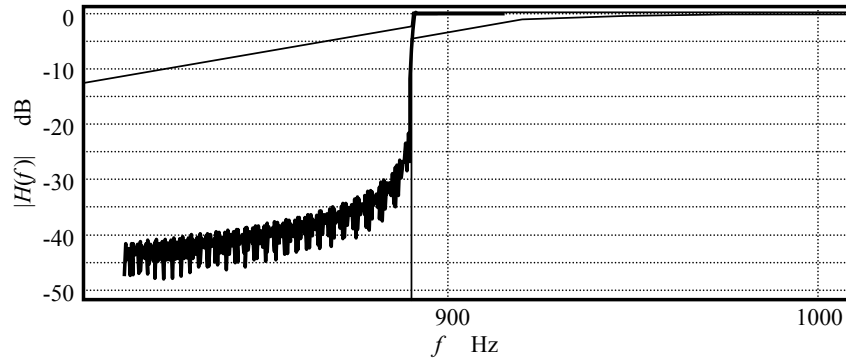
**Figura 110.** Comparación de la respuesta del filtro FFT ( $N = 8192$ ) con la plantilla de banda de octava de 1 kHz para Clase 2 según IEC 61260. Para comparación se muestra la respuesta del B&K 2250.



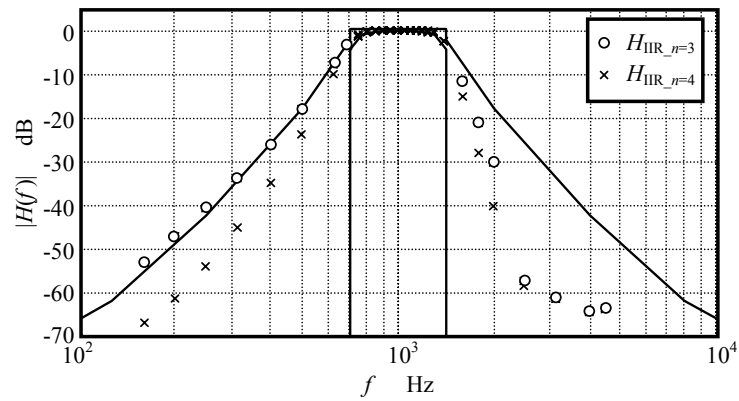
**Figura 111.** Comparación de la respuesta del filtro FFT ( $N = 65536$ ) con la plantilla de banda de tercio de octava de 1 kHz para Clase 2 según IEC 61260. Para comparación se muestra la respuesta del B&K 2250.

En la figura 112 se ha realizado un trazado más minucioso de la respuesta en frecuencia de este filtro en las proximidades del borde de la banda pasante. Se aprecia un comportamiento lobular a la vez que una transición muy rápida en la banda de atenuación.

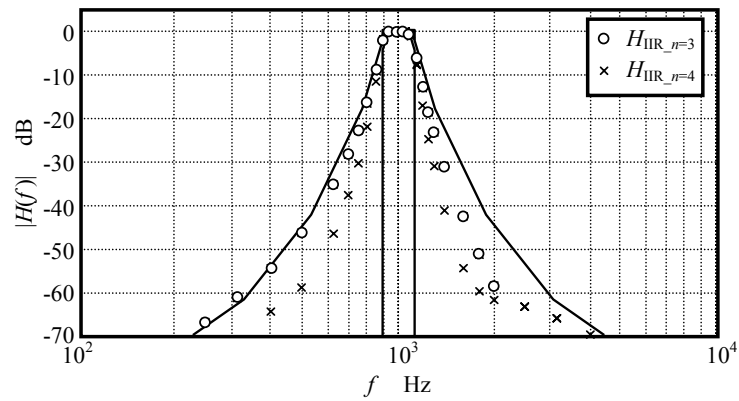
En las figuras 113 y 114 se muestra la respuesta de los filtros IIR de Butterworth de orden 3 y 4 para las bandas de octava y de tercio de octava centradas en 1 kHz. Se observa que los filtros de tercer orden cumplen marginalmente con los requerimientos de la Clase 0 en baja frecuencia y los filtros de cuarto orden los satisfacen holgadamente.



**Figura 112.** Detalle del comportamiento del filtro FFT de tercio de octava ante una sucesión de frecuencias alrededor de uno de los extremos de la banda de paso.



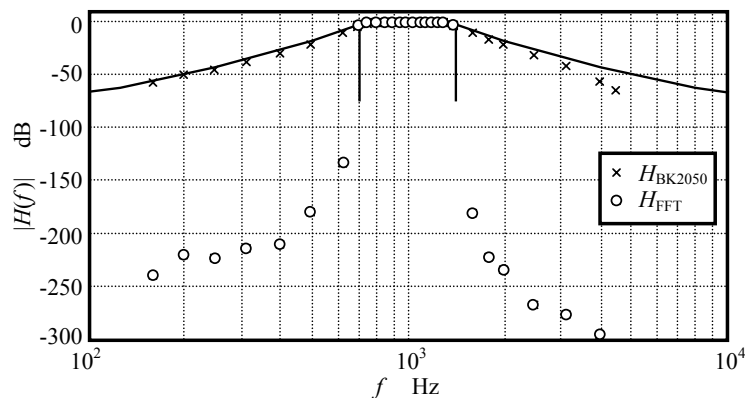
**Figura 113.** Comparación de la respuesta del filtro IIR (Butterworth orden 3 y 4) con la plantilla de banda de octava de 1 kHz para Clase 0, IEC 61260.



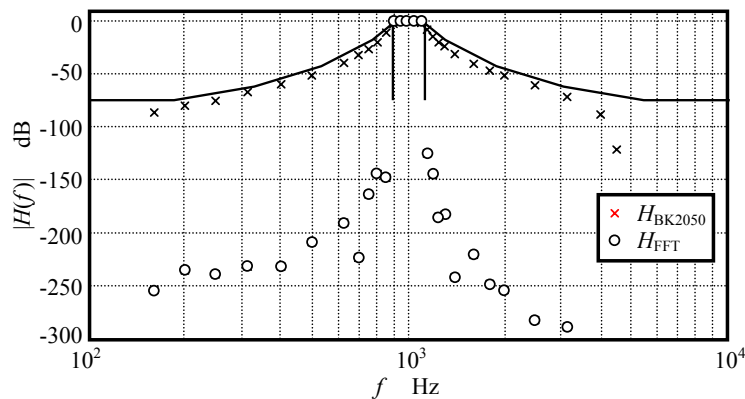
**Figura 114.** Comparación de la respuesta del filtro IIR (Butterworth de orden 3 y 4) con la plantilla de banda de tercio de octava de 1 kHz para Clase 0 según IEC 61260.

Es posible implementar un filtro híbrido que se comporte como de tercer orden en alta frecuencia y de cuarto orden en baja frecuencia, con lo que se lograría una respuesta más simétrica, aunque la ventaja real sería una ligera reducción del tiempo de cómputo.

En las figuras 115 y 116 se grafican los resultados del análisis de espectro por FFT y clasificación de líneas. Se observa una caída muy abrupta fuera de la banda de paso que no se explica por la atenuación de 58 dB del primer lóbulo lateral de la ventana de Blackman. Una posible explicación es que cada línea FFT dentro de la banda se comporta como un filtro con respuesta lobular similar a la de la figura 95 (sólo que con lóbulos laterales menores por tratarse de la ventana de Blackman, véase la Tabla 8.1). Una frecuencia fuera de la banda de interés cae en diversas posiciones relativas con respecto al lóbulo central de cada uno de esos filtros, por lo cual es afectada por una atenuación y una fase cuasi aleatorias. La superposición de gran número de estas respuestas, todas de igual frecuencia pero de amplitudes y fases aleatorias produce estadísticamente una tendencia a la cancelación, de allí la baja respuesta observada.



**Figura 115.** Comparación de la respuesta del analizador por clasificación de líneas FFT ( $N = 4096$ , ventana de Blackman) con la plantilla de banda de octava de 1 kHz para Clase 0 según IEC 61260. Se compara con la respuesta del B&K 2250.



**Figura 116.** Comparación de la respuesta del analizador por clasificación de líneas FFT ( $N = 4096$ , ventana de Blackman) con la plantilla de banda de tercio de octava de 1 kHz para Clase 0 según IEC 61260. Se compara con la respuesta del B&K 2250.

En la Tabla 8.3 se da una comparativa de los tiempos de cómputo requeridos por cada algoritmo en relación con la duración de la señal analizada utilizando una computadora con procesador Intel Centrino core duo T2350 con una velocidad de 1,86 GHz con sistema operativo de 32 bits y 1 Gb de memoria RAM. Se observa que los filtros IIR son los que requieren menor tiempo de cómputo, incluso teniendo en cuenta que se requiere un análisis de todas las bandas. Los filtros FFT son ventajosos en las proximidades de la banda de paso. Los métodos basados en FFT y clasificación de líneas espectrales requieren la determinación de la FFT una sola vez por cada cuadro de análisis pero no realizan un filtrado de la señal.

**Tabla 8.3.** Tiempos de cálculo de cada método para una banda centrada en 1000 Hz

Método	Duración de la señal	Tiempo de cálculo 1/1 octava	Tiempo de cálculo 1/3 octava
Filtro FFT	60 s	79,5 s	73,5 s
Filtro IIR n=3	60 s	0,8 s	0,8 s
Filtro IIR n=4	60 s	0,8 s	0,8 s
FFT	47,5 s	46,7 s	46,6 s

### 8.9.8 Conclusión

Del análisis realizado se puede concluir que los algoritmos para análisis de espectro descriptos, complementados con un instrumento sencillo capaz de proveer una señal de audio calibrada, permiten sustituir el uso de instrumentos de medición más costosos en el análisis de espectro de bandas de octava y tercio de octava.<sup>112</sup>

<sup>112</sup> Esta sección se basa en (Miyara et al., 2009), trabajo realizado dentro de un proyecto de investigación financiado por la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT - PICT N° 38109)

## Capítulo 9

# ENSAYO DE GRABADORES DIGITALES

### 9.1 Introducción<sup>113</sup>

La propuesta desarrollada en este libro involucra la adquisición sistemática de señales mediante un grabador digital y su posterior análisis por software en una computadora de propósito general. En el capítulo anterior se detallaron los ensayos realizados sobre el software, suponiendo que la señal registrada reflejaba fielmente la señal eléctrica que entrega el sonómetro con salida calibrada.

Existe, sin embargo, una legítima preocupación en cuanto a si la calidad de la señal digital es suficiente para cumplir con la normativa. Normas internacionales tales como las ISO 1996-2:2007 (y su versión argentina, la IRAM 4113-2) especifican que para que una pieza del equipo sea idónea para propósitos de medición, el sistema completo de instrumentación que la incluye deberá cumplir plenamente con la norma IEC 61672 sobre sonómetros. Esto significa, en particular, que los límites de tolerancia para la respuesta en frecuencia son muy estrictos. Además, se incluyen otras condiciones sobre ruido auto-generado, linealidad y respuesta transitoria.

Con el fin de verificar el cumplimiento de tales requisitos, se realizó una serie de ensayos sobre tres unidades de un grabador digital para uso general de gran difusión comercial, el Zoom H4, según se describe en las secciones siguientes. En principio, este tipo de ensayos debería realizarse sobre cualquier unidad de cualquier marca y modelo que se destine a este uso, ya que se trata de líneas de producción destinadas al usuario general y no a cubrir necesidades metrológicas, por lo tanto los diseños o incluso especificaciones podrían modificarse de una partida a la siguiente sin previo aviso.

### 9.2 Especificaciones del grabador digital

El grabador digital Zoom H4 permite el uso de tres tasas de muestreo: 44,1 kHz, 48 kHz, y 96 kHz, y dos resoluciones: 16 bit y 24 bit. Aunque el muestreo a 96 kHz puede mejorar la respuesta en frecuencia, en la mayoría de los casos el uso de una resolución de 24 bit es innecesaria y sólo incrementará el tamaño del archivo, ya que la mejor resolución no está acompañada por una mejora real en el rango dinámico. El formato de grabación puede ser .WAV o .MP3. Este último no se recomienda porque usa una compresión de audio con pérdidas que podría suprimir componentes espectrales significativas. El nivel de la señal de entrada de línea se especifica como -10 dBm. Sorprendentemente, no se provee información con respecto a la relación señal-ruido ni a la distorsión (Zoom Corporation, s/f).

El grabador tiene dos entradas combinadas (coaxiales) XLR balanceada y de 1/4" no balanceada. Tiene también dos micrófonos montados en una configuración X-Y, pero no hay posibilidad de calibrarlos y no se garantiza la respuesta en frecuencia ni la estabilidad a largo plazo de manera que su uso no está recomendado para propósitos de medición.

---

<sup>113</sup> Este capítulo está basado en (Miyara et. al., 2010a y 2010b)

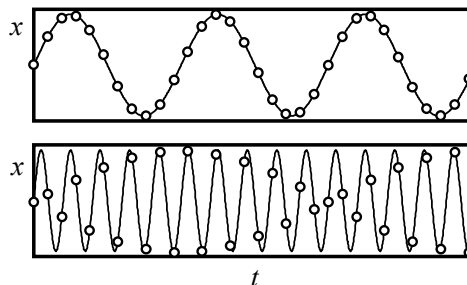
### 9.3 Ensayos

Se han realizado varias pruebas sobre el grabador en lo que concierne a su respuesta en frecuencia, respuesta transitoria, ruido y linealidad.

#### 9.3.1 Respuesta en frecuencia

Para analizar la respuesta en frecuencia se emplearon dos tipos de señales: tonos puros y barridos de frecuencia. Los tonos son las señales ideales de prueba, pero se necesitarían muchos de ellos para detectar posibles rizados en las cercanías de la frecuencia de Nyquist (la mitad de la tasa de muestreo). En cambio, los barridos permiten cubrir todas las frecuencias, siempre que la frecuencia instantánea de la señal de prueba se incremente a una tasa suficientemente baja como para evitar un excesivo ensanchamiento espectral.

Cuando la frecuencia de entrada se acerca al límite de Nyquist, la secuencia discreta tiene poco más de dos muestras por ciclo. La forma de onda deja de ser evidente a partir de la observación de las muestras, que parecen agruparse visualmente en patrones de tipo Moiré (un aliasing visual), tal como se muestra en la figura 117. Sin embargo, dado que la señal de entrada ha pasado por un filtro antialiasing antes de muestrearse, es teóricamente posible recuperar la señal original (vía interpolación sinc, véase apéndice 10).

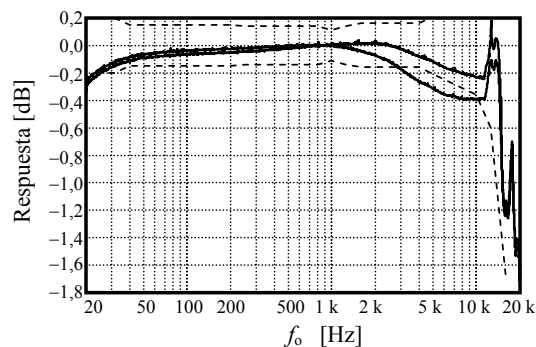


**Figura 117.** Patrón de muestreo para tonos muy por debajo (figura superior) y cercanos (figura inferior) al límite de Nyquist.

Nos interesa solamente el valor eficaz (RMS) de la señal grabada. Se podría, por ejemplo, sobremuestrear y aplicar un filtro de suavizado. Sin embargo, esto es innecesario ya que el valor eficaz se puede calcular a partir de las muestras originales, siempre que se consideren varios períodos de la señal. Esto es cierto aun para las secuencias con frecuencia cercana al límite de Nyquist. Otro enfoque consiste en detectar el máximo de las muestras en un intervalo alrededor del instante deseado, lo cual se aproxima al pico de la onda senoidal ideal. De este máximo se detecta el valor eficaz buscado.

Se verifica por simulación que ambos métodos son adecuados. Sin embargo, el enfoque del máximo requiere menos muestras para alcanzar el valor correcto con una tolerancia dada. Esto es más importante en el caso de barridos de frecuencia, pues se debe asegurar la detección del valor eficaz antes de que haya un cambio significativo en la frecuencia. Se usaron barridos logarítmicos de 100 s. A fin de mantener la resolución temporal (y por ende frecuencial), se adoptaron intervalos de alrededor de 200 muestras, excepto en el extremo de bajas frecuencias donde se trabajó con dos períodos. En algunos casos, especialmente en altas frecuencias, se encontraron máximos de pequeña amplitud en secuencias largas. Éstos no representan disminución alguna en la respuesta en frecuencia y se corrigen fácilmente mediante la quita de muestras que representen una disminución brusca de la respuesta con respecto a las frecuencias vecinas en un umbral previamente asignado en 0,02 dB.

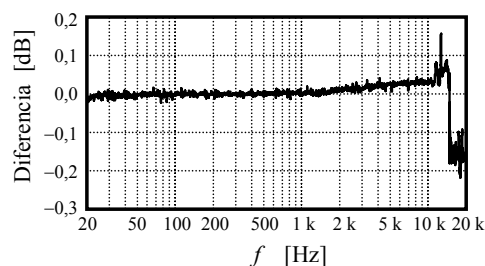
La mejor y la peor respuesta se ilustran en la figura 118. El aspecto algo ruidoso se debe a una pequeña distorsión del sintetizador de señales, al ruido eléctrico y a la influencia del ruido en el método de detección de amplitud. Dado que el mencionado ruido posee niveles muy inferiores a los valores principales, prácticamente no influye.



**Figura 118.** Respuesta en frecuencia del mejor y del peor grabador digital Zoom H4 obtenido con un barrido en frecuencia sintetizado con el SRS DS 345 a  $T = 24\text{ }^{\circ}\text{C}$ . La tasa de muestreo es 44,1 kHz. Las líneas de trazos muestran 1/10 de los límites de tolerancia superior e inferior mencionados en el Estándar Internacional IEC 61672-1 para la respuesta en frecuencia de un sonómetro clase 1.

Para una dada unidad de H4 la respuesta es esencialmente invariable en sucesivas las pruebas. Con fines comparativos, también se muestra un décimo de las tolerancias de la respuesta en frecuencia especificado en la Norma Internacional IEC 61672-1 para un instrumento clase 1.

Cuando la temperatura aumenta de  $24\text{ }^{\circ}\text{C}$  a  $40\text{ }^{\circ}\text{C}$ , se observa una pequeña diferencia en la respuesta en frecuencia, como se muestra en la figura 119.

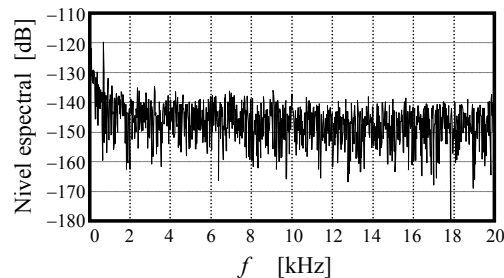


**Figura 119.** Diferencia entre las respuestas en frecuencias del Zoom H4 obtenidas con  $T = 40\text{ }^{\circ}\text{C}$  y  $T = 24\text{ }^{\circ}\text{C}$ .

Como se puede apreciar, la respuesta en frecuencia es aceptable dado que aún el peor caso está por debajo de los límites de tolerancia. La caída brusca en las proximidades de 16 kHz se debe a la alta sensibilidad causada por la pendiente muy negativa de la respuesta en esa frecuencia.

### 9.3.2 Ruido

El ruido interno está compuesto por ruido analógico (debido a fuentes misceláneas como ruido térmico y ruido de emisión) y ruido digital incluyendo el de discretización y *dithering*. Con el fin de medir el ruido térmico, se grabó con la entrada cortocircuitada. Al archivo .WAV se lo procesó vía software de PC para obtener su valor eficaz y analizarlo vía FFT (transformada rápida de Fourier). También por medio de PC se le aplicó ponderación A al ruido estudiado. La figura 120 muestra la densidad espectral del ruido detectado en la unidad H4 más ruidosa, con tasa de muestreo de 44,1 kHz y resolución de 16 bit. A pesar de que hay diferencias entre las unidades analizadas, tasas de muestreo o resoluciones, éstas parecen ser aleatorias y no se detecta ningún efecto significativo producido por alguno de estos factores.



**Figura 120.** Ejemplo del peor caso de nivel espectral de ruido del Zoom H4 referenciado a la máxima salida digital. La señal se grabó con la entrada cortocircuitada y la resolución adoptada fue de 16 bit.

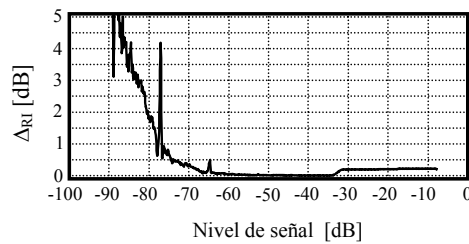
El peor caso de relación señal-ruido no ponderado fue de 91,7 dB mientras que la relación señal / ruido ponderada A fue de 95,9 dBA, en ambos casos referidos a una onda senoidal sin distorsión de máxima amplitud. Esto significa que si se graba un tono de calibración de 94,0 dB a máxima amplitud, se pueden grabar satisfactoriamente señales 20 dB menores con un error debido al ruido menor a 0,1 dB. La mayoría de los sonómetros poseen un ruido inherente mínimo del orden de 20 dB.

### 9.3.3 Linealidad

Se evalúa la linealidad de dos formas diferentes. Primero, comparando su efecto en la señal grabada para varios niveles de entrada. Ésta es la prueba estándar especificada en la norma IEC 61672. En la segunda forma, se calcula la distorsión total armónica (THD) para varios niveles de entrada. Se realizaron ambas pruebas usando un tono modulado por un leve decaimiento exponencial (de duración  $T = 80$  s; constante de tiempo  $\tau = 8,69$  s; rango dinámico: 80 dB).

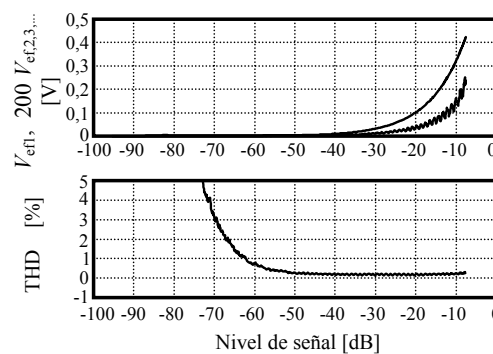
En el primer caso (figura 121), se calculó la envolvente de la energía con un filtro digital de primer orden con  $\tau = 0,05$  s y se la comparó en escala logarítmica con la respuesta de amplitud teórica sin distorsión. Se la expresó como la diferencia  $\Delta_{RI}$  entre la respuesta real y una respuesta ideal. Se puede observar un error inesperado cercano a los -32 dB, donde hay un cambio abrupto en el error de alrededor de 0,2 dB. Esto no afecta el cumplimiento de la normativa IEC, ya que la máxima diferencia después de un cambio de nivel entre 1 dB y 10 dB debe ser menor que  $\pm 0,6$  dB. Para niveles inferiores a -70 dB la diferencia  $\Delta_{RI}$  se incrementa rápidamente como consecuencia de la distorsión digital. Además, hay un efecto espurio del ruido intrínseco.





**Figura 121.** Respuesta real menos respuesta ideal representado como  $DR = 20 \log(V_{RMS-real} / V_{RMS-ideal})$ .

En el segundo caso (figura 122), se aisló la componente fundamental de los restantes armónicos por medio de un filtro FFT multibanda de alta selectividad con el fin de reducir el efecto del ruido residual de banda ancha. Luego se calcularon y compararon los valores eficaces de estas componentes.



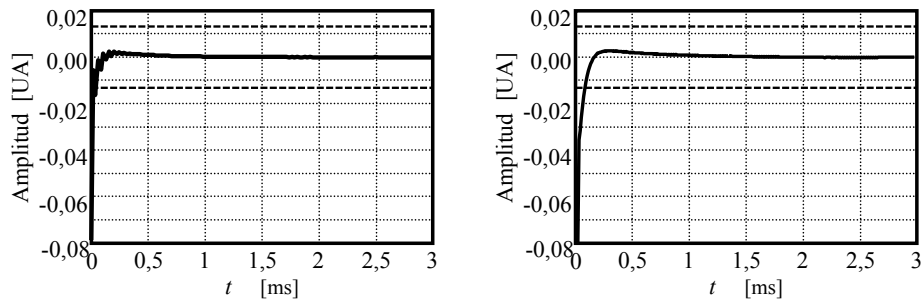
**Figura 122.** Análisis de linealidad. Arriba: Primer armónico y armónicos superiores (multiplicados por un factor de 200 a fines comparativos). Abajo: Distorsión total armónica. En ambos casos la frecuencia es de 1 kHz. La referencia de audio digital es la señal digital máxima.

Como se puede ver, los armónicos 2, 3, ... disminuyen más rápidamente que la fundamental al disminuir el nivel de la señal, por lo cual la THD decrece suavemente hasta alrededor del 0,02 %. Si sigue disminuyendo el nivel de la señal, la THD comienza a incrementarse nuevamente debido a la distorsión digital en niveles muy bajos. Para niveles inferiores a -60 dB la THD se incrementa rápidamente y en -73 dB (referido al máximo nivel digital), alcanza un 5 %, que equivale a un error de 0,42 dB.

En el extremo superior del espectro (por encima de 10 kHz) se ha observado también un efecto de aliasing que indicaría que hay distorsión después del filtro antialias. Este efecto desaparece completamente para niveles de señal por debajo de -20 dB respecto al máximo digital.

### 9.3.4 Respuesta transitoria

Se evaluó la respuesta transitoria usando onda cuadrada. Se encontró que la respuesta en frecuencias altas no es exactamente la misma en distintas unidades del modelo H4, pero se halló una cota superior del tiempo de establecimiento del 5 %. Las respuestas son oscilantes sólo para la tasa de muestreo de 44,1 kHz como se muestra en la figura 123. Se determinó que el tiempo de establecimiento es a lo sumo 104  $\mu$ s para 96 kHz y 136  $\mu$ s para 44,1 kHz.



**Figura 123.** Respuesta al escalón para 44,1 kHz (izquierda) y 96 kHz (derecha). En línea de trazos, los límites del 5 %.

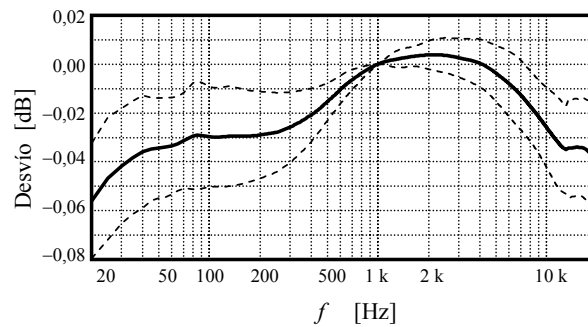
### 9.3.5 Incertidumbre

Con el fin de cumplir con la norma IEC 61672, el grabador debería exhibir una incertidumbre de a lo sumo 0,1 dB. Se generaron las señales de prueba usando un generador de funciones sintéticas Stanford Research Systems DS 345. Según sus especificaciones (Stanford Research Systems, s/f), la precisión de la onda senoidal es de solamente  $\pm 0,4$  dB. A pesar de que el uso del grabador requiere la grabación de una señal calibrada, esto carecería de importancia siempre que la respuesta en frecuencia sea suficientemente plana. Sin embargo, no se encontró información concerniente a cuán plana es la respuesta en frecuencia, razón por la cual se la corroboró usando un multímetro digital Hewlett-Packard HP 974A (Hewlett Packard, 1995), cuya precisión del valor eficaz verdadero se brinda en la Tabla 1 (incluyendo porcentaje de lectura + cuentas). Como se ilustra en la figura 124, el desvío con respecto a la respuesta plana del generador DS 345 llega a ser a lo sumo  $\pm 0,045$  dB (con 96 % de confianza) que, al ser combinado con la precisión del HP 974A da una incertidumbre resultante de 0,103 dB. Se puede reducir esta incertidumbre a 0,095 dB tomando en cuenta el desvío sistemático del DS 345 en cada frecuencia.

Se especifica la distorsión del DS345 como  $-55$  dBc. Esto representa un error máximo de  $20 \log(1+10^{-55/20}) = 0,015$  dB cuando los picos de las armónicas están en fase con la señal principal.

**Tabla 9.1.** Precisión del valor eficaz verdadero del multímetro HP 974A según el rango de frecuencia.

Rango de frecuencia (Hz)	Precisión (dB)
20 - 50	$\pm 0,092$
50 - 10 000	$\pm 0,049$
10 000 - 20 000	$\pm 0,093$



**Figura 124.** Desvío de la respuesta en frecuencia plana del generador SRS DS345. En líneas de trazos se indican los intervalos de confianza del 96 % ( $\pm 2\sigma$ ). Se supone calibración exacta a 1 kHz.

### 9.3.6 Conclusiones

Varios parámetros muestran que las unidades evaluadas del modelo H4 son adecuadas para grabar señales de audio con propósitos de medición. No hay diferencias significativas entre los distintos formatos de grabación (16 bit ó 24 bit). Así, el ruido parece tener una cota mínima condicionada por el ruido analógico más que por el de digitalización. El modelo exhibe una cierta distorsión que causa, en el extremo de las frecuencias altas (por sobre 10 kHz), algún aliasing perceptible. Se recomienda ajustar el nivel de grabación máximo en alrededor de  $-20$  dB con el fin de minimizar la distorsión, quedando todavía un rango dinámico de unos 60 dB hasta llegar a un error del orden de 0,1 dB con respecto al máximo debido al ruido.

## Referencias<sup>114</sup>

- Accolti, Ernesto; Miyara, Federico (2008). “Combinación digital controlada de ruidos diversos” VI Congreso Iberoamericano de Acústica, FIA 2008, Pg. A090. Bs As, Argentina.
- Accolti, Ernesto; Miyara, Federico (2009). “Fast and controllable box-shaped room impulse response algorithm” XIII Reunión de Trabajo en Procesamiento de la Información y Control RPIC 2009. Rosario, Argentina.
- Accolti, Ernesto; Miyara, Federico (2009b). “Fluctuation Strength of Mixed Fluctuating sound Sources” Mecánica Computacional Vol XXVIII, Pg. 9-22. ENIEF 2009. Tandil, Argentina.
- Analog Devices. “Ultralow Noise Microphone with Bottom Port and PDM Digital Out-put”. Hoja de datos. 2012.  
[http://www.analog.com/static/imported-files/data\\_sheets/ADMP521.pdf](http://www.analog.com/static/imported-files/data_sheets/ADMP521.pdf)
- Analog Devices. “Microphone Beamforming Simulation Tool”. Software de simulación. 2013.  
<http://wiki.analog.com/resources/tools-software/microphone/beamforming-simulation-tool>
- Anón. “FFT window and overlap”. Sin fecha. Consultado 06 ago 2013.  
<http://www.katjaas.nl/FFTwindow/FFTwindow.html>
- Audacity team. “Formatos de archivo”. Sitio web del proyecto Audacity. Consultado el 02/07/13. <http://audacity.sourceforge.net/xml/>
- Bayer, R.; McCreight, E. “Organization and Maintenance of Large Ordered Indexes”. Acta Informatica, Vol. 1, Fasc. 3, 1972, pp. 173-189
- Bayer, R.; McCreight, E. “Organization and Maintenance of Large Ordered Indexes”. Mathematical and Information Sciences Report No. 20. Mathematical and Information Sciences. Laboratory Boeing Scientific Research Laboratories. July 1970.  
[http://infolab.usc.edu/csci585/Spring2010/den\\_ar/indexing.pdf](http://infolab.usc.edu/csci585/Spring2010/den_ar/indexing.pdf)
- Bell, Stephanie “Measurement Good Practice Guide No. 11 (Issue 2). A Beginner's Guide to Uncertainty of Measurement”. National Physical Laboratory. Teddington, Middlesex, United Kingdom, 2001.  
[http://resource.npl.co.uk/cgi-bin/download.pl?area=npl\\_publications&path\\_name=/npl\\_web/pdf/mgpg11.pdf](http://resource.npl.co.uk/cgi-bin/download.pl?area=npl_publications&path_name=/npl_web/pdf/mgpg11.pdf)
- Beranek, Leo. “Acoustical Measurements”. Acoustical Society of America. Cambridge, 1993
- Bergero, Federico; Miyara, Federico. “Diseño de una interfaz gráfica para ingreso de una respuesta en frecuencia variable en Audacity”. Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario. 2013. Documento interno inédito.
- Birch, Keith. “Measurement Good Practice Guide No. 36. Estimating Uncertainties in Testing. An Intermediate Guide to Estimating and Reporting Uncertainty of Measurement in Testing”. British Measurement and Testing Association. Teddington, Middlesex, United Kingdom, 2003  
<http://www.dit.ie/media/physics/documents/GPG36.pdf>
- Blu-ray Disc Association. “White Paper Blu-ray Disc™ Format - General - 3rd Edition. December, 2012”. Consultado 22 jul 2012.  
[http://www.blu-raydisc.com/Assets/Downloadablefile/White\\_Paper\\_General\\_3rd\\_Dec%202012\\_20121210.pdf](http://www.blu-raydisc.com/Assets/Downloadablefile/White_Paper_General_3rd_Dec%202012_20121210.pdf)
- Boggino, Luciano; Miyara, Federico. “Especificación de un filtro variable en el tiempo mediante una herramienta de dibujo en Audacity y su implementación”. Universidad Nacional de Rosario, FCEIA, 2013. Documento interno inédito.

<sup>114</sup> Las URL de Internet de los documentos incluidos en esta lista de referencias estaban vigentes al momento de consultarlos (aproximadamente entre enero y julio de 2013). Sin embargo, por la dinámica de la red es posible que los mismos estén transitoria o definitivamente inaccesibles en cualquier momento, sin previo aviso, o que experimenten actualizaciones o correcciones. En algunos casos los documentos pueden haber sido cambiados de ubicación en los servidores, por lo cual una búsqueda con el nombre del archivo (la cadena de caracteres después de la última barra) puede conducir a la URL actualizada. En otros casos se puede hacer una búsqueda con el título del documento.

- Boser, B.; Wooley, B. "The Design Of Sigma-Delta Modulation Analog- To- Digital Converters". IEEE Journal Of Solid-State Circuits, Vol. 23. No. 6, December 1988  
[http://www.uio.no/studier/emner/matnat/ifi/INF4420/v10/undervisningsmateriale/Boser\\_Wooley.pdf](http://www.uio.no/studier/emner/matnat/ifi/INF4420/v10/undervisningsmateriale/Boser_Wooley.pdf)
- Buus, Søren: "Auditory Masking". En Crocker, Malcolm (ed), "Encyclopedia of Acoustics". John Wiley & Sons. 1997.
- Bray, Tim; Paoli, Jean; Sperberg-McQueen, C. M.; Maler, Eve; Yergeau; François (compiladores). "Extensible Markup Language (XML) 1.0 (Fifth Edition)". 2008. Disponible en <http://www.w3.org/TR/REC-xml/>
- Brüel & Kjær. "Microphone Handbook. Brüel & Kjær Falcon™ Range of Microphone Products". BA 5105–12. Revision February 1995
- Brüel & Kjær. "Microphone Handbook - Vol. 1: Theory". Technical Documentation Brüel & Kjær BE 1447–11. Nærum, Denmark, 1996. <http://www.bksv.es/doc/be1447.pdf>
- Coalson, Josh. "FLAC - Format", 2008. Internet: <http://flac.sourceforge.net/format.html>
- COGUANOR. "Introducción a la Incertidumbre de las mediciones". Comisión Guatemalteca de Normas. S/F
- Cooley, J. W.; Tuckey, J. W. (1965). "An algorithm for the machine computation of complex Fourier Series". Math. Comp., Vol 19, pp.297-301.
- COVENIN 2552:1999. "Vocabulario internacional de términos básicos y generales de metrología"
- Cox, M. G.; Harris, P. M. "Best Practice Guide No. 6. Software Support for Metrology. Uncertainty Evaluation". NPL Report DEM-ES-011. Teddington, UK, 2006.  
[publications.npl.co.uk/npl\\_web/pdf/dem\\_es11.pdf](http://publications.npl.co.uk/npl_web/pdf/dem_es11.pdf)
- Cox, M. G.; Forbes, A. B.; Harris, P. M. "Best Practice Guide No. 11. Numerical analysis for algorithm design in metrology". Technical report, National Physical Laboratory, Teddington, UK, 2004. <http://www.npl.co.uk/ssfm/download/bpg.html#ssfmbpg11>
- Crochiere, R. (1980). A weighted overlap-add method of short-time Fourier analysis/synthesis. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-28(2):99–102.
- Crook, James. "How Noise Removal Works". Audacity Team, 2011. Visitado 27 jul 2013.  
[http://wiki.audacityteam.org/wiki/How\\_Noise\\_Removal\\_Works](http://wiki.audacityteam.org/wiki/How_Noise_Removal_Works)
- de la Rosa, José M. "Sigma-Delta Modulators: Tutorial Overview, Design Guide, and State-of-the-Art Survey". IEEE Transactions On Circuits And Systems—I: Regular Papers, Vol. 58, No. 1, January 2011  
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5672380>
- Developer Connection. "Technical Note TN1150. HFS Plus Volume Format". Apple Inc., 2004. Consultado 23 jul 2013.  
<http://dubeiko.com/development/FileSystems/HFSPLUS/tn1150.html>
- Developer Connection. "Inside Macintosh: Files". Apple Inc., 2007. Consultado 23 jul 2013.  
[http://www.dubeyko.com/development/FileSystems/HFS/inside\\_macintosh/inside\\_macintosh.htm](http://www.dubeyko.com/development/FileSystems/HFS/inside_macintosh/inside_macintosh.htm)
- ECA. "EA-4/02 Expression of the Uncertainty of Measurement in Calibration". European Cooperation for Accreditation. December 1999  
<http://www.european-accreditation.org/publication/ea-4-02-m>
- ECMA-167. "Volume and File Structure for Write-Once and Rewritable Media using Non-Sequential Recording for Information Interchange". 1997. Consultado el 23 jul 2013.  
<http://www.ecma-international.org/publications/files/ECMA-ST/Ecma-167.pdf>
- ECMA-377. "Information Interchange on Holographic Versatile Disc (HVD) Recordable Cartridges Capacity: 200 Gbytes per Cartridge. 2007.  
<http://www.ecma-international.org/publications/standards/Ecma-377.htm>  
<http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-377.pdf>
- Gade, Svend; Herlufsen, Henrik. "Use of Weighting Functions in DFT/FFT Analysis (Part I)" Brüel & Kjær. Technical Review. 1987 N° 3.  
[http://bruel.ru/UserFiles/File/Review3\\_87.pdf](http://bruel.ru/UserFiles/File/Review3_87.pdf)

- Gade, Svend; Herlufsen, Henrik. "Use of Weighting Functions in DFT/FFT Analysis (Part II)" Brüel & Kjaer. Technical Review. 1987 N° 4.  
[http://bruel.ru/UserFiles/File/Review4\\_87.pdf](http://bruel.ru/UserFiles/File/Review4_87.pdf)
- Gauss, Carl Friedrich. "Theoria interpolationis methodo nova tractata", Werke, Band 3, 265–327 (Königliche Gesellschaft der Wissenschaften, Göttingen, 1866)
- Gray, Robert M. "Oversampled Sigma-Delta Modulation". IEEE Transactions On Communications, Vol. Com-35, No. 5, May 1987  
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1096814>
- Heideman, Michael T.; Johnson, Don H.; Burrus, C. Sidney. "Gauss and the History of the Fast Fourier Transform". IEEE ASSP MAGAZINE. OCTOBER 1984. pp 14-21.  
[http://www.cis.rit.edu/class/simg320/Gauss\\_History\\_FFT.pdf](http://www.cis.rit.edu/class/simg320/Gauss_History_FFT.pdf)
- Hewlett-Packard Company. HP 974A Multimeter User's Guide. 1995.  
<http://cp.literature.agilent.com/litweb/pdf/00974-90002.pdf>
- IEC 60318-1 ed2.0 "Electroacoustics - Simulators of human head and ear - Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones." IEC, Ginebra, 31/08/2009
- IEC 61183 ed1.0 1994. "Electroacoustics - Random-incidence and diffuse-field calibration of sound level meters"
- IEC 61260:1995 "Octave-Band and Fractional-Octave-Band Filters"
- IEC 61672-1. "Electroacoustics - Sound level meters - Part 1: Specifications". International Electrotechnical Commission. 2002.
- IEC 60537. "Frequency weighting for measurement of aircraft noise (D-weighting)". International Electrotechnical Commission. 1976. Norma retirada.
- IRAM 4063-3 "Acústica. Medición del aislamiento acústico en los edificios y de los elementos de construcción. Parte 3: Medición en laboratorio del aislamiento acústico al ruido aéreo de los elementos de construcción".
- IRAM 4109-1 "Acústica. Medición de parámetros acústicos en recintos. Parte 1 – Salas de espectáculos"ISO-GUM. "Guide to the Expression of Uncertainty in Measurement, prepared by ISO Technical Advisory Group 4 (TAG 4), Working Group 3 (WG 3), October 1993.
- ISO/IEC Guide 99:2007, International vocabulary of metrology — Basic and general concepts and associated terms (VIM)
- ISO 140-3:1995 "Acoustics. Measurement of sound insulation in building element. Part 3: Laboratory measurements of airborne sound insulation of building elements"
- ISO/TR 389-5:1998, Acoustics – Reference zero for the calibration of audiometric equipment – Part 5: Reference equivalent threshold sound pressure levels for pure tones in the frequency range 8 kHz to 16 kHz.
- ISO/DIS 389-8:2004, "Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones."
- ISO 1996-2: Acoustics - Description, measurement and assessment of environmental noise - Part 2: Determination of environmental noise levels. Geneve, Switzerland, 2007.
- ISO 3382-1:2009 "Acoustics. Measurement of room parameters. Part 1: Performance Spaces".
- ISO 7196:1995 "Acoustics - Frequency-weighting characteristic for infrasound measurements"
- ISO 11904-1:2002 "Acoustics Determination of sound immission from sound sources placed close to the ear - Part 1: Technique using a microphone in a real ear (MIRE technique)"
- ISO 11904-2:2004 "Acoustics Determination of sound immission from sound sources placed close to the ear - Part 2: Technique using a manikin"
- Janssen, E.; van Roremund, A. "Look-Ahead Based Sigma-Delta Modulation". Analog Circuits and Signal Processing. Springer, 2011, Capítulo 2.  
[http://www.springer.com/cda/content/document/cda\\_downloadaddocument/9789400713864-c2.pdf](http://www.springer.com/cda/content/document/cda_downloadaddocument/9789400713864-c2.pdf)
- Jarman, David. "A Brief Introduction to Sigma Delta Conversion". Intersil, May, 1995.  
<http://www.intersil.com/content/dam/Intersil/documents/an95/an9504.pdf>

- JCGM 100:2008. "Evaluation of measurement data — Guide to the expression of uncertainty in measurement". Working Group 1 of the Joint Committee for Guides in Metrology (JCGM/WG 1).  
[http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf)
- Johnson, Richard. "Probabilidad y Estadística para Ingenieros de Miller y Freund (5ta Ed.)". Prentice Hall Hispanoamericana. Naucalpan de Juárez, México, 1997
- Kendall, M.; Stuart, A. "The advanced theory of Statistics", vol 1. Charles Griffin & Co., Londres, 1977.
- Kholodov, Igor. "CIS-24 The Microcomputer Environment". Bristol Community College. Fall River, Massachussets, USA, 2010.  
<http://www.c-jump.com/CIS24/Slides/FAT/lecture.html>
- Kogan, Pablo. "Análisis de la eficiencia de la ponderación A para evaluar los efectos del ruido en el ser humano". Tesis de grado. Escuela de Ingeniería Acústica. Universidad Austral de Chile. Valdivia, 2004. <http://www.fceia.unr.edu.ar/acustica/biblio/kogan.pdf>
- Kozierok, Charles M. "New Technology File System". The PC Guide. 2004.  
<http://www.pcguides.com/ref/hdd/file/ntfs/index.htm>
- Lee, Justin. "MEMS Microphone". Course notes MECH 207. 2009.  
[http://mech207.engr.scu.edu/SensorPresentations/Lee%20-%20MEMS\\_Microphone%20Combined.pdf](http://mech207.engr.scu.edu/SensorPresentations/Lee%20-%20MEMS_Microphone%20Combined.pdf)
- Lewis, Jerad. "Microphone Specifications Explained". Analog Devices. Application Note AN-1112. 2012. [http://www.analog.com/static/imported-files/application\\_notes/AN-1112.PDF](http://www.analog.com/static/imported-files/application_notes/AN-1112.PDF)
- Lewis, Jerad "Microphone Array Beamforming". Analog Devices, Application Note AN-1140. 2012a. [http://www.analog.com/static/imported-files/application\\_notes/AN-1140.pdf](http://www.analog.com/static/imported-files/application_notes/AN-1140.pdf)
- Longares, Javier. "Sistemas de archivos FAT16 / FAT32 para las tarjetas SD / MMC". Octubre de 2011. Consultado: 20/07/13. Disponible en: <http://www.javierlongares.com/arte-en-8-bits/sistemas-de-archivos-fat16-fat32-para-las-tarjetas-sd-mmc/>
- McDougall, Richard. "Getting to know the Solaris filesystem, Part 1". Sun World, May 1999. <http://sunsite.uakom.sk/sunworldonline/swol-05-1999/swol-05-filesystem.html>
- Marengo Rodriguez, Fernando; Roveri, Eriberto; Rodríguez Guerrero, Juan M.; Treffiló, Mauro. "Análisis comparativo de codificadores de audio sin pérdidas y una herramienta gráfica para su selección y predicción de su desempeño". Mecánica Computacional Vol XXX, págs. 3167-3186 (artículo completo) Oscar Möller, Javier W. Signorelli, Mario A. Storti (Eds.) Rosario, Argentina, 1-4 Noviembre 2011
- Mazzoni, Dominic. "Audacity: A Digital Audio Editor. Noise removal.cpp". (sin fecha)  
<https://audacity.googlecode.com/svn/audacity-src/trunk/src/effects/NoiseRemoval.cpp>
- Microsoft Corporation. "Microsoft Extensible Firmware Initiative - FAT32 File System Specification - FAT: General Overview of On-Disk Format. Version 1.03, December 6, 2000. Disponible en <http://staff.washington.edu/dittrich/misc/fatgen103.pdf>
- Miloslav, Nic. "DTD Tutorial" Traducción: Iván García Alcaide.Zvon.org (sin fecha). Disponible en [http://www.zvon.org/xxl/DTDTutorial/General\\_spa/book.html](http://www.zvon.org/xxl/DTDTutorial/General_spa/book.html)
- Miyara, Federico. "Modelización del ruido del tránsito automotor". INGEACUS 99. Valdivia, Chile, 1999. <http://www.fceia.unr.edu.ar/biblio/MRT/MRT.htm>  
<http://www.fceia.unr.edu.ar/acustica/biblio/mrt.pdf>
- Miyara, Federico. "¿Ruido o señal? La otra información. En defensa del registro digital del ruido urbano". Cuarta Jornada Regional sobre Ruido Urbano, Montevideo, Uruguay, 2001. En CD-ROM. <http://www.fceia.unr.edu.ar/acustica/biblio/reg-dig.pdf>
- Miyara, F. Grabación digital: ¿DAT o MiniDisc? Tecnopolitan. Año 2 N° 15., 2001.  
<http://www.fceia.unr.edu.ar/acustica/biblio/md-vs-dat.pdf>
- Miyara, Federico. "Filtros Activos". Monografía de la cátedra de Electrónica III, Universidad Nacional de Rosario. Publicación interna B09.00. 2004.  
<http://www.fceia.unr.edu.ar/enica3/filtros-t.pdf>
- Miyara, Federico. "Introducción a la Electracústica". Monografía Cátedra de Fundamentos de Audio UNR. 2007

- Miyara, Federico; Miechi, Pablo; Pasch, Vivian; Cabanellas, Susana; Yanitelli, Marta; Accolti, Ernesto. "Tiempos de estabilización del espectro del ruido de tránsito". VI Congreso Iberoamericano de Acústica FIA2008. Trabajo A033. Rosario, 2008  
<http://www.fceia.unr.edu.ar/acustica/biblio/A033-Miyara.pdf>
- Miyara, F.; Pasch, V.; Yanitelli, M.; Accolti, E.; Cabanellas, S.; Miechi, P. Contratación de algoritmos de análisis de espectro con un instrumento normalizado. Actas de las Primeras Jornadas Regionales de Acústica AdAA 2009, Rosario, Argentina, 2009.  
[http://www.fceia.unr.edu.ar/acustica/biblio/A032%20\(Miyara\)%20Contraste%20algoritmos%20analizador%20normalizado.pdf](http://www.fceia.unr.edu.ar/acustica/biblio/A032%20(Miyara)%20Contraste%20algoritmos%20analizador%20normalizado.pdf)
- Miyara, Federico; Accolti, Ernesto; Pasch, Vivian; Cabanellas, Susana; Yanitelli, Marta; Miechi, Pablo; Marengo, Fernando; Mignini, Ezequiel. (2010a) "Suitability of a consumer digital recorder for use in acoustical measurements" Internoise 2010.  
[http://www.fceia.unr.edu.ar/acustica/biblio/consumer\\_recorders\\_for\\_noise\\_measurement\\_INTERNOISE\\_2010\\_993.pdf](http://www.fceia.unr.edu.ar/acustica/biblio/consumer_recorders_for_noise_measurement_INTERNOISE_2010_993.pdf)
- Miyara, F.; Accolti, E.; Pasch, V.; Cabanellas, S.; Yanitelli, M.; Miechi, P.; Marengo, F.; Mignini, E. (2010b) "Idoneidad de un grabador digital comercial para su empleo en mediciones acústicas". 2º Congreso Internacional de Acústica UNTREF. 2010.  
[http://www.fceia.unr.edu.ar/acustica/biblio/mediciones\\_acusticas\\_con\\_grabador\\_digital\\_comercial\\_UNTREF\\_2010.pdf](http://www.fceia.unr.edu.ar/acustica/biblio/mediciones_acusticas_con_grabador_digital_comercial_UNTREF_2010.pdf)
- Miyara, Federico. "Vibración de membranas". Monografía Cátedra de Fundamentos de Audio UNR. 2013
- Mohr, Peter J; Taylor, Barry N; Newell, David B. "CODATA Recommended Values of the Fundamental Physical Constants: 2006". National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8420, USA, 2007.
- National Park Service. "Listen to Edison Sound Recordings" Sin fecha. Consulta-do 29 jul 2013. <http://www.nps.gov/edis/photosmultimedia/the-recording-archives.htm>
- National Semiconductor. "An Introduction to the Sampling Theorem". Application Note AN-236. Linear Application Handbook. Santa Clara, 1986.
- OSTA. "Universal Disk Format® Specification". Revision 2.01. Marzo de 2000. Consultado el 23 jul 2013. <http://www.osta.org/specs/pdf/udf201.pdf>
- Park, Sangil. "Principles of Sigma-delta modulation for Analog to Digital Converters". Motorola APR8/D Rev. 1.  
[http://xanthippi.ceid.upatras.gr/people/psarakis/courses/DSP\\_APL/demos/APR8-sigma-delta.pdf](http://xanthippi.ceid.upatras.gr/people/psarakis/courses/DSP_APL/demos/APR8-sigma-delta.pdf)
- Peres, Michael R. (compilador) "Focal Encyclopedia of Photography. Digital Imaging, Theory and Applications, History, and Science, Fourth Edition". Focal Press - Elsevier. Burlington, Massachussets, USA, 2007.
- Pohlmann, Ken. "Principios de audio digital". McGraw-Hill Profesional. Madrid, 2002
- Pohlmann, Ken. "Advanced Digital Audio". SAMS - Prentice Hall. Carmel, Indiana, USA, 1993
- Proakis, John G.; Manolakis, Dimitris G. (1998). "Tratamiento digital de señales. Principios, algoritmos y aplicaciones". Prentice Hall, Madrid, España.
- Randall R. B., Upton R. "Digital Filters and FFT Technique in Real-time Analysis". Brüel & Kjaer. Technical Review 1978 N° 1
- Randall, R. B. "Application of B&K Equipment to Frequency Analysis". Brüel & Kjaer. Naerum, Denmark, 1977.  
<http://www.ingelec.uns.edu.ar/pds2803/Materiales/LibrosPDF/Randall/TOC.htm>
- Raynaudo, Ezequiel; Treffiló, Mauro. "Registro digital de ruido ambiental con transferencia inalámbrica de datos y post-procesamiento basado en software y hardware libre". Proyecto final de la Carrera de Ingeniería Electrónica, FCEIA-UNR. Rosario, 2013.
- Robinson, Tony. "Simple Lossless and Near-Lossless Waveform Compression". Technical Report CUED/F-INFENG/TR.156, Cambridge University, 1994. Disponible en Internet en <http://citeseer.nj.nec.com/robinson94shorten.html>.
- Roveri, Eriberto. "Estudio y comparación de codificadores de audio sin pérdidas". Proyecto final Ingeniería Electrónica. Univ. Nac. Rosario, FCEIA. Noviembre de 2011.



- Salomon, David. "Data Compression. The Complete Reference". Springer-Verlag. Londres, 2007
- Sennheiser Electronic Corporation. "HDA 200. Audiometric Headphone" (sin fecha)  
<http://en-us.sennheiser.com/downloads/a135337001d00197026b38d9c278f44d.pdf>
- Silva-Martinez, Jose. "Sigma-Delta Modulators Design Issues". Texas A&M University Electrical And Computer Engineering.  
<http://amesp02.tamu.edu/~jsilva/610/Lecture%20notes/Sigma-Delta-1.3.pdf>
- Smirnov, N.V.; Dunin-Barkowski, I. "Cálculo de probabilidades y estadística matemática". Ed. Paraninfo. Madrid, 1978.
- Stanford Research Systems. "Synthesized Function Generators - DS345 Specifications". Sin fecha. <http://www.thinksrs.com/downloads/PDFs/Manuals/DS345m.pdf>
- Tan, Kay. "25 Free Digital Audio Editors You Should Know". Sin fecha, Consultado 24 jul 2013. <http://www.hongkiat.com/blog/25-free-digital-audio-editors/>
- Thibodeau, Kennith. "Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years, Conference Proceedings: The State of Digital Preservation: An International Perspective (2002).
- Tuxera. "NTFS-3G + Ntfsprogs". 2003. <http://www.tuxera.com/community/ntfs-3g-download/>
- UKAS. "M3003 - The Expression of Uncertainty and Confidence in Measurement, Edition 2". United Kingdom Accreditation Service. 2007  
<http://www.ukas.com/library/Technical-Information/Pubs-Technical-Articles/Pubs-List/M3003.pdf>
- Watkinson, John. "El Arte del Audio Digital". Instituto Oficial de la RadioTelevisión Española (IORTV). 1993.
- Watkinson, John. "Audio digital". Editorial Paraninfo. Madrid, 1996.
- Wikipedia (2013). "Extent (file systems)". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 27 feb 2013. Consultado 21 jul 2013.  
[http://en.wikipedia.org/wiki/Extent\\_\(file\\_systems\)](http://en.wikipedia.org/wiki/Extent_(file_systems))
- Wikipedia (2013a). "B+ trees". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 30 may 2013. Consultado 22 jul 2013. [http://en.wikipedia.org/wiki/B%2B\\_tree](http://en.wikipedia.org/wiki/B%2B_tree)
- Wikipedia (2013b). "B trees". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 15 jul 2013. Consultado 22 jul 2013. <http://en.wikipedia.org/wiki/B-tree>
- Wikipedia (2013c). "Blu-ray Disc". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 21 jul 2013. Consultado el 22 jul 2013. [https://en.wikipedia.org/wiki/Blu-ray\\_Disc](https://en.wikipedia.org/wiki/Blu-ray_Disc)
- Wikipedia (2013d). "Ext4". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 26 jun 2013. Consultado el 22 jul 2013. <http://en.wikipedia.org/wiki/Ext4>
- Wikipedia (2013e). "HFS+". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 06 may 2013. Consultado el 23 jul 2013. <http://es.wikipedia.org/wiki/HFS%2B>
- Wikipedia (2013f). "Universal disc Format". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 12 jun 2013. Consultado el 23 jul 2013.  
[http://en.wikipedia.org/wiki/Universal\\_Disk\\_Format](http://en.wikipedia.org/wiki/Universal_Disk_Format)
- Wikipedia (2013g). "Cooley–Tukey FFT algorithm". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc., 11 jul 2013. Consultado el 28 jul 2013.  
[http://en.wikipedia.org/wiki/Universal\\_Disk\\_Format](http://en.wikipedia.org/wiki/Universal_Disk_Format)
- Williams, Ross N. "A painless guide to CRC error detection algorithms". 1993. Internet: [http://www.ross.net/crc/download/crc\\_v3.txt](http://www.ross.net/crc/download/crc_v3.txt)
- Yanitelli, Marta; Pasch, Vivian; Mosconi, Patricia; Cabanellas, Susana; Vazquez, Jorge; Rall, Juan Carlos; Miyara, Federico. "Manchas acústicas: Ruido de aeropuertos". Cuartas Jornadas Internacionales Multidisciplinarias sobre Violencia Acústica. Rosario, Argentina, 22 al 24 de octubre de 2001. <http://www.fceia.unr.edu.ar/biblio/aeropuer.pdf>
- Zoom Corporation. "Handy Recorder H4 Operation Manual". (fecha no especificada)  
[http://www.zoom.co.jp/archive/English\\_Manual/E\\_H4.pdf](http://www.zoom.co.jp/archive/English_Manual/E_H4.pdf)
- Zwicker, Eberhard; Fastl, Hugo. "Psychoacoustics - Facts and Models". Springer-Verlag. Berlin, 1999.

## Apéndice 1

### Glosario y definiciones sobre metrología<sup>115</sup>

**Cantidad:** Una propiedad de un fenómeno o de un sistema susceptible de ser representada numéricamente relativa a una referencia.

NOTA: También se suele utilizar el término *magnitud*. Ver, por ejemplo, la norma venezolana COVENIN 2552

**Mensurando:** Cantidad objeto de la medición.

NOTA: La definición de un mensurando particular requiere establecer una serie de condiciones que permitan especificarlo de manera única. Muchas veces la definición de esas condiciones es incompleta ya que es imposible o impráctico especificar todos los factores que podrían afectar el valor del mensurando. En general se procura al menos especificar aquellos factores que presumiblemente tienen un impacto significativo en el valor del mensurando.

**Realización del mensurando:** Una instancia física del mensurando susceptible de ser sometida al proceso de medición.

**Medición:** Proceso mediante el cual se obtiene una aproximación o estimación del valor de un mensurando, así como alguna expresión de la incertidumbre.

**Principio de medición:** Fenómeno físico en el que se basa la medición

NOTA: Por ejemplo, transducción capacitiva de presión en tensión eléctrica

**Método de medición:** Descripción general de los pasos a seguir en la medición

**Procedimiento de medición:** Descripción detallada del proceso de medición. Incluye todos los aspectos a tener en cuenta, desde el tipo y características del instrumental hasta los cálculos y fórmulas a aplicar, pasando por la adecuada selección de las muestras y el relevamiento de todas aquellas variables que puedan influir en el resultado de la medición.

**Modelo de la medición:** Relación matemática entre las variables involucradas en una medición. Puede ser una fórmula explícita, una expresión implícita o un algoritmo.

**Valor verdadero:** Valor de un mensurando que se obtendría, en condiciones especificadas, mediante un instrumento de medición ideal. El valor verdadero en general no se conoce exactamente. Sólo se lo estima, aproxima o acota.

NOTA 1: Algunas excepciones son las magnitudes involucradas en la definición de las unidades base del Sistema Internacional de Unidades (SI). Por ejemplo, la velocidad de la luz en el vacío es exactamente 299 792 458 m/s.

NOTA 2: La *Guía para la expresión de la incertidumbre* de la ISO (ISO-GUM) evita el uso de la locución “valor verdadero” y en su lugar utiliza “valor del mensurando”, considerando que el adjetivo “verdadero” es redundante.

---

<sup>115</sup> En general estas definiciones concuerdan conceptualmente con las ideas expresadas en la “Guía para la expresión de la incertidumbre en las mediciones” (Guide to the Expression of Uncertainty of Measurements, GUM) de la ISO (ISO-GUM). Cuando no hay concordancia o se presentan otras alternativas, se aclara. En muchos casos se incluyen notas con ejemplos y comentarios

NOTA 3: La GUM también aclara que en general no hay un único valor del mensurando, ya que al haber ambigüedad en la definición del mensurando por la imposibilidad de especificar todas las condiciones que podrían influir, podría haber más de un valor que satisface la definición. Esto da origen a una incertidumbre intrínseca que no es atribuible al instrumento ni al procedimiento de medición. Un ejemplo es la medición del nivel equivalente de un ruido blanco extendido a un tiempo  $T$ . En la ecuación (A2.6) del apéndice se muestra la dispersión estadística de la presión eficaz cuadrática, que se traslada al nivel equivalente causando una incertidumbre intrínseca. Dicha componente de la incertidumbre desaparece si dentro de la definición del mensurando incluimos el instante exacto de inicio y fin de la medición, pero subsiste una incertidumbre debido a que no es posible verificar con ningún reloj real dichos instantes en forma exacta.

NOTA 4: En algunas situaciones se interpreta el valor verdadero como un valor de referencia convencional. Un ejemplo es en el contexto de la calibración, donde se dispone de un patrón que se supone más exacto que el instrumento a calibrar. Este valor de referencia se ha determinado con una incertidumbre mucho menor que la incertidumbre esperada debido al procedimiento de medición.

**Precisión:** Grado de coincidencia de una serie de mediciones con su media. Se expresa en general como el desvío estándar de los valores medidos en condiciones de repetibilidad, es decir, igualdad de procedimiento, lugar, operador, instrumento y condiciones ambientales.

NOTA: Por ejemplo, un sonómetro tiene una precisión de 0,1 dB si varias mediciones de un mismo sonido en idénticas condiciones tienen un desvío estándar de 0,1 dB.

**Exactitud:** Grado de coincidencia de una medición con el valor verdadero.

NOTA 1: Por ejemplo, un sonómetro tiene una exactitud de 0,1 dB si cualquier medición difiere del valor verdadero en menos de 0,1 dB. La exactitud depende del estado de calibración del instrumento. Sin embargo, la ISO (International Organization for Standardization) y otros organismos asociados, como el NIST (National Institute of Standards and Technology de EEUU) atribuyen a la exactitud un carácter cualitativo, no susceptible de ser asociada a un valor numérico.

NOTA 2: Una razón para esto es que la teoría de la incertidumbre aceptada por la ISO, el NIST y otros organismos implica siempre alguna probabilidad (aunque extremadamente baja) de que el valor medido difiera del valor verdadero en una cantidad arbitrariamente grande.

**Resolución:** Para un instrumento dado, mínima diferencia posible entre dos lecturas diferentes.

NOTA 1: La resolución no implica exactitud, pero sí impone una cota a la exactitud posible. Así, un sonómetro con una resolución de 0,1 dB no puede garantizar una exactitud mejor que 0,05 dB. En efecto, si el valor verdadero de un nivel de presión sonora es de 62,35 dB, el valor observado será, en el mejor de los casos, 62,3 dB o 62,4 dB, que difieren del valor verdadero en 0,05 dB.

NOTA 2: En general la resolución de cualquier instrumento es mucho menor que la componente de incertidumbre debida al instrumento, de modo que el problema señalado en la nota anterior no es importante.

NOTA 3: En el caso del sonómetro, la resolución puede mejorarse si en lugar de tomar la lectura directamente se calcula por software a partir de una grabación digital, ya que para la misma se discretiza la señal con muchos bits (en general 16 o incluso 24). Ello no implica, sin embargo que mejore demasiado la incertidumbre general de la medición.

**Tolerancia:** Máxima dispersión admisible entre el valor real y el valor especificado de una cantidad relativa a un producto.

NOTA 1: La dispersión se expresa estadísticamente como una cantidad de veces el desvío estándar para garantizar una confiabilidad dada (por ejemplo, 95 %).

NOTA 2: Si la tolerancia se definiera como *la máxima diferencia entre el valor real y el valor especificado*, no habría forma de verificar, en un caso concreto, si el producto satisface la tolerancia, debido a la incertidumbre de la medición del valor real.

**Repetición:** Reiteración de la medición de un mismo mensurando, por ejemplo, la longitud de una barra en idénticas condiciones.

NOTA 1: Las condiciones incluyen las variables atmosféricas, el instrumento, el operador y su ubicación en relación con la fuente y los objetos circundantes.

NOTA 2: La repetición de una medición y la consecuente promediación de los resultados permiten reducir la incertidumbre de la medición.

NOTA 3: Conceptualmente, en el caso de mediciones acústicas no es posible en general la repetición de mediciones dado que el mensurando es variable en el tiempo. Una excepción son las condiciones de laboratorio, donde la fuente se puede controlar para producir esencialmente la misma señal.

**Repetibilidad:** Grado de coincidencia entre las diferentes mediciones del mismo mensurando realizadas en las mismas condiciones (procedimiento, lugar, operador, instrumento, condiciones ambientales) expresado en términos de la dispersión estadística de los resultados.

**Condiciones de repetibilidad:** Condiciones en las que se evalúa la repetibilidad de un proceso de medición (igualdad de procedimiento, lugar, operador, instrumento y condiciones ambientales).

**Réplica:** Reiteración de la ejecución de una prueba.

NOTA 1: La réplica difiere de la repetición en que la ésta se hace sobre el mismo mensurando, en cambio la réplica se hace sobre versiones diferentes del mensurando. Por ejemplo, la medición de la longitud de varias barras obtenidas de un mismo proceso de producción.

NOTA 2: La replicación permite determinar o estimar la dispersión debida al proceso de producción, que debe distinguirse de la incertidumbre en la medición.

**Reproducibilidad:** Grado de coincidencia entre diferentes mediciones del mismo mensurando realizadas variando algunas condiciones especificadas (por ejemplo, el laboratorio, el instrumento, el operador), expresado en términos de la dispersión estadística de los resultados.

NOTA: El propósito de la reproducibilidad es tener una medida de la robustez de un proceso de medida frente a cambios de algunas condiciones que se espera puedan cambiar al intentar replicar un experimento. Por ejemplo, en estudios conjuntos interlaboratorios, en la verificación y confirmación de resultados de investigación científica o técnica, en pericias legales, en la preparación de hojas de especificaciones y en la verificación independiente de éstas. En todos estos casos pueden variar, como mínimo, el instrumento (incluyendo marca, modelo y número de serie, aun cuando cumplan con la misma norma técnica), el operador, el lugar, el material de investigación e incluso condiciones ambientales (como temperatura, humedad). También podrían variar el protocolo o secuencia de medición. Cuantas más condiciones puedan variar, la reproducibilidad empeorará, es decir, aumentará la dispersión de los valores medidos.

**Condiciones de reproducibilidad:** Condiciones en las que se evalúa la reproducibilidad de un proceso de medición, especificando cuáles condiciones admiten cambios (por ejemplo, el operador, el instrumento y el laboratorio) y cuáles no.

**Error de medición:** Diferencia entre el valor medido y el valor verdadero.

$$\varepsilon = x_{\text{medido}} - x_{\text{verdadero}} .$$

NOTA 1: El error no es conocido, pero en algunos casos se lo puede acotar o especificar estadísticamente por la incertidumbre. Sin embargo, la teoría de la incertidumbre en general adoptada no permite una cota absoluta para el error sino estadística. Ver el apartado Cota del error.

NOTA 2: El error de medición es la suma del error sistemático y el error aleatorio.

NOTA 3: A veces el error de medición se define con respecto a un valor de referencia convencional y no con respecto al valor verdadero.

**Error de observación:** Error de medición.

**Error sistemático:** Media estadística de los errores de medición de un mismo mensurando en las mismas condiciones.

$$\varepsilon_{\text{sistemático}} = E(x_{\text{medido}} - x_{\text{verdadero}}) .$$

NOTA: El error sistemático puede deberse a una descalibración del instrumento, o a vicios o particularidades del procedimiento de medición. Por ejemplo si una medición de un ruido de baja frecuencia se realiza muy cerca de una superficie reflectora, la medición arrojará sistemáticamente valores más altos que lo que se mediría lejos de tales superficies.

**Error aleatorio:** Diferencia entre el valor de una medición y la media estadística (esperanza matemática) de las mediciones

$$\varepsilon_a = x_{\text{medido}} - E(x_k) = x_{\text{medido}} - (x_{\text{verdadero}} + \varepsilon_{\text{sistemático}}) .$$

NOTA: El error aleatorio puede deberse a limitaciones del instrumento de medición, a condiciones no controladas dentro del proceso de medición (por ejemplo condiciones ambientales, lugar de medición, operador), a la variabilidad inherente del mensurando o a su carácter aleatorio. De importancia en metrología acústica es la variabilidad del mensurando con el tiempo.

**Cota del error:** Máximo valor absoluto del error de medición bajo condiciones estipuladas.

NOTA: La Guía de Expresión de la Incertidumbre (Guide to the Expression of Uncertainty of Measurements, GUM) de la ISO, así como el NIST, tácitamente presumen que no es posible dar una cota del error, ya que sólo admiten para su cuantificación parámetros estadísticos como la incertidumbre expandida, que estipula los márgenes dentro de los cuales se encontrará el valor verdadero con respecto al valor medido con un porcentaje dado de probabilidad (por ejemplo 95 %). Pero no ofrecen ningún parámetro que determine un intervalo dentro del cual se encuentre el valor verdadero con certeza absoluta (100 %).

**Densidad de probabilidad:** Función tal que la probabilidad de que una variable aleatoria  $X$  esté entre  $x_1$  y  $x_2$  es

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x) dx .$$

**Media (valor esperado):** Valor medio de todos los valores posibles de una variable aleatoria  $X$ :

$$\mu = \int_{-\infty}^{\infty} x f(x) dx .$$

donde  $f(x)$  es la densidad de probabilidad.

NOTA 1: La media también se llama *esperanza matemática* y, a veces, *media poblacional*.

**Desvío estándar:** Raíz cuadrada del valor medio de los cuadrados de las diferencias entre los valores de la variable aleatoria y la media:

$$\sigma = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx} .$$

donde  $f(x)$  es la densidad de probabilidad.

**Varianza:** Es el cuadrado del desvío estándar de una variable aleatoria.

**Densidad de probabilidad conjunta:** Dadas dos variables aleatorias  $X, Y$ , es una función tal que

$$P(x_1 < x < x_2, y_1 < y < y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y) dx dy .$$

Alternativamente, en una versión diferencial de la definición, es una función tal que

$$P(x < X < x + dx, y < Y < y + dy) = f(x, y) dx dy .$$

**Covarianza:** Dadas dos variables aleatorias  $X$  e  $Y$ , es el valor medio del producto de las diferencias entre cada variable y su media:

$$\text{cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f(x, y) dx dy$$

donde  $f(x, y)$  es la densidad de probabilidad conjunta entre  $X$  e  $Y$ .

NOTA 1: La covarianza de una variable aleatoria consigo misma es igual a su varianza. En tal caso la densidad de probabilidad es

$$f(x, y) = \delta(y - x) f(x),$$

donde  $\delta$  es la función impulso de Dirac.

NOTA 2: Cuando dos variables aleatorias  $X, Y$  son independientes, es decir tales que la densidad de probabilidad conjunta  $f_{XY}$  es el producto de las densidades de probabilidad  $f_X$  y  $f_Y$ :

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

entonces

$$\text{cov}(X, Y) = 0.$$

La covarianza es tanto mayor cuanto más parecidas son las variables (es decir cuando los resultados de mediciones simultáneas son similares)

**Media muestral:** Dadas  $n$  muestras  $x_1, \dots, x_n$  de una variable aleatoria  $X$ , es el valor medio de dichas muestras:

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k .$$

NOTA 1: Este parámetro aproxima a la media o valor esperado. Es, pues un *estimador* de la media. Dado que  $\bar{x}$  es a su vez una variable aleatoria, se puede calcular su media  $\mu_{\bar{x}}$  y su desvío estándar  $\sigma_{\bar{x}}$  que valen

$$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} .$$

NOTA 2: Dado que el desvío estándar se reduce al aumentar  $n$ , es posible reducir la dispersión en la determinación de  $\mu$  repitiendo la medición muchas veces.

**Desvío estándar muestral:** Dadas  $n$  muestras  $x_1, \dots, x_n$  de una variable aleatoria  $X$ , es:

$$s = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2} .$$

NOTA: Este parámetro aproxima al desvío estándar de  $X$ . Es, por lo tanto, un *estimador* del desvío estándar.

**Covarianza muestral:** Dados  $n$  pares de muestras  $(x_1, y_1), \dots, (x_n, y_n)$  de las variables aleatorias  $X, Y$ , es

$$m_{xy} = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})(y_k - \bar{y})$$

NOTA 1: Se supone que los pares  $(x_k, y_k)$  son resultados de una misma observación o réplica en la que se han recogido *simultáneamente* ambos datos. Por ejemplo, la temperatura ambiente y el nivel de presión sonora.

NOTA 2: Este parámetro aproxima la covarianza de  $X$  e  $Y$ . Es, por lo tanto, un *estimador* de la covarianza.

NOTA 3: La covarianza muestral de una variable consigo misma es igual a la varianza muestral de la variable.

NOTA 4: La covarianza muestral de dos variables aleatorias independientes se aproxima a 0 pero no necesariamente es igual a 0.

**Incertidumbre:** Noción cualitativa o expresión cuantitativa del grado de desconocimiento que se tiene, luego de una medición, de la diferencia entre el valor medido (estimado) y el valor verdadero del mensurando (es decir, el error de medición). Generalmente está conformada por una serie de componentes atribuibles a diversos aspectos del proceso de medición (instrumento, operador, naturaleza del fenómeno observado). Puede expresarse como un parámetro estadístico.

**Incertidumbre estándar:** Desvío estándar estimado de una serie de mediciones de un mismo mensurando bajo condiciones especificadas:

$$u = s = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2}.$$

NOTA: La incertidumbre estándar estará integrada en general por varias componentes asociadas a diferentes causas (ver *Incertidumbre estándar combinada*).

**Incertidumbre tipo A:** Cualquier componente de la incertidumbre obtenida por métodos estadísticos a partir de los propios datos medidos, en general luego de varias repeticiones o réplicas de la medición en condiciones de repetibilidad especificadas. Se pueden utilizar métodos generalmente aceptados de tratamiento estadístico de datos.

**Incertidumbre tipo B:** Cualquier componente de la incertidumbre obtenida por otros métodos. Entre ellos, tenemos: el conocimiento o estimación previa por consenso científico o técnico (por ejemplo incertidumbres en constantes físicas o en procedimientos), información provista en normas, especificaciones del instrumental empleado, información provista en certificados de calibración, simulaciones por software.

**Incertidumbre estándar combinada:** Dadas  $N$  variables aleatorias  $X_1, \dots, X_N$  de las que depende funcionalmente un mensurando  $Y$  según

$$Y = f(X_1, \dots, X_N),$$

supongamos que deseamos obtener el valor de  $y$  para unos determinadas valores medidos  $x_1, \dots, x_N$ . La *incertidumbre estándar combinada* es, entonces

$$u_c = \sqrt{\sum_{h=1}^N \sum_{k=1}^N \frac{\partial f}{\partial x_h} \frac{\partial f}{\partial x_k} u(x_h, x_k)}$$

donde  $\partial f / \partial x_h$  y  $\partial f / \partial x_k$  son las derivadas parciales de  $f$  calculadas en los valores observados (estimados)  $x_1, \dots, x_N$  y  $u(x_h, x_k)$  son las covarianzas estimadas

$$u(x_h, x_k) = m_{x_h y_k} = \frac{1}{n-1} \sum_{i=1}^n (x_{hi} - \bar{x}_h) (x_{ki} - \bar{x}_k)$$

NOTA 1: La fórmula de  $u_c$  se obtiene aplicando una aproximación de Taylor multivariable de primer orden centrada en los valores observados de las variables  $x_1, \dots, x_N$ . Se la conoce también como *fórmula de propagación de errores*.

NOTA 2: Para  $h = k$  se tiene

$$u(x_h, x_h) = \sigma_{x_h}^2.$$

NOTA 3: Si todas las variables son independientes entre sí, la incertidumbre estándar combinada es

$$u_c = \sqrt{\sum_{h=1}^N \left( \frac{\partial f}{\partial x_h} \right)^2 \sigma_{x_h}^2}.$$

ya que las covarianzas entre variables diferentes se anulan (la *matriz de covarianza* es diagonal).



NOTA 4: La función  $f$  puede contener el modelo físico para determinar indirectamente una magnitud, por ejemplo el volumen  $V$  de un paralelepípedo a partir de sus aristas  $a$ ,  $b$ ,  $c$ :

$$V = a b c.$$

En este caso es una función de tres variables. También podría describir el modelo del proceso de medición. Por ejemplo, si se desea medir el nivel equivalente de un ruido estacionario a partir de 10 mediciones, tendremos

$$L_{eq} = 10 \log_{10} \left( \frac{10^{L_{p1}/10} + \dots + 10^{L_{p10}/10}}{10} \right).$$

En este caso es una función de 10 variables.

**Incertidumbre expandida:** Es la incertidumbre estándar combinada multiplicada por un factor de cobertura  $k$

$$U = k u_c$$

NOTA 1: Suponiendo una distribución normal, el factor de cobertura es tal que permite garantizar una determinada probabilidad  $P$  de que el valor del mensurando se encuentre en un entorno de  $\pm U$  del valor estimado  $y$ , es decir,

$$y - U \leq Y \leq y + U$$

Si la densidad de probabilidad es gaussiana (normal) y  $k = 2$ , entonces  $P = 0,96$  lo cual significa que hay una probabilidad de sólo 4 % de que el valor estimado del mensurando difiera más de  $U$  del valor verdadero.

NOTA 2: La incertidumbre expandida, aun con un factor de cobertura muy alto, no permite definir un intervalo en el que se pueda garantizar con probabilidad  $P = 1$  que se en él se encuentra el valor verdadero del mensurando. Por ejemplo, para una distribución normal un factor de cobertura  $k = 5$  sólo da una confiabilidad de 0,99999944. Probablemente la razón para adoptar este enfoque en la expresión de la incertidumbre sea que el formalismo estadístico detrás del concepto de incertidumbre permite un cálculo relativamente riguroso del efecto de las diferentes fuentes de incertidumbre, el cual resulta, además bastante simple en el caso en que se pueda aceptar que las variables tienen distribuciones normales. Pero elimina toda certeza en las mediciones. Por ejemplo, al medir una barra de 1,2 m de largo existiría una probabilidad no nula de obtener un valor de 2 m, 10 km o, incluso, de  $-1$  m. Esta situación ha dado origen a controversias. Ver *Cota del error*.

**Trazabilidad:** Cadena de incertidumbres conocidas que conectan el resultado de una medición (estimación del valor del mensurando) con un patrón nacional o internacional reconocido por la comunidad metrológica.

NOTA 1: En ausencia de trazabilidad no existe forma de cuantificar la incertidumbre de una medición. Por ejemplo, un instrumento de precisión podría estar descalibrado y entonces se desconoce no sólo su error sistemático sino su posible distribución estadística.

**Patrón de medición:** Puede ser una materialización o realización de una unidad de una magnitud (por ejemplo 1 m si la magnitud es la longitud), o un instrumento que permite medirla, con el fin de definirla, conservarla o reproducirla.

**Patrón primario (estándar primario):** Patrón de medición de una magnitud de la máxima calidad técnicamente posible. Típicamente cada país tiene un patrón primario con trazabilidad al prototipo internacional, cuando el mismo existe.

**Patrón de referencia (patrón de laboratorio):** Patrón de medición de una magnitud con trazabilidad a un patrón primario a través de un proceso de calibración. El uso típico de este patrón es para la calibración de patrones de trabajo

**Patrón de trabajo:** Patrón de medición de una magnitud con trazabilidad a un patrón de referencia a través de un proceso de calibración con el fin de realizar mediciones de campo o de laboratorio.

**Calibración:** Proceso por el cual se determina la incertidumbre de un instrumento en varias condiciones de operación con referencia a un patrón con trazabilidad a un patrón nacional o internacional reconocido por la comunidad metrológica.

NOTA: Si se trata de un sonómetro, las condiciones de operación pueden incluir su ensayo a varias frecuencias y en varios rangos.

**Series de Renard:** Series de valores preferidos introducidos por Charles Renard en 1870. Están distribuidos en forma aproximadamente uniforme en una década en escala logarítmica. Se designan por la letra R seguida por la cantidad de valores por década, que pueden ser 5, 10, 20 ó 40. Así, la R10 introduce los siguientes 10 valores en una década: 1,00 1,25 1,60 2,00 2,50 3,15 4,00 5,00 6,30 8,00. Las series se extienden a otras décadas como múltiplos y submúltiplos por potencias de 10. Los valores son aproximaciones de las series  $10^{k/n}$ , donde  $k = 0, \dots, n-1$  y  $n$  es la cantidad de valores por década.

NOTA 1: Dado que  $10^{3/10} \cong 2$ , las series R10, R20 y R40 tienen la propiedad de que los valores también permiten subdividir exacta o casi exactamente la octava. Por ello se han adoptado como series normalizadas para los filtros de uso en acústica de octava y tercio de octava.

NOTA 2: Las series de Renard completas son

R5:	1,00	1,60	2,50	4,00	6,30																																			
R10:	1,00	1,25	1,60	2,00	2,50	3,15	4,00	5,00	6,30	8,00																														
R20:	1,00	1,12	1,25	1,40	1,60	1,80	2,00	2,24	2,50	2,80	3,15	3,55	4,00	4,50	5,00	5,60	6,30	7,10	8,00	9,00																				
R40:	1,00	1,06	1,12	1,18	1,25	1,32	1,40	1,50	1,60	1,70	1,80	1,90	2,00	2,12	2,24	2,36	2,50	2,65	2,80	3,00	3,15	3,35	3,55	3,75	4,00	4,25	4,50	4,75	5,00	5,30	5,60	6,00	6,30	6,70	7,10	7,50	8,00	8,50	9,00	9,50

## Apéndice 2

### Conceptos fundamentales de Estadística

#### A2.1 Estadística descriptiva

Dado que el análisis estadístico es una parte importante de todo proceso de medición, es conveniente repasar algunos conceptos fundamentales de Estadística.

En primer lugar, tengamos en cuenta que uno de los objetivos de esta disciplina es obtener propiedades globales de un conjunto de entes (denominado *población*) cuya determinación individual no es posible en forma exacta para todos los entes. Un ejemplo es la altura de las personas de una determinada región. Tomada una persona al azar, no es posible predecir exactamente cuál será su altura. Pero sí puede determinarse un *descriptor global*, tal como la *altura promedio* de tal conjunto de personas, e inclusive puede especificarse la *dispersión* de la altura, es decir, qué tan diferente puede llegar a ser la altura real de la altura promedio. Lo más interesante de los métodos estadísticos es la posibilidad de *estimar* dichas propiedades globales a través del examen de una adecuada *muestra*<sup>116</sup> de la población, es decir un subconjunto reducido de la misma. Esta tarea se denomina *inferencia estadística*.

Las propiedades asociadas a los entes que forman la población se expresan en general como *variables aleatorias*, es decir variables que asumen un valor determinado (pero en general desconocido) para cada elemento de la población. En el ejemplo anterior, la altura de los individuos de una dada región geográfica constituye una variable aleatoria. En el caso de la metrología acústica, la variable aleatoria de interés será el *resultado de una medición*.

El comportamiento de las variables aleatorias está determinado por su *distribución estadística*. Para esclarecer este concepto, supongamos que estamos interesados en el nivel umbral de audición a 4000 Hz de un conjunto de 50 personas. En primer lugar subdividimos el rango total de dicha variable en subintervalos de 5 dB en 5 dB. Luego simplemente contamos la cantidad de individuos cuyos umbrales se encuentran en cada uno de estos subintervalos. Obtenemos una tabla como la A2.1, en la cual también se expresaron las cantidades como porcentajes:

Las cantidades de personas en cada “clase” se denominan *frecuencias*,<sup>117</sup> obteniendo de esta manera la denominada *distribución de frecuencias*. Los porcentajes, por otra parte, dan origen a la *distribución porcentual*. Ambas pueden representarse gráficamente como se indica en la figura A2.1.

La distribución porcentual es una manera más conveniente de representar la distribución estadística de una variable aleatoria, ya que en primer lugar permite comparar los comportamientos de diferentes conjuntos de datos, y en segundo lugar da origen a la *inferencia estadística*, ya que si de acuerdo con algún criterio podemos considerar que la muestra (en este caso los niveles umbrales de las 50 personas) es representativa de la población podemos extrapolar los resultados obtenidos en la muestra a toda la población y por ende a otras muestras extraídas de la misma.

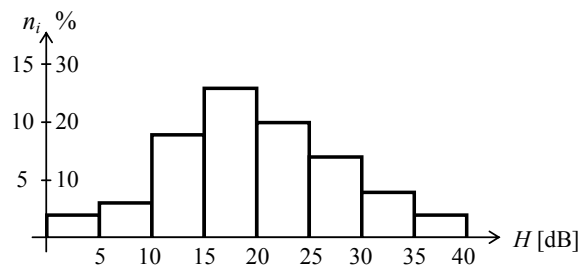
<sup>116</sup> Nótar que en Estadística *muestra* tiene un significado diferente que en procesamiento de señales discretas. En Estadística corresponde a una cierta cantidad  $n$  de instancias de una variable aleatoria, mientras que en procesamiento de señales es un único valor de la señal discreta (tomado de la señal analógica que la señal discreta representa).

<sup>117</sup> No confundir la *frecuencia de aparición* de una clase de valores con el concepto físico de *frecuencia* de un tono o de cualquier otro fenómeno periódico

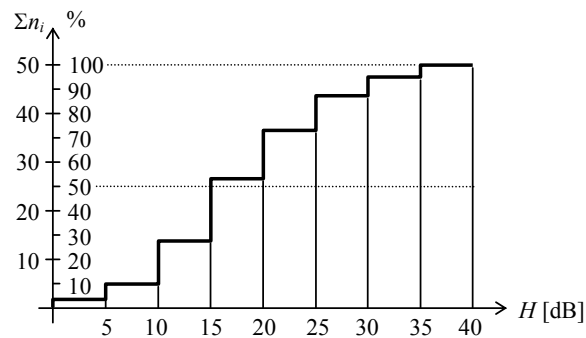
**Tabla A2.1.** Ejemplo de distribución de frecuencias: umbral auditivo de un conjunto de 50 personas.

Intervalo [dB]	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Cantidad	2	3	9	13	10	7	4	2
Porcentaje [%]	4	6	18	26	20	14	8	4

También es posible hablar de la *distribución acumulada de frecuencias*. En este caso se van sumando las cantidades correspondientes a niveles umbrales crecientes, lo cual se ilustra en la figura A2.2. Si se representan como porcentajes se obtiene la *distribución porcentual acumulada*.



**Figura A2.1.** Representación gráfica de las distribuciones de frecuencias y porcentual correspondientes al ejemplo de la tabla 5.



**Figura A2.2.** Representación gráfica de la distribución acumulada de frecuencias y la distribución porcentual acumulada correspondiente al ejemplo de la tabla A2.1.

A cada variable aleatoria es posible asociarle además dos tipos de medidas numéricas: las de *tendencia central* y las de *dispersión*. Las medidas de tendencia central determinan, según diversos criterios, un valor “central” alrededor del cual se agrupan, aglutinan o distribuyen los valores de la variable correspondientes a todos los integrantes de la

población. Las medidas de dispersión indican, también según una variedad de criterios, cuán cerca o lejos se encuentran, en general, los valores de la variable del valor “central”.

Las tres medidas de tendencia central más utilizadas son el *valor medio* (o *promedio*, o *media*), la *mediana* y la *moda*. El *valor medio*  $\bar{x}$  de una variable aleatoria  $x$  se obtiene como

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k, \quad (\text{A2.1})$$

donde  $N$  es la cantidad de integrantes de la población y  $x_k$  el valor de la variable para el  $k$ -ésimo integrante. En el caso en que el rango total de la variable aleatoria se haya subdividido en una cantidad (relativamente pequeña)  $m$  de clases (como en el ejemplo de la tabla 5), la ecuación (A2.1) puede reescribirse

$$\bar{x} = \frac{1}{N} \sum_{i=1}^m n_i x_i, \quad (\text{A2.2})$$

donde  $x_i$  es el valor de la variable aleatoria asociado a la clase  $i$ -ésima y  $n_i$  la frecuencia correspondiente. Aplicando esta ecuación a las frecuencias de la tabla 5, y suponiendo que la variable en cada intervalo toma el valor medio de sus extremos (por ejemplo en el intervalo 20-25, supondremos  $x_i = 22,5$ ), resulta  $\bar{x} = 19,8$  dB.

La *mediana*,  $\text{med}(x)$ , es un valor tal que un 50% de los integrantes tienen valores de la variable menores que  $\text{med}(x)$  y el restante 50% tienen valores mayores que  $\text{med}(x)$ . Es fácil determinar la mediana a partir de la distribución porcentual acumulada, viendo cuál es el valor más cercano al 50%. En el ejemplo anterior, de la figura A2.2 se desprende que dicho valor es el valor intermedio entre 15 y 20, es decir 17,5 dB.

Por último, la *moda* de una distribución,  $\text{moda}(x)$ , es el valor en el cual la distribución de frecuencias alcanza un máximo. En el ejemplo anterior, de la figura 5 se observa que dicho valor es también el valor medio entre 15 y 20, es decir 17,5 dB. Es de notar que una distribución puede tener más de una moda, lo cual se da cuando existe una gran polarización en torno a dos o más valores. Un ejemplo es el nivel umbral de audición entre los trabajadores de una empresa industrial: los empleados administrativos, que trabajan en ambientes relativamente silenciosos tendrán umbrales variados pero bajos; los operarios que trabajan en producción con maquinarias y procesos ruidosos exhibirán también diversos umbrales, pero grupalmente más elevados. Este tipo de distribución se denomina *bimodal*.

La utilización de una u otra medida de tendencia central depende de varios factores. El valor medio es analíticamente más simple, y permite el aprovechamiento de los resultados de la Teoría de la Probabilidad, por su correspondencia con la *esperanza matemática*. La principal desventaja es que es muy sensible a los valores extremos, sobre todo en variables que pueden asumir valores muy disímiles. Un caso típico lo constituye el nivel equivalente  $L_{eq}$ . Por ejemplo, si el ruido de fondo es de alrededor de 55 dBA, el paso de un camión emitiendo 90 dBA durante 10 s llevará el nivel equivalente a 82,3 dBA, mucho más cercano a 90 dBA que a 55 dBA, a pesar de la escasa duración del ruido más intenso. Cuando la distribución es simétrica alrededor de un valor, el valor medio y la mediana coinciden. En las distribuciones fuertemente asimétricas la mediana suele dar mejores resultados. La moda es un buen descriptor a la hora de comuni-

car resultados, más que al procesarlos. Una desventaja es que ignora lo que sucede con todos los valores de la variable excepto el más frecuente.<sup>118</sup>

La medida de dispersión más utilizada es el *desvío estándar*,  $\sigma_x$ , definido como la raíz cuadrada del promedio de los desvíos con respecto a la media elevados al cuadrado:

$$\sigma_x = \sqrt{\frac{\sum_{k=1}^N (x_k - \bar{x})^2}{N}}, \quad (\text{A2.3})$$

donde  $x_k$  son los valores que asume la variable aleatoria en la población,  $N$  la cantidad total de integrantes de la población y  $\bar{x}$  el valor medio de  $x$ . En caso de que la variable haya sido agrupada en subintervalos, también puede calcularse como

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^m n_i (x_i - \bar{x})^2}{N}}. \quad (\text{A2.4})$$

donde  $x_i$  es el valor correspondiente al  $i$ -ésimo subintervalo y  $n_i$  la frecuencia de aparición de valores en dicho subintervalo, siendo  $N = \sum n_i$ . A modo de ejemplo, podemos calcular el desvío estándar correspondiente al nivel umbral de audición del ejemplo de la tabla A2.1. Resulta  $\sigma_x = 8,2$  dB. En general una parte importante de los valores de una distribución se encuentran comprendidos entre  $\bar{x} - 2\sigma_x$  y  $\bar{x} + 2\sigma_x$ . Cuanto mayor sea el desvío estándar, más difusa y por lo tanto menos concentrada será la distribución de la variable aleatoria.

También es de interés en Estadística la *varianza*, definida como el cuadrado del desvío estándar, ya que tiene propiedades aditivas frente a causas independientes de dispersión de una variable aleatoria. La única desventaja es que dimensionalmente no es compatible con la variable, dado que las unidades están elevadas al cuadrado.

Tanto el valor medio de una variable aleatoria como su desvío estándar relativos a una población pueden ser *estimados* a partir de una *muestra* seleccionada apropiadamente. En general, para que una muestra sea adecuada debe seleccionarse aleatoriamente<sup>119</sup> y de un tamaño suficiente para garantizar una alta probabilidad de que el error en la estimación no supere cierto valor (por ejemplo, que la probabilidad de tener un error menor que 2 dB sea de un 95 %). La discusión detallada de los criterios de selección de la muestra y de las pruebas preliminares requeridas excede el propósito introductorio de este apéndice.

El *valor medio poblacional* puede ser estimado directamente por el *valor medio muestral*, que se calcula en forma idéntica a aquél, es decir

<sup>118</sup> La moda es útil cuando se requiere analizar variables aleatorias nominales, es decir no numéricas; por ejemplo, cuando se pretende averiguar mediante una encuesta qué tipo de ruido resulta más molesto entre el de las motos con escape libre, el de una discoteca, y el de los martillos neumáticos. No es posible en casos como éste promediar los tres tipos de ruido ni ordenarlos de “menor” a “mayor”.

<sup>119</sup> Por ejemplo, si se desea obtener el umbral de audición promedio (a cierta frecuencia) de los trabajadores de una industria, debería elegirse la muestra a examinar procurando que las variables demográficas (edad, sector en el que trabajan, antigüedad, etc.) aparezcan en proporciones similares a las del total de la planta. Sería erróneo, en cambio, seleccionar todos los empleados en una misma sección o de cierto segmento etario, excepto si se pretenden resultados correspondientes a dichos subconjuntos.

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k, \quad (\text{A2.5})$$

donde  $n$  es el *tamaño de la muestra*. El *desvío estándar poblacional* se estima por medio del *desvío estándar muestral*, que se calcula en forma ligeramente diferente:

$$s = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n - 1}}. \quad (\text{A2.6})$$

La diferencia en el denominador se vuelve pequeña para muestras grandes.<sup>120</sup> Por ejemplo, para muestras de más de 10 elementos, la diferencia es menor al 5 %, y para muestras de más de 50 elementos, es menor al 1 %.

Aun cuando la variable aleatoria que hemos estado utilizando en los ejemplos (el nivel umbral de audición) es inherentemente *continua*, en el sentido de que sus valores pueden ser en principio números reales cualesquiera, para su tratamiento estadístico la transformamos en una variable *discreta*, ya que de cada subintervalo se tomó como representativo su valor central. Hay dos razones para ello. La primera es que en general no es posible una determinación tan precisa del umbral, ya que para una misma persona el umbral puede variar de un momento a otro y además el examen audiométrico ofrece varias dificultades.<sup>121</sup> La segunda razón es que al reducir la cantidad de valores posibles para la variable aleatoria se simplifica el manejo estadístico de los datos.

## A2.2 Probabilidad

A efectos de la modelización y el análisis teórico, en general es más conveniente trabajar con distribuciones continuas, especialmente cuando las poblaciones son grandes y de dimensiones indeterminadas. En ese caso se considera un concepto en cierta medida análogo al de distribución porcentual. Debido a que existen infinitos valores posibles, es poco probable que uno cualquiera de ellos corresponda a algún elemento de la población. Por ejemplo: ¿Cuántas personas tendrán un umbral de audición a 1 kHz exactamente igual a 4,572349 dB? Probablemente, ninguna. Es más razonable preguntarse cuántas personas tendrán un umbral entre 4 dB y 5 dB. Con mayor generalidad, podemos preguntarse qué proporción del total de la población tendrá un umbral entre  $x$  y  $x + \Delta x$ . En otras palabras, nos preguntamos qué *probabilidad* hay de que un individuo al azar tenga su umbral entre  $x$  y  $x + \Delta x$ . Es evidente que cuanto menor sea  $\Delta x$  menor será dicha probabilidad. Pero si dividimos la probabilidad por  $\Delta x$  obtendremos una magnitud que tiende (para cada  $x$ ) a un valor bien definido que denominamos *densidad de probabilidad*, es decir

$$f(x_0) = \lim_{\Delta x \rightarrow 0} \frac{P(x_0 < X < x_0 + \Delta x)}{\Delta x}, \quad (\text{A2.7})$$

<sup>120</sup> La razón para dividir por  $n - 1$  en lugar de por  $n$  es que de esa forma se obtiene un *estimador insesgado* de la varianza. Un estimador es insesgado si el promedio de todas las estimaciones obtenidas mediante el mismo a partir de diferentes muestras coincide con el parámetro a estimar. Nótese que el desvío estándar muestral no es un estimador insesgado del desvío estándar poblacional.

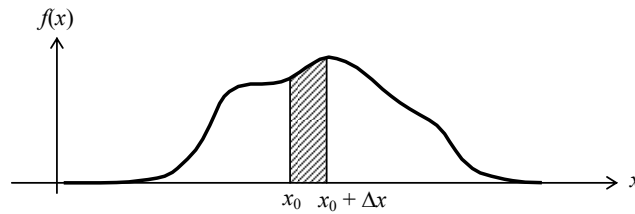
<sup>121</sup> Por ejemplo, la posición de los auriculares, o la atención y la fatiga del examinado, pueden afectar apreciablemente el resultado de las mediciones.

donde  $P(A)$  denota la probabilidad de que la condición  $A$  se cumpla.<sup>122</sup> Si consideramos como dominio todos los posibles valores de  $x_0$ , tendremos la *función de densidad de probabilidad*, ilustrada para un caso particular en la figura A2.3.

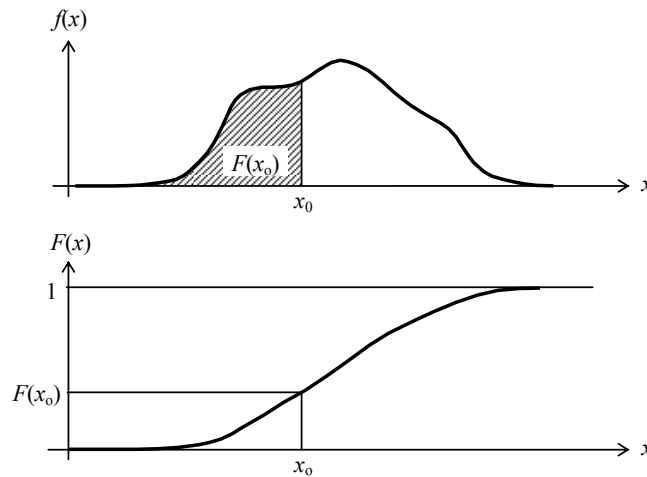
Definimos la *función de probabilidad acumulada*,  $F(x_0)$ , como la probabilidad de que  $x$  sea menor que  $x_0$ . Puede obtenerse como

$$F(x_0) = \int_{-\infty}^{x_0} f(x) dx. \quad (\text{A2.8})$$

En la figura A2.4 se muestra el ejemplo de función de probabilidad acumulada que corresponde a la densidad de probabilidad de la figura A2.3.



**Figura A2.3.** Un ejemplo de función de densidad de probabilidad. El área rayada representa la probabilidad de que la variable aleatoria se encuentre entre  $x_0$  y  $x_0 + \Delta x$ .



**Figura A2.4.** Función de densidad de probabilidad  $p(x)$  y su función de probabilidad acumulada. El área rayada representa la probabilidad de que la variable aleatoria sea menor que  $x_0$ .

<sup>122</sup> La probabilidad es 0 cuando la condición no se da nunca (por ejemplo que al arrojar en un dado se obtenga un 7, y es 1 cuando la condición se cumple siempre, por ejemplo que el resultado de arrojar el dado esté entre 1 y 6).



En el caso de las variables aleatorias continuas, el valor medio puede calcularse mediante la fórmula

$$\bar{x} = \int_{-\infty}^{\infty} x f(x) dx \quad (\text{A2.9})$$

Asimismo, el desvío estándar puede calcularse como

$$\bar{x} = \int_{-\infty}^{\infty} x f(x) dx \quad (\text{A2.10})$$

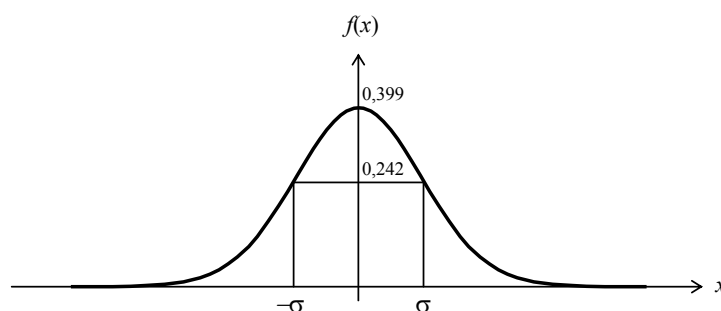
Estas expresiones son formalmente equivalentes a las que corresponden a las variables aleatorias discretas, habiéndose reemplazado las sumatorias por integrales.

### A2.3 Distribución normal

Un caso sumamente importante es la *distribución normal* o *distribución de Gauss*, cuya función de densidad de probabilidad es

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (\text{A2.11})$$

Esta función, mostrada en la figura A2.5, tiene un valor medio nulo y un desvío estándar igual a 1.



**Figura A2.5.** Función de densidad de probabilidad de la distribución normal (distribución de Gauss) con  $\mu = 0$  y  $\sigma = 1$ .

Es posible tener una distribución normal con valor medio  $\mu$  y desvío estándar  $\sigma$  por medio de la siguiente función de densidad de probabilidad:

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x - \mu)^2}{2 \sigma^2}}. \quad (\text{A2.12})$$

Una de las propiedades interesantes que tiene la distribución normal de Gauss es que el valor medio de las muestras de tamaño  $n$  de cualquier distribución (sea o no Gaussiana)

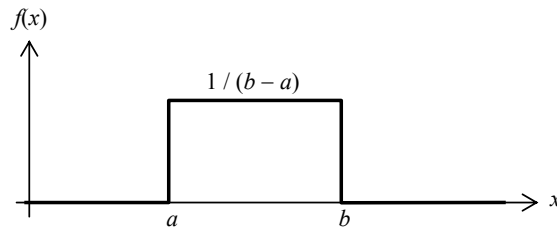
tiende a estar distribuida normalmente cuando  $n \rightarrow \infty$ . Esta propiedad, denominada *Teorema central del límite*, tiene importantes aplicaciones en la teoría de errores.

#### A2.4 Distribución uniforme

Otra distribución de interés en metrología es la distribución uniforme. En esta distribución la variable puede asumir valores en un intervalo acotado  $a < X < b$  y lo hace de tal manera que todos los valores son igualmente probables, lo cual significa que la densidad de probabilidad es constante en ese intervalo y 0 en el resto:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a \text{ ó } b < x \end{cases} \quad (\text{A2.13})$$

El valor  $1/(b-a)$  se adopta para que la probabilidad total sea 1. En la figura A2.6 se ilustra esta distribución.



**Figura A2.6.** Función de densidad de probabilidad de una distribución uniforme.

La distribución uniforme se presenta naturalmente en las mediciones con cualquier instrumento de resolución finita, particularmente, los instrumentos con indicación digital, aunque también aquéllos en los que se tiene una escala con divisiones. Al observar una indicación entre dos divisiones, sólo será posible, si no se dispone de alguna forma de interpolación (como en el caso del sistema Vernier), asignar el valor correspondiente a la división más cercana. Dado que no hay en principio razón para suponer que el mensurando debe estar a mayor o menor distancia de una división, la distribución de los posibles valores es equiprobable.

La media de esta distribución es la semisuma de los extremos,

$$\mu = \frac{a+b}{2}. \quad (\text{A2.14})$$

El desvío estándar es

$$\sigma = \sqrt{\int_a^b \frac{1}{b-a} \left(x - \frac{b+a}{2}\right)^2 dx} = \frac{b-a}{\sqrt{12}}. \quad (\text{A2.15})$$

### A2.5 Distribución de la media de una muestra

Si de una población se toma una muestra formada por  $n$  instancias de la variable aleatoria, la media de esos  $n$  valores es a su vez una nueva variable aleatoria denominada *media muestral*:

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k, \quad (\text{A2.16})$$

En otras palabras, si repetimos muchas veces la toma de una muestra y cada vez calculamos la media muestral, esas medias constituyen sucesivas instancias de una variable aleatoria.

Nos interesa conocer su distribución estadística, para lo cual hace falta conocer su media (la media de la media muestral) y su desvío estándar (el desvío estándar de la media muestral). La media coincide con la media de la población original:

$$\mu_{\bar{x}} = \mu. \quad (\text{A2.17})$$

El desvío estándar es igual al desvío estándar muestral dividido por  $\sqrt{n}$ :

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}. \quad (\text{A2.18})$$

Es decir, cuanto mayor sea la muestra, menor será el “error” cometido al estimar el valor medio poblacional con el valor medio muestral. Esta es la base para la inferencia estadística.

Estas propiedades se cumplen en general para cualquier variable aleatoria, ya sea su distribución normal o cualquier otra.

Ahora bien, en lo anterior supusimos que conocíamos  $\mu$  y  $\sigma$ , pero en la generalidad de los casos dichos valores son desconocidos y a lo sumo se los puede estimar. De hecho  $\mu$  se estima mediante la media muestral  $\bar{x}$  a través de la ecuación (A2.5) y  $\sigma$  se estima mediante el desvío estándar muestral de la ecuación (A2.6)

$$s = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n - 1}}. \quad (\text{A2.19})$$

Cuando  $n \gg 1$ , la estimación es muy buena, no así para tamaños pequeños de la muestra. El desconocimiento de la dispersión real de la que se parte tiende incrementar la dispersión que se calcularía por medio de (A2.18).

Se demuestra que la variable auxiliar normalizada

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \quad (\text{A2.20})$$

tiene una distribución conocida como distribución  $t$  de Student con  $v = n - 1$  grados de libertad. En la figura A2.7 se muestra la distribución  $t$  para tres grados de libertad y su

comparación con la distribución normal. Para  $v \geq 30$  se acepta en general que se puede reemplazar con ventaja por una distribución normal con media 0 y desvío estándar 1.

La función de densidad de probabilidad de la distribución  $t$  está dada por la ecuación siguiente (Smirnov et al., 1978)

$$f(t, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v} \Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}}, \quad (\text{A2.21})$$

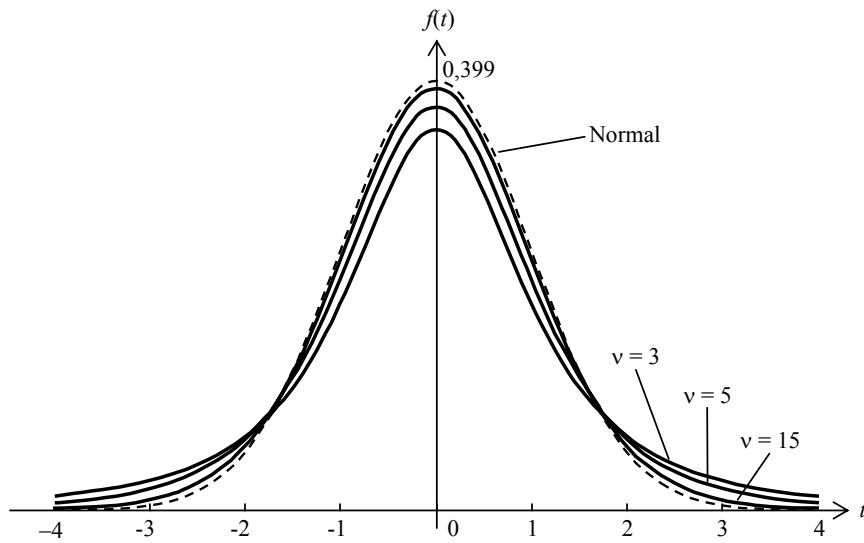
donde  $\Gamma(x)$  es la función gamma, definida como

$$\Gamma(x) = \int_0^{\infty} z^{x-1} e^{-z} dz. \quad (\text{A2.22})$$

La distribución  $t$  tiene media nula y desvío estándar (para  $v > 2$ )

$$\sigma_t = \sqrt{\frac{v}{v-2}}. \quad (\text{A2.23})$$

Este desvío estándar varía entre 1,73 para  $v = 3$  y 1 para  $v = \infty$ .

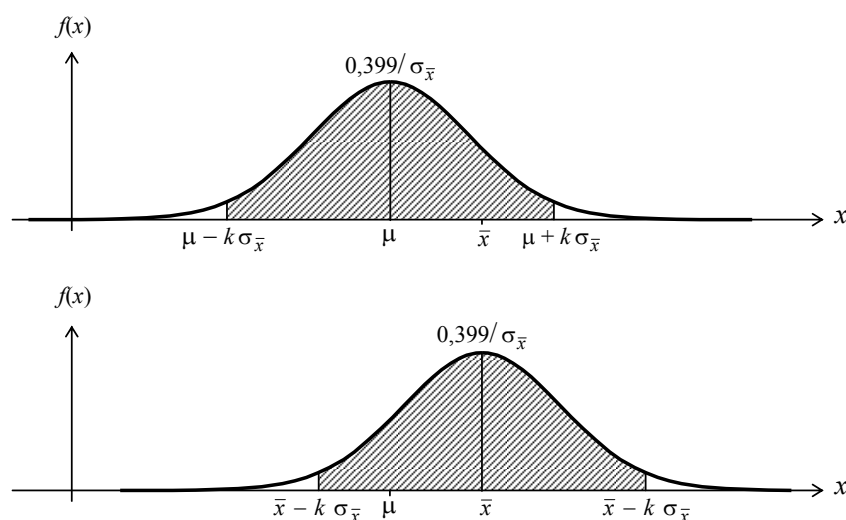


**Figura A2.7.** Función de densidad de probabilidad  $t$  de Student para tres diferentes grados de libertad, comparadas con la distribución normal (en línea de trazos).

## A2.5 Inferencia relativa a la media poblacional

La media muestral  $\mu_{\bar{x}}$  puede utilizarse para estimar la media de la población; y su distribución, particularmente su desvío estándar, para estimar el error de la estimación. Ahora bien, si la distribución de la media muestral es, por ejemplo,  $t$  de Student o normal, ello implica que podría tomar cualquier valor, incluso valores muy alejados de la verdadera media. ¿Significa esto que la estimación es completamente inútil? No, si podemos determinar un *intervalo de confianza*, esto es, un intervalo en el que se encuentra la media con una probabilidad especificada. Si la probabilidad es razonablemente alta (por ejemplo, 95 % ó 99 %), implica que sería un evento muy raro que nuestra estimación cayera fuera de ese intervalo.

Consideremos en primer lugar el caso en que se conoce  $\sigma$  o se la puede estimar con buena exactitud. Supongamos que queremos obtener un intervalo alrededor de nuestra estimación  $\bar{x}$  en el cual se encuentre la verdadera media  $\mu$  con probabilidad  $1 - \alpha$ , o, expresado de otra forma, de modo que la probabilidad de que *no* se encuentre en ese intervalo (probabilidad residual) sea  $\alpha$ . El problema es equivalente a buscar un intervalo de igual amplitud alrededor de  $\mu$  que contenga a la media estimada  $\bar{x}$  con una probabilidad  $1 - \alpha$ .<sup>123</sup> En particular significa, que podemos usar exactamente la misma distribución sólo que cambiando la media, como se muestra en la figura A2.8.



**Figura A2.8.** Dos interpretaciones de la proximidad entre la media poblacional  $\mu$  y la media muestral  $\bar{x}$ .

<sup>123</sup> En efecto, veamos que  $\mu$  y  $\bar{x}$  son intercambiables:

$$\begin{aligned} P(\bar{x} - L < \mu < \bar{x} + L) &= P(-L < \mu - \bar{x} < L) = \\ &= P(-\mu - L < -\bar{x} < -\mu + L) = \\ &= P(\mu - L < \bar{x} < \mu + L) \end{aligned}$$

Llamando  $z_{\alpha/2}$  al valor de la variable de la distribución normal estándar ( $\mu = 0$ ,  $\sigma = 1$ ) tal que  $P(x > z_{\alpha/2}) = \alpha/2$ , debido a la simetría el intervalo de confianza tal que  $P(|x| < z_{\alpha/2}) = 1 - \alpha$  puede expresarse como

$$\bar{x} - z_{\alpha/2} \sigma_{\bar{x}} < \mu < \bar{x} + z_{\alpha/2} \sigma_{\bar{x}},$$

o bien, teniendo en cuenta (A2.18),

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}. \quad (\text{A2.24})$$

donde  $\sigma$  es el desvío estándar poblacional.

Consideremos ahora el caso en que no se conoce  $\sigma$  pero se lo estima a partir del desvío estándar muestral de una muestra de tamaño  $n$ . La distribución es entonces la  $t$  de Student con  $n - 1$  grados de libertad. Llamando  $t_{n-1, \alpha/2}$  al valor de  $t$  tal que  $P(x > t_{n-1, \alpha/2}) = \alpha/2$ , el intervalo de confianza resulta

$$\bar{x} - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}. \quad (\text{A2.25})$$

En la tabla A2.2 se dan los valores de  $z_{\alpha/2}$  y  $t_{v, \alpha/2}$  para diversos valores de los parámetros (probabilidad residual  $\alpha$  y grados de libertad  $v$ )

**Tabla A2.2.** Valores de  $z_{\alpha/2}$  y  $t_{v, \alpha/2}$  para diversos valores de probabilidad residual  $\alpha$  y grados de libertad  $v$  para la determinación de intervalos de confianza para la media poblacional

$\alpha$	0,1	0,05	0,025	0,010	0,005	0,001
$z_{\alpha/2}$	1,6449	1,9600	2,2414	2,5758	2,8070	3,2905
$t_{2, \alpha/2}$	2,9200	4,3027	6,2053	9,9248	14,089	31,599
$t_{3, \alpha/2}$	2,3534	3,1824	4,1765	5,8409	7,4533	12,924
$t_{5, \alpha/2}$	2,0150	2,5706	3,1634	4,0321	4,7733	6,8688
$t_{10, \alpha/2}$	1,8125	2,2281	2,6338	3,1693	3,5814	4,5869
$t_{15, \alpha/2}$	1,7531	2,1314	2,4899	2,9467	3,2860	4,0728
$t_{20, \alpha/2}$	1,7247	2,0860	2,4231	2,8453	3,1534	3,8495
$t_{25, \alpha/2}$	1,7081	2,0595	2,3846	2,7874	3,0782	3,7251
$t_{30, \alpha/2}$	1,6973	2,0423	2,3596	2,7500	3,0298	3,6460
$t_{40, \alpha/2}$	1,6839	2,0211	2,3289	2,7045	2,9712	3,5510

**EJEMPLO 1:** Supongamos que se han medido los siguientes valores de nivel de presión sonora:

73,6	80,2	76,7	75,5	76,7	75,3	75,4	77,8	77,7	77,7
------	------	------	------	------	------	------	------	------	------

El valor medio es 76,66 dB y el desvío estándar muestral, 1,83 dB. El intervalo de confianza para una probabilidad del 95 % se determina usando la distribución  $t$  con 9 gra-

dos de libertad y  $\alpha = 0,05$ . Interpolando en la tabla A2.2 entre 5 y 10 grados de libertad resulta

$$t_{9, 0,05/2} \cong 2,2966,$$

de donde el intervalo de confianza que contiene la media poblacional con una probabilidad del 95 % será

$$76,66 - 2,2966 \frac{1,83}{\sqrt{10}} < \mu < 76,66 + 2,2966 \frac{1,83}{\sqrt{10}}$$

es decir,

$$75,33 \text{ dB} < \mu < 77,99 \text{ dB}.$$

### A2.6 Distribución de la varianza de una muestra

La varianza muestral, es decir el cuadrado del desvío estándar muestral, dado por

$$s^2 = \frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n - 1}, \quad (\text{A2.26})$$

es un *estimador insesgado* de la varianza poblacional. Insesgado significa que el promedio extendido un gran número de muestras tiende a la varianza poblacional. Nos interesa determinar cómo se distribuye estadísticamente. La media de  $s^2$ , por ser un estimador insesgado, es precisamente  $\sigma^2$ . Si la variable aleatoria de la cual se extrae la muestra es normal, entonces la variable  $\chi^2$ , definida como

$$\chi^2 = \frac{(n-1) s^2}{\sigma^2}, \quad (\text{A2.27})$$

responde a una distribución ji-cuadrada (también llamada a veces chi-cuadrada, según ortografía inglesa de la letra griega  $\chi$ ), con  $v = n - 1$  grados de libertad, dada por

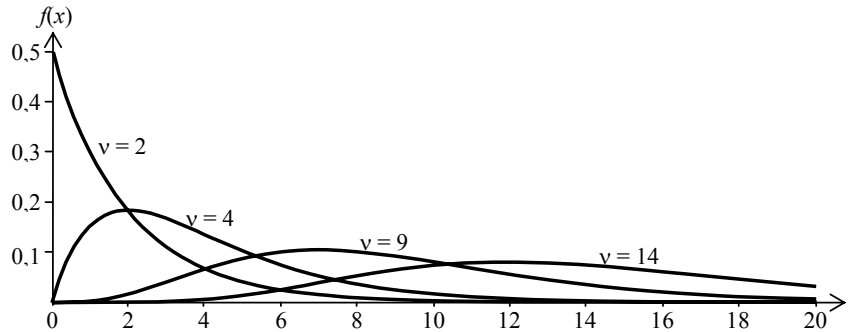
$$f(x, v) = \begin{cases} \frac{1}{2^{\frac{v}{2}} \Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (\text{A2.28})$$

El valor nulo para valores negativos se debe a que representa la distribución de una variable positiva, como lo es la varianza muestral.

La media de la distribución ji-cuadrado es  $v$  y su varianza es  $2v$ . En la figura A2.9 se muestra la densidad de probabilidad de esta distribución para algunos valores de  $v$ . Se observa la asimetría y la presencia de una moda (máximo) en  $v - 2$ .

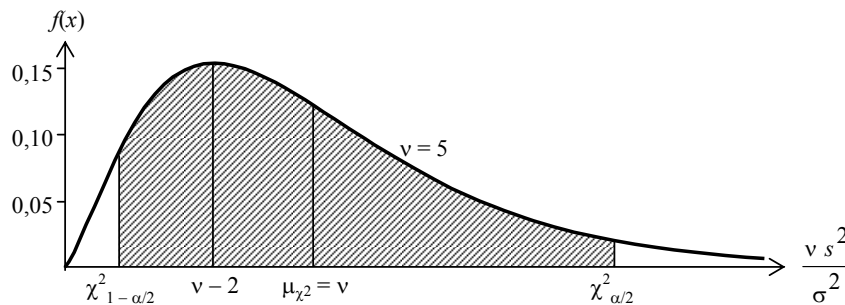
### A2.7 Inferencia relativa a la varianza poblacional

La varianza muestral  $s^2$  puede usarse para estimar la varianza de la población  $\sigma^2$ , y su distribución permite estudiar el error de la estimación. Igual que para la media, podemos obtener un intervalo de confianza donde se hallará la varianza con una probabilidad  $1 - \alpha$ .



**Figura A2.9.** Densidad de probabilidad de la distribución ji-cuadrado para cuatro valores de  $v$ .

Supongamos que queremos obtener un intervalo alrededor de la estimación  $s^2$  de la varianza dentro del cual se encuentre la verdadera varianza  $\sigma^2$  con una probabilidad  $1 - \alpha$ . Debido a que la distribución ji-cuadrado es asimétrica, deberemos obtener dos valores,  $\chi^2_{1-\alpha/2}$  y  $\chi^2_{\alpha/2}$  que describen los límites a uno y otro lado de la media, como se muestra en la figura A2.10. En efecto,  $\chi^2_{1-\alpha/2}$  deja a su derecha una fracción  $1 - \alpha/2$  de instancias de la variable  $\chi^2$ , en tanto que  $\chi^2_{\alpha/2}$  deja a su derecha una fracción  $\alpha/2$ . La diferencia es el valor buscado  $1 - \alpha$ .



**Figura A2.10.** Obtención de un intervalo de confianza para la varianza normalizada a partir de la distribución ji-cuadrado. Entre  $\chi^2_{1-\alpha/2}$  y  $\chi^2_{\alpha/2}$  se encuentra la varianza normalizada con probabilidad  $1 - \alpha$ . El ejemplo es para  $v = 5$  y  $\alpha = 0,1$ . La media se da en  $s^2 = \sigma^2$ .

Podemos expresar lo anterior en la forma,



$$\chi^2_{1-\alpha/2} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2},$$

que, operando para aislar  $\sigma^2$ , se transforma en

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}. \quad (\text{A2.29})$$

En la tabla A2.3 se dan valores de  $\chi^2_{1-\alpha/2}$  y  $\chi^2_{\alpha/2}$  para algunos valores de los parámetros  $\alpha$  y  $v$ .

**Tabla A2.3.** Valores de  $\chi^2_{1-\alpha/2}$  y  $\chi^2_{\alpha/2}$  para diversos valores de probabilidad residual  $\alpha$  y grados de libertad  $v$  para la determinación de intervalos de confianza para la varianza poblacional

$v$	$\alpha$	0,1	0,05	0,025	0,010	0,005	0,001
3	$\chi^2_{\alpha/2}$	7,8147	9,3484	10,8610	12,8380	14,3200	17,7300
	$\chi^2_{1-\alpha/2}$	0,3519	0,2158	0,1338	0,0717	0,0449	0,0153
5	$\chi^2_{\alpha/2}$	11,0700	12,8330	14,5440	16,7500	18,3860	22,1050
	$\chi^2_{1-\alpha/2}$	1,1455	0,8312	0,6109	0,4117	0,3075	0,1581
10	$\chi^2_{\alpha/2}$	18,3070	20,4830	22,5580	25,1880	27,1120	31,4200
	$\chi^2_{1-\alpha/2}$	3,9403	3,2470	2,7072	2,1559	1,8274	1,2650
15	$\chi^2_{\alpha/2}$	24,9960	27,4880	29,8430	32,8010	34,9500	39,7190
	$\chi^2_{1-\alpha/2}$	7,26090	6,2621	5,4569	4,6009	4,0697	3,1075
20	$\chi^2_{\alpha/2}$	31,4100	34,1700	36,7610	39,9970	42,3360	47,4980
	$\chi^2_{1-\alpha/2}$	10,8510	9,5908	8,5563	7,4338	6,7228	5,3981

EJEMPLO 2: Con los datos del ejemplo 1, determinemos el intervalo de confianza para tener una probabilidad del 95 % de que el mismo incluya el verdadero desvío estándar poblacional. En este caso  $n = 10$ , por lo tanto  $v = 9$ . Además,  $\alpha = 1 - 0,95 = 0,05$ . Los valores interpolados entre los correspondientes a  $v = 5$  y  $v = 10$  de la tabla son

$$\chi^2_{1-\alpha/2} = 3,3813$$

$$\chi^2_{\alpha/2} = 16,86$$

de donde

$$\frac{9 \times 1,83^2}{16,86} \text{ dB}^2 < \sigma^2 < \frac{9 \times 1,83^2}{3,3813} \text{ dB}^2.$$

Resulta, así, con una probabilidad del 95 %, que

$$1,377 \text{ dB} < \sigma < 2,986 \text{ dB}.$$

### Apéndice 3

#### Dispersión estadística del valor eficaz de un ruido estacionario en función del tiempo de integración

Consideremos un ruido blanco acotado en banda  $p(t)$  de valor eficaz  $P_{\text{ef}}$ . La densidad de probabilidad de sus valores instantáneos es

$$f(p) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{p^2}{2\sigma^2}}. \quad (\text{A3.1})$$

Por ergodicidad, el promedio temporal es igual al promedio, para un instante, extendido a todas las posibles realizaciones o instancias del ruido

$$\begin{aligned} P_{\text{ef}}^2 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} p^2(t) dt = \int_{-\infty}^{\infty} p^2 f(p) dp \\ P_{\text{ef}}^2 &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} p^2 e^{-\frac{p^2}{2\sigma^2}} dp \end{aligned} \quad (\text{A3.2})$$

Esta integral se puede calcular por partes descomponiendo  $p^2$  como producto de  $p$  por sí misma, y asignar una de ellas a acompañar a la exponencial, que entonces se integra fácilmente. Resulta  $P_{\text{ef}}^2 = \sigma^2$ , de donde

$$P_{\text{ef}} = \sigma. \quad (\text{A3.3})$$

Nos interesa calcular también la varianza de la presión cuadrática. Para ello tengamos en cuenta que por la propia definición de desvío estándar, el valor medio de  $p^2$  es, precisamente,  $\sigma^2$ . Tenemos

$$\begin{aligned} \sigma_{p^2}^2 &= \int_{-\infty}^{\infty} (p^2 - \sigma^2)^2 f(p) dp = \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} (p^2 - \sigma^2)^2 e^{-\frac{p^2}{2\sigma^2}} dp = \\ &= \frac{1}{\sqrt{2\pi} \sigma} \int_{-\infty}^{\infty} (p^4 - 2p^2\sigma^2 + \sigma^4) e^{-\frac{p^2}{2\sigma^2}} dp \end{aligned}$$

Esta integral se puede descomponer en tres integrales. Las dos primeras se calculan por partes y la tercera es la integral de la función de probabilidad normal. Resulta

$$\sigma_{p^2}^2 = 2\sigma^4,$$

de donde,

$$\sigma_{p^2} = \sqrt{2} \sigma^2 = \sqrt{2} P_{\text{ef}}^2. \quad (\text{A3.4})$$

Esta es la dispersión estadística de la presión cuadrática (aclaremos que la presión cuadrática no tiene distribución normal).

Nos interesa ahora calcular la dispersión estadística del valor eficaz extendido a un tiempo  $T$  en función del  $T$ . Para ello supongamos que el ruido blanco está limitado en banda a  $f_{\text{máx}}$ . Entonces es posible muestrearlo a una frecuencia de muestreo  $F_s = 2 f_{\text{máx}}$ . Llamando  $K = [F_s T]$ , donde  $[ ]$  es la parte entera, se puede calcular la presión eficaz cuadrática como

$$P_{\text{ef}, T}^2 = \frac{1}{K} \sum_{k=1}^K p^2(k).$$

Esta expresión es interesante porque permite obtener la presión eficaz cuadrática mediante una suma discreta de variables aleatorias  $p^2(k)$ . De hecho se trata de un promedio de variables aleatorias. El desvío estándar de una media es igual al desvío estándar de las variables individuales dividido por la raíz cuadrada de la cantidad de variables:

$$\sigma_{P_{\text{ef}, T}^2} = \sqrt{\frac{2}{K}} P_{\text{ef}, \infty}^2, \quad (\text{A3.5})$$

donde  $P_{\text{ef}, \infty}^2$  es el valor eficaz cuadrático a largo plazo. En otras palabras,

$$\sigma_{P_{\text{ef}, T}^2} \cong \sqrt{\frac{1}{f_{\text{máx}} T}} P_{\text{ef}, \infty}^2. \quad (\text{A3.6})$$

Vemos que la dispersión estadística de la presión eficaz cuadrática se reduce inversamente con la raíz cuadrada del tiempo de integración. Este fenómeno se conoce como *estabilización* de la presión eficaz y es útil para determinar el tiempo mínimo a medir para garantizar con cierta confiabilidad una cota de error dada.

## Apéndice 4

### Dispersión estadística del valor eficaz de un ruido no estacionario en función del tiempo de integración

Consideremos el ejemplo de un ruido de tránsito  $p(t)$ . Este ruido es la superposición de una serie de eventos constituidos por el paso de  $Q$  vehículos individuales por unidad de tiempo (flujo vehicular). Si la presión máxima medida a una distancia  $d$  de la trayectoria del vehículo es  $P_{\text{máx}}$  y su velocidad es  $v$ , se puede demostrar (Miyara, 1999) que la presión eficaz cuadrática extendida a un tiempo  $T$  centrado en el instante de mayor proximidad está dada por

$$P_{\text{ef},T}^2 = \frac{P_{\text{máx}}^2}{T} \frac{d}{v} 2 \arctg \frac{vT}{2d}. \quad (\text{A4.1})$$

Cuando  $T \gg 2d/v$ , el argumento de la arcotangente es muy grande y ésta tiende a  $\pi/2$ , de donde resulta

$$P_{\text{ef},T}^2 \cong P_{\text{máx}}^2 \frac{\pi d}{vT}. \quad (\text{A4.2})$$

Vemos, entonces, que para  $T$  grande la presión eficaz cuadrática es inversamente proporcional a  $T$ . Ello se debe a que virtualmente toda la energía acústica que aporta el vehículo al punto receptor ya ha sido emitida pero se reparte en un tiempo cada vez mayor.

La presión eficaz cuadrática difiere de un vehículo a otro debido a la dispersión estadística. Podemos suponer que la presión eficaz cuadrática correspondiente al  $k$ -ésimo vehículo es

$$P_{\text{ef},k,T}^2 = K_k P_{\text{ef},0,T}^2 \quad (\text{A4.3})$$

donde  $P_{\text{ef},0,T}^2$  es el valor central y  $K_k$  una variable aleatoria adimensional independiente del tiempo de integración<sup>124</sup> de valor medio 1. Podemos pasar a nivel equivalente dividiendo por la presión de referencia al cuadrado y aplicando 10 log:

$$L_{\text{eq},k,T} = L_{\text{eq},0,T} + 10 \log K_k = L_{\text{eq},0,T} + \Delta L_{\text{eq},k},$$

donde

$$\Delta L_{\text{eq},k} = 10 \log K_k.$$

Esto permite reescribir (A3.3) como

---

<sup>124</sup>  $K$  depende esencialmente de la potencia acústica, la velocidad, y la distancia, que son variables aleatorias.

$$P_{\text{ef},k,T}^2 = P_{\text{ef},o,T}^2 10^{\Delta L_{\text{eq},k}/10} \quad (\text{A4.4})$$

La exponencial puede aproximarse linealmente cuando el exponente es pequeño:

$$P_{\text{ef},k,T}^2 = P_{\text{ef},o,T}^2 \left( 1 + \frac{\ln 10}{10} \Delta L_{\text{eq},k} \right) \quad (\text{A4.5})$$

Esta expresión permite obtener la presión eficaz cuadrática debida a los  $Q.T$  vehículos que circularon durante el tiempo  $T$  por superposición energética de fuentes incoherentes:

$$P_{\text{ef},T}^2 = \sum_{k=1}^{Q.T} P_{\text{ef},k,T}^2 = Q.T P_{\text{ef},o,T}^2 + P_{\text{ef},o,T}^2 \frac{\ln 10}{10} \sum_{k=1}^{Q.T} \Delta L_{\text{eq},k}$$

Podemos expresar esta ecuación en términos de la ecuación (A3.2)

$$\begin{aligned} P_{\text{ef},T}^2 &= Q.T P_{\text{máx},o}^2 \frac{\pi d_o}{v_o T} + P_{\text{máx},o}^2 \frac{\pi d_o}{v_o T} \frac{\ln 10}{10} \sum_{k=1}^{Q.T} \Delta L_{\text{eq},k} = \\ &= Q P_{\text{máx},o}^2 \frac{\pi d_o}{v_o} + Q P_{\text{máx},o}^2 \frac{\pi d_o}{v_o} \frac{\ln 10}{Q.T} \sum_{k=1}^{Q.T} \Delta L_{\text{eq},k} = \\ &= P_{\text{ef},\infty}^2 \left( 1 + \frac{1}{Q.T} \frac{\ln 10}{10} \sum_{k=1}^{Q.T} \Delta L_{\text{eq},k} \right) \\ &= P_{\text{ef},\infty}^2 \left( 1 + \frac{1}{Q.T} \frac{\ln 10}{10} \Delta L_{\text{eq}} \right) \end{aligned}$$

donde  $P_{\text{ef},\infty}$  es la presión eficaz a largo plazo. El incremento  $\Delta L_{\text{eq}}$  tiene un desvío estándar  $\sqrt{Q.T}$  veces mayor que los incrementos  $\Delta L_{\text{eq},k}$  de cada vehículo. Podemos reescribir la ecuación anterior en términos del nivel equivalente:

$$L_{\text{eq},T} = L_{\text{eq},\infty} + \frac{\Delta L_{\text{eq}}}{Q.T} = L_{\text{eq},\infty} + \Delta L.$$

Podemos acotar con una confiabilidad del 99,7 %

$$|\Delta L_{\text{eq}}| < 3 \sigma_{\Sigma L_{\text{eq},k}} = 3\sqrt{Q.T} \sigma_{L_{\text{eq},k}}$$

por lo tanto

$$|\Delta L| = \left| \frac{\Delta L_{\text{eq}}}{Q T} \right| < \frac{3}{\sqrt{Q T}} \sigma_{L_{\text{eq},k}}$$

Vemos, así, que el error al integrar durante un tiempo finito  $T$  con respecto al valor que se obtendría integrando durante un tiempo infinito disminuye con la raíz cuadrada del tiempo de integración  $T$  y con la raíz cuadrada del número de vehículos por unidad de tiempo,  $Q$ .

Puede obtenerse el tiempo de estabilización para tener un error menor que  $\Delta L_{\text{máx}}$  planteando

$$\frac{3}{\sqrt{Q T}} \sigma_{L_{\text{eq},k}} < \Delta L_{\text{máx}}$$

Resulta:

$$T_{\text{estab}} = \frac{9}{Q} \left( \frac{\sigma_{L_{\text{eq},k}}}{\Delta L_{\text{máx}}} \right)^2$$

## Apéndice 5

### Dispersión estadística de los percentiles

Supondremos una variable aleatoria  $X$  con distribución normal de media  $\mu$  y desvío estándar  $\sigma$ , es decir, cuya función de densidad de probabilidad es

$$f(X) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(X-\mu)^2}{2\sigma^2}}. \quad (\text{A5.1})$$

Podemos hallar los percentiles obteniendo primero la función de probabilidad acumulada

$$F(X) = \int_{-\infty}^X \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx. \quad (\text{A5.2})$$

El percentil  $p$  se puede obtener como el valor  $X_p$  tal que

$$F(X_p) = \frac{p}{100}, \quad (\text{A5.3})$$

es decir

$$X_p = F^{-1}\left(\frac{p}{100}\right). \quad (\text{A5.4})$$

La función  $F^{-1}$  puede obtenerse a partir de la *función error inversa*, que está disponible en los paquetes de software matemático. La *función error* se define como

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx, \quad (\text{A5.5})$$

por lo que luego de alguna manipulación algebraica se puede expresar  $F(X)$  en términos de  $\text{erf}(x)$ , lo cual permite despejar

$$X_p = \sqrt{2}\sigma \text{erf}^{-1}\left(2\frac{p}{100} - 1\right) + \mu. \quad (\text{A5.6})$$

$X_p$  es el percentil  $p$  poblacional de la variable  $X$ , es decir, extendido a la totalidad de las realizaciones de  $X$ . ¿Qué sucede si en lugar de analizar la población analizamos una muestra de  $n$  valores de  $X$ , es decir,  $X_1, \dots, X_n$ ? Para cada muestra se obtendrá un percentil  $p$  diferente. La media de estos percentiles muestrales coincide con el valor anterior, y

su variabilidad puede expresarse en términos del desvío estándar mediante la siguiente ecuación (Kendall y Stuart, 1977):

$$\sigma_{X_p} = \sqrt{\frac{\frac{p}{100} \left(1 - \frac{p}{100}\right)}{n}} \frac{1}{f(X_p)}. \quad (\text{A5.7})$$

Esta fórmula indica que, para un dado percentil muestral  $p$  el desvío estándar se reduce con la recíproca de la raíz cuadrada del tamaño de la muestra. También se observa que para percentiles extremos, si bien el numerador tiende a 0, el denominador lo hace más rápidamente por lo cual la dispersión estadística es mayor.

Particularmente, el mínimo (percentil 0) y el máximo (percentil 100) están sujetos a una variabilidad infinita, razón por la que, cualquiera sea el tamaño de la muestra, no tienen ningún tipo de representatividad.



## Apéndice 6

### Envolvente de un ruido filtrado

Sea  $n(t)$  un ruido blanco gaussiano y supongamos que es filtrado mediante un filtro pasabanda ideal con frecuencias de corte  $f_1$  y  $f_2$ . La respuesta  $x(t)$  se obtiene como la convolución con la respuesta al impulso  $h(t)$  del filtro:

$$x(t) = n(t) * h(t) \quad (\text{A6.1})$$

La respuesta al impulso es

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) e^{j\omega t} dt = \frac{1}{2\pi} \left( \int_{2\pi f_1}^{2\pi f_2} e^{j\omega t} dt + \int_{-2\pi f_2}^{-2\pi f_1} e^{j\omega t} dt \right).$$

Calculando las integrales se obtiene

$$h(t) = 2(f_2 - f_1) \cos \left( 2\pi \frac{f_2 + f_1}{2} t \right) \frac{\sin \left( 2\pi \frac{f_2 - f_1}{2} t \right)}{2\pi \frac{f_2 - f_1}{2} t} \quad (\text{A6.2})$$

Si  $\Delta f = f_2 - f_1$  es pequeña, la envolvente de  $x(t)$  tiene una distribución aproximadamente de Rayleigh. Dicha envolvente está dada por<sup>125</sup>

$$Y(t) = \sqrt{2 \int_{-\infty}^t x^2(\theta) e^{-\frac{\theta-t}{\tau}} d\theta} \quad (\text{A6.3})$$

La función de densidad de probabilidad de Rayleigh para  $Y$  es

$$f(Y) = \frac{Y}{\sigma^2} e^{-\frac{Y^2}{2\sigma^2}}, \quad (\text{A6.4})$$

donde  $\sigma$  es el desvío estándar del ruido filtrado. El desvío estándar coincide con el valor eficaz, que puede calcularse integrando la densidad espectral (que para un ruido blanco filtrado con un filtro ideal es constante en la banda de paso):

$$\sigma = \sqrt{\int_{f_1}^{f_2} n^2(f) df} = \sqrt{(f_2 - f_1) n^2} \quad (\text{A6.5})$$

<sup>125</sup> Es la envolvente energética obtenida filtrando  $x^2(t)$  con un filtro pasabajos de primer orden con constante de tiempo  $\tau$ , (que se calcula como la convolución con la respuesta al impulso), multiplicada por  $\sqrt{2}$  para obtener la envolvente de los máximos (esto es posible porque al ser  $\Delta f$  pequeño el ruido filtrado se comporta localmente como una senoide de amplitud lentamente variable).

Si el ruido blanco antes de pasar por el filtro está acotado en banda entre 0 y  $f_{\text{máx}}$  y su valor eficaz es  $N_{\text{ef}}$ , entonces

$$\sigma = N_{\text{ef}} \sqrt{\frac{f_2 - f_1}{f_{\text{máx}}}}. \quad (\text{A6.6})$$

La distribución de Rayleigh tiene esperanza

$$\mu_Y = \sqrt{\frac{\pi}{2}} \sigma = 1,2533 \sigma, \quad (\text{A6.7})$$

y desvío estándar

$$\sigma_Y = \sqrt{\frac{4 - \pi}{2}} \sigma = 0,6551 \sigma. \quad (\text{A6.8})$$

## Apéndice 7

### Respuesta transitoria de un filtro pasabanda de orden 3

Si la función de transferencia de un filtro pasabanda de tercer orden está dada por

$$G(s) = K \prod_{k=1}^3 \frac{a_{1k}s}{s^2 + a_{1k}s + a_{2k}}, \quad (\text{A7.1})$$

al excitarlo con una entrada senoidal de frecuencia angular  $\omega_0$  la transformada de Laplace de la respuesta será

$$R(s) = K \frac{\omega_0}{s^2 + \omega_0^2} \prod_{k=1}^3 \frac{a_{1k}s}{s^2 + a_{1k}s + a_{2k}}. \quad (\text{A7.2})$$

Esta expresión puede desarrollarse en fracciones simples de la siguiente forma:

$$R(s) = \frac{A_{01} + A_{02}s}{s^2 + \omega_0^2} + \sum_{k=1}^3 \frac{A_{k1} + A_{k2}s}{s^2 + a_{1k}s + a_{2k}} \quad (\text{A7.3})$$

Para obtener los coeficientes  $A_{kn}$  se suman las fracciones de (A7.3) formando un denominador común y se iguala el correspondiente numerador al de la ecuación (A7.2). Se obtiene un sistema de 8 ecuaciones y 8 incógnitas del que se pueden despejar los  $A_{kn}$ .

Resulta entonces fácil obtener la transformada inversa, que toma la forma

$$r(t) = B_0 \sin(\omega_0 t + \varphi) + \sum_{k=1}^3 B_k e^{-t/\tau_k} \sin(\omega_k t + \psi_k) \quad (\text{A7.4})$$

donde

$$\omega_k = \sqrt{a_{k2} - \left(\frac{a_{k1}}{2}\right)^2} \quad (\text{A7.5})$$

y

$$\tau_k = \frac{2}{a_{k1}} \quad (\text{A7.6})$$

Las amplitudes  $B_k$  y los desfases  $\varphi$  y  $\psi_k$  dependen de las constantes  $A_{kn}$ , aunque en general carecen de importancia.  $B_0$  y  $\varphi$  están dados por la respuesta en frecuencia en

amplitud y fase del filtro, y pueden obtenerse sustituyendo  $s = j\omega$  en (A7.2). Utilizando la notación de la ecuación (1.58) y reordenando,

$$\omega_k = \frac{\omega_{ok}}{2Q} \sqrt{4Q_k^2 - 1} \quad (\text{A7.7})$$

y

$$\tau_k = \frac{2Q_k}{\omega_{ok}} \quad (\text{A7.8})$$

## Apéndice 8

### Incertidumbre combinada

Sea  $Y$  una variable aleatoria que depende funcionalmente de otras  $N$  variables aleatorias independientes:

$$Y = g(X_1, \dots, X_N). \quad (\text{A8.1})$$

Dado un valor o instancia  $(x_{10}, \dots, x_{N0})$  de la  $n$ -upla  $(X_1, \dots, X_N)$ , podemos expresar  $Y$  en las proximidades de  $(x_{10}, \dots, x_{N0})$  mediante una aproximación de Taylor  $N$ -dimensional de primer orden:

$$Y \cong g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0})(X_i - x_{i0}). \quad (\text{A8.2})$$

Analizaremos dos casos: poblacional y muestral. En el primero se trabajará con los parámetros poblacionales y en el segundo con los muestrales.

#### A8.1 Incertidumbre combinada poblacional

Veamos el caso poblacional. Primeramente tengamos en cuenta que podemos definir una densidad de probabilidad conjunta  $f(x_1, \dots, x_N)$  tal que

$$P(x_{i0} < x_i < x_{i0} + dx_i, i = 1, \dots, N) = f(x_{10}, \dots, x_{N0}) dx_1 \dots dx_N \quad (\text{A8.3})$$

Tenemos, entonces, que la media de una variable aleatoria  $Y$  que depende de  $X_1, \dots, X_N$  será

$$\mu_Y = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y(x_1, \dots, x_N) f(x_1, \dots, x_N) dx_1 \dots dx_N \quad (\text{A8.4})$$

Reemplazando,

$$\mu_Y = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left( g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0})(x_i - x_{i0}) \right) f(x_1, \dots, x_N) dx_1 \dots dx_N$$

Teniendo en cuenta que la media de  $g(x_{10}, \dots, x_{N0})$ , por ser constante, es igual a ella, y conmutando la suma y la integración,

$$\begin{aligned} \mu_Y &= g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - x_{i0}) f(x_1, \dots, x_N) dx_1 \dots dx_N = \\ &= g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) (\mu_i - x_{i0}) \end{aligned}$$

Si elegimos las coordenadas del punto alrededor del cual se desarrolla iguales a la media de cada variable, es decir,  $\bar{x}_i = x_{i0}$ , resulta que la media es igual a la función  $g$  calculada en las medias de las variables independientes.

$$\mu_Y = g(\mu_1, \dots, \mu_N) \quad (\text{A8.5})$$

Análogamente, su varianza (desvío estándar cuadrático) será

$$\begin{aligned} \sigma_Y^2 &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (y(x_1, \dots, x_N) - \mu_Y)^2 f(x_1, \dots, x_N) dx_1 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left( \sum_{i=1}^N \frac{\partial g}{\partial x_i}(\mu_i) (x_i - \mu_i) \right)^2 f(x_1, \dots, x_N) dx_1 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left( \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\mu_i) \frac{\partial g}{\partial x_j}(\mu_j) (x_i - \mu_i)(x_j - \mu_j) \right) f(x_1, \dots, x_N) dx_1 \dots dx_N \end{aligned}$$

Invirtiendo el orden entre la integración y las sumatorias

$$\sigma_Y^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\mu_i) \frac{\partial g}{\partial x_j}(\mu_j) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - \mu_i)(x_j - \mu_j) f(x_1, \dots, x_N) dx_1 \dots dx_N$$

Las integrales son, por definición, las covarianzas poblacionales entre  $X_i$  y  $X_j$ , de donde

$$\sigma_Y^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\mu_i) \frac{\partial g}{\partial x_j}(\mu_j) \text{cov}(X_i, X_j) \quad (\text{A8.6})$$

El análisis poblacional es útil para determinación de incertidumbre combinada tipo B, cuando es posible conocer la matriz de covarianzas por algún método previo a la toma de muestras, por ejemplo algún estudio técnico o científico anterior.

### A8.2 Incertidumbre combinada muestral

Supongamos ahora que hemos hecho  $n$  mediciones de las  $n$  variables, es decir, que disponemos de  $n$  instancias de la  $N$ -upla,  $(x_{1k}, \dots, x_{Nk})$ , donde  $k = 1, \dots, n$ . Podemos expresar, para la  $k$ -ésima instancia,

$$y_k \cong g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) (x_{ik} - x_{i0}). \quad (\text{A8.7})$$

El valor medio de los valores  $y_k$  vale

$$\bar{y} = \frac{1}{n} \sum_{k=1}^n y_k \cong g(x_{10}, \dots, x_{N0}) + \frac{1}{n} \sum_{k=1}^n \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) (x_{ik} - x_{i0}). \quad (\text{A8.8})$$

Podemos invertir el orden de las sumatorias y reordenar, obteniendo

$$\begin{aligned} \bar{y} &\cong g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) \frac{1}{n} \sum_{k=1}^n (x_{ik} - x_{i0}). \\ \bar{y} &\cong g(x_{10}, \dots, x_{N0}) + \sum_{i=1}^N \frac{\partial g}{\partial x_i}(x_{i0}) (\bar{x}_i - x_{i0}). \end{aligned} \quad (\text{A8.9})$$

Particularmente, si elegimos  $x_{i0} = \bar{x}_i$  (es decir, buscamos una aproximación para valores de los argumentos de  $g$  iguales a los valores medios de los valores medidos de cada argumento), resultará

$$\bar{y} \cong g(\bar{x}_1, \dots, \bar{x}_N). \quad (\text{A8.10})$$

En otras palabras, la función evaluada en los estimadores de los mensurandos es un estimador del valor esperado de  $y$ .

Determinemos ahora la varianza muestral,  $s_y^2$ , de los  $y_k$ :

$$\begin{aligned} s_y^2 &= \frac{1}{n-1} \sum_{k=1}^n (y_k - \bar{y})^2 = \frac{1}{n-1} \sum_{k=1}^n \left( \sum_{i=1}^N \frac{\partial g}{\partial x_i}(\bar{x}_i) (x_{ik} - \bar{x}_i) \right)^2 = \\ &= \frac{1}{n-1} \sum_{k=1}^n \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\bar{x}_i) \frac{\partial g}{\partial x_j}(\bar{x}_j) (x_{ik} - \bar{x}_i) (x_{jk} - \bar{x}_j) \end{aligned}$$

Introduciendo la sumatoria respecto a  $k$  dentro de las otras dos se puede escribir

$$s_y^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\bar{x}_i) \frac{\partial g}{\partial x_j}(\bar{x}_j) \frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i) (x_{jk} - \bar{x}_j)$$

La sumatoria

$$\frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i) (x_{jk} - \bar{x}_j) \quad (\text{A8.11})$$

es la covarianza muestral  $c_{ij}$  entre las variables  $x_i$  y  $x_j$ , es decir, un estimador de la covarianza poblacional  $\text{cov}(X_i, X_j)$ , de modo que podemos escribir

$$s_y^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial g}{\partial x_i}(\bar{x}_i) \frac{\partial g}{\partial x_j}(\bar{x}_j) c_{ij}, \quad (\text{A8.12})$$

Esta fórmula resulta idéntica a la (A8.6) para el caso poblacional, sólo que se reemplazan las covarianzas por sus estimadores. Es útil para el tratamiento de componentes de incertidumbre tipo A, si las covarianzas muestrales pueden obtenerse a partir de los datos.

Si las variables son independientes sólo resultan no nulas las covarianzas de cada variable consigo mismas, es decir, las varianzas, por lo que

$$s_y^2 = \sum_{i=1}^N \left( \frac{\partial g}{\partial x_i}(\bar{x}_i) \right)^2 s_i^2. \quad (\text{A8.13})$$

NOTA: Aun si las variables son independientes, al calcular las covarianzas muestrales entre pares de variables diferentes el resultado no será nulo. La razón es la misma por la cual si se toma una muestra de ruido blanco durante un tiempo corto, su valor medio muy probablemente será diferente de 0.



## Apéndice 9

### Ejemplo de cálculo de incertidumbre en caso de un error sistemático no lineal

Consideremos un sistema de adquisición de datos que registra valores de presión sonora  $p$  a partir de la conversión analógica/digital de la tensión  $v$  de un micrófono y posterior aplicación de la ecuación

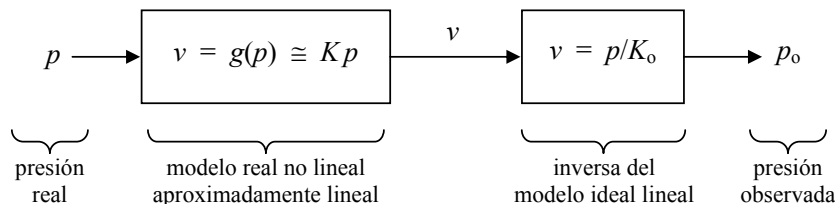
$$p = K_o v \quad (\text{A9.1})$$

donde  $K_o$  es la sensibilidad nominal del micrófono. Supongamos que la sensibilidad tiene un valor de 0,040 V/Pa y una incertidumbre estándar de 0,002 V/Pa y que las mediciones de presión sonora tienen una incertidumbre equivalente a 0,4 dB (lo cual implicará que la incertidumbre en la presión es proporcional a ésta).

La relación entre  $p$  y  $v$  es, en realidad, no lineal, a través de una ecuación

$$p = Kv + K_1v^2 + K_2v^3 \quad (\text{A9.2})$$

El modelo de medición es el ilustrado en el diagrama de bloques de la figura A9.1.



**Figura A9.1.** Modelo de la medición a partir de un proceso de transducción no lineal pero aproximadamente lineal a través de una constante  $K$ .

Disponemos de tres referencias nominales de 74 dB, 94 dB y 114 dB con una incertidumbre estándar de 0,2 dB en todos los casos y supondremos que las mismas no poseen error sistemático respecto a lo especificado. Dichos valores corresponden a los valores de presión sonora que se indican en la tercera columna de la tabla A9.1.

**Tabla A9.1.** Nivel de presión sonora de referencia nominal, presión de referencia nominal, presión de referencia medida y nivel de presión sonora que le correspondería a la presión de referencia medida

$k$	$L_{pk}$ [dB]	$P_k$ [Pa]	$P_{ko}$ [Pa]	$L_{pko}$ [dB]
1	74,0	0,10024	0,10029	74,004
2	94,0	1,0024	1,0066	94,036
3	114,0	10,024	9,7204	113,73

Proponemos un modelo inverso también cúbico para obtener el valor corregido de la presión sonora en función de la presión sonora medida:

$$p \cong ap_o + bp_o^2 + cp_o^3. \quad (\text{A9.3})$$

Plantearemos ahora el sistema de ecuaciones (2.72), pero a efectos de su resolución numérica lo escribiremos en forma matricial

$$\begin{bmatrix} P_{1o} & P_{1o}^2 & P_{1o}^3 \\ P_{2o} & P_{2o}^2 & P_{2o}^3 \\ P_{3o} & P_{3o}^2 & P_{3o}^3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix}. \quad (\text{A9.4})$$

Este sistema es de la forma

$$A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = B,$$

donde, hechos los cálculos,

$$A = \begin{bmatrix} 0,10029 & 0,010057 & 0,0010086 \\ 1,0066 & 1,0132 & 1,0199 \\ 9,7204 & 94,486 & 918,45 \end{bmatrix}$$

$$B = \begin{bmatrix} 0,10024 \\ 1,0024 \\ 10,024 \end{bmatrix}$$

Resulta<sup>126</sup>

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = A^{-1}B = \begin{bmatrix} 1 \\ -0,0050166 \text{ Pa}^{-1} \\ 0,00084634 \text{ Pa}^{-2} \end{bmatrix} \quad (\text{A9.5})$$

Con esto disponemos del modelo para corregir los efectos sistemáticos no lineales. Debemos ahora calcular la incertidumbre. Para ello necesitaremos las 9 derivadas parciales de  $a$ ,  $b$  y  $c$  con respecto a  $P_{1,o}$ ,  $P_{2,o}$ ,  $P_{3,o}$ . Notemos que  $a$ ,  $b$  y  $c$  son funciones implícitas de dichas variables definidas por el sistema de ecuaciones (A9.3). En lugar de intentar despejarlas analíticamente y luego derivar, lo cual sería tedioso, conviene obtener las derivadas implícitamente. El sistema (A9.4) puede escribirse como

<sup>126</sup> Para no sobrecargar la escritura, en los resultados anteriores se omitieron las unidades por ser resultados intermedios

$$\begin{cases} f_1(a, b, c, P_{1,o}) = 0 \\ f_2(a, b, c, P_{2,o}) = 0 \\ f_3(a, b, c, P_{3,o}) = 0 \end{cases} \quad (\text{A9.6})$$

donde, para  $k = 1, 2, 3$ ,

$$f_k = P_{k,o} a + P_{k,o}^2 b + P_{k,o}^3 c - P_k. \quad (\text{A9.7})$$

Dado que

$$\begin{cases} a = a(P_{1,o}, P_{2,o}, P_{3,o}) \\ b = b(P_{1,o}, P_{2,o}, P_{3,o}) \\ c = c(P_{1,o}, P_{2,o}, P_{3,o}) \end{cases} \quad (\text{A9.8})$$

podemos derivar cada  $f_k$  como función compuesta con respecto a las tres variables  $P_{1,o}$ ,  $P_{2,o}$ ,  $P_{3,o}$ , utilizando la regla de la cadena. Esta indica que, dadas, por ejemplo, unas funciones tales como  $f(x, y, z)$ ,  $x(u, v)$ ,  $y(u, v)$ , entonces las derivadas de su función compuesta

$$g(u, v) = f(x(u, v), y(u, v), z(u, v))$$

serán

$$\frac{\partial}{\partial u} g(u, v) = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial u}$$

$$\frac{\partial}{\partial v} g(u, v) = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial v}$$

Luego, para las derivadas con respecto a  $P_{\text{ref } 1, o}$ , se forma el sistema de 3 ecuaciones con 3 incógnitas que permiten despejar las derivadas  $\partial a / \partial P_{1,o}$ ,  $\partial b / \partial P_{1,o}$ ,  $\partial c / \partial P_{1,o}$ :

$$\begin{cases} \frac{\partial f_1}{\partial a} \frac{\partial a}{\partial P_{1,o}} + \frac{\partial f_1}{\partial b} \frac{\partial b}{\partial P_{1,o}} + \frac{\partial f_1}{\partial c} \frac{\partial c}{\partial P_{1,o}} + \frac{\partial f_1}{\partial P_{1,o}} = 0 \\ \frac{\partial f_2}{\partial a} \frac{\partial a}{\partial P_{1,o}} + \frac{\partial f_2}{\partial b} \frac{\partial b}{\partial P_{1,o}} + \frac{\partial f_2}{\partial c} \frac{\partial c}{\partial P_{1,o}} = 0 \\ \frac{\partial f_3}{\partial a} \frac{\partial a}{\partial P_{1,o}} + \frac{\partial f_3}{\partial b} \frac{\partial b}{\partial P_{1,o}} + \frac{\partial f_3}{\partial c} \frac{\partial c}{\partial P_{1,o}} = 0 \end{cases} \quad (\text{A9.9})$$

El último término de la primera ecuación aparece porque  $f_1$  depende de  $P_{1,o}$ , no así  $f_2$  ni  $f_3$ . El sistema puede reescribirse matricialmente como

$$\begin{bmatrix} \frac{\partial f_1}{\partial a} & \frac{\partial f_1}{\partial b} & \frac{\partial f_1}{\partial c} \\ \frac{\partial f_2}{\partial a} & \frac{\partial f_2}{\partial b} & \frac{\partial f_2}{\partial c} \\ \frac{\partial f_3}{\partial a} & \frac{\partial f_3}{\partial b} & \frac{\partial f_3}{\partial c} \end{bmatrix} \begin{bmatrix} \frac{\partial a}{\partial P_{1o}} \\ \frac{\partial b}{\partial P_{1o}} \\ \frac{\partial c}{\partial P_{1o}} \end{bmatrix} = \begin{bmatrix} -\frac{\partial f_1}{\partial P_{1o}} \\ 0 \\ 0 \end{bmatrix} \quad (\text{A9.10})$$

Las derivadas que intervienen en la matriz del sistema se calculan como

$$\frac{\partial f_1}{\partial a} = P_{1o} \quad (\text{A9.11})$$

$$\frac{\partial f_1}{\partial b} = P_{1o}^2 \quad (\text{A9.12})$$

$$\frac{\partial f_1}{\partial c} = P_{1o}^3 \quad (\text{A9.13})$$

directamente a partir de los datos. Vemos que resulta la misma matriz del sistema para el cálculo de  $a$ ,  $b$  y  $c$ . La derivada que aparece en el término independiente es

$$\frac{\partial f_1}{\partial P_{1o}} = a + 2P_{1o}b + 3P_{1o}^2c, \quad (\text{A9.14})$$

Además de los datos figuran aquí también  $a$ ,  $b$ , y  $c$ , que se han obtenido numéricamente en (A9.5). Es posible, entonces, determinar las derivadas necesarias para cálculos de incertidumbre mediante la resolución numérica del sistema (A9.10) y otros dos análogos para las derivadas respecto a  $P_{2o}$  y  $P_{3o}$ . Los valores obtenidos por software matricial se encuentran en la tabla A9.2.

**Tabla A9.2.** Valores numéricos necesarios para determinar la incertidumbre

Derivada	$a$	$b$	$c$
$\frac{\partial}{\partial P_{1o}}$	-11,179	12,256	-1,1425
$\frac{\partial}{\partial P_{2o}}$	0,1217	-1,2261	0,12485
$\frac{\partial}{\partial P_{3o}}$	-0,00014341	0,0015725	-0,0014206

Podemos ahora calcular la corrección y su incertidumbre estándar combinada en una medición cualquiera. Supongamos que hemos medido una presión sonora  $P_o = 8$  Pa con una incertidumbre estándar de 0,35 Pa. La presión sonora corregida será:

$$P_c = 8 \times 1 - 0,0050166 \times 8^2 + 0,00084634 \times 8^3 = 8,1123 \text{ Pa}$$

Para calcular la incertidumbre utilizaremos la ecuación (2.77) conjuntamente con (2.78) y (2.79). La (2.78) toma la forma

$$\frac{\partial P_c}{\partial P_o} = a + 2bP_o + 3cP_o^2 = 1,0822.$$

Las (2.79),

$$\frac{\partial P_c}{\partial P_{1o}} = \frac{\partial a}{\partial P_{1o}} P_o + \frac{\partial b}{\partial P_{1o}} P_o^2 + \frac{\partial c}{\partial P_{1o}} P_o^3 = 109,97$$

$$\frac{\partial P_c}{\partial P_{2o}} = \frac{\partial a}{\partial P_{2o}} P_o + \frac{\partial b}{\partial P_{2o}} P_o^2 + \frac{\partial c}{\partial P_{2o}} P_o^3 = -13,574$$

$$\frac{\partial P_c}{\partial P_{3o}} = \frac{\partial a}{\partial P_{3o}} P_o + \frac{\partial b}{\partial P_{3o}} P_o^2 + \frac{\partial c}{\partial P_{3o}} P_o^3 = -0,6278$$

Las incertidumbres que aparecen en (2.77) se pueden determinar a partir de lo datos. Veamos la relación entre la incertidumbre en dB y la incertidumbre en la presión. Dado que

$$L_p = 20 \log \frac{P}{P_{\text{ref}}} = \frac{20}{\ln 10} \ln \frac{P}{P_{\text{ref}}} \quad (\text{A9.15})$$

Entonces

$$u_{L_p} = \frac{20}{\ln 10} \frac{1}{P} u_P \quad (\text{A9.16})$$

o bien

$$u_P = \frac{\ln 10}{20} P u_{L_p} \quad (\text{A9.17})$$

La incertidumbre en el valor leído de cada referencia tiene dos componentes, la del valor de la propia referencia (0,2 dB) y la de la medición (0,4 dB):

$$u_{P_{io}}^2 = \left( \frac{\ln 10}{20} P_1 u_{L_{p1}} \right)^2 + \left( \frac{\ln 10}{20} P_{1o} u_{L_{p1o}} \right)^2 \quad (\text{A9.18})$$

de modo que

$$\begin{aligned}
u_{P_{10}} &= 0,005163 \text{ Pa} \\
u_{P_{20}} &= 0,051783 \text{ Pa} \\
u_{P_{30}} &= 0,50364 \text{ Pa}
\end{aligned}$$

La incertidumbre en el valor leído del mensurando es sólo la correspondiente a las mediciones (0,4 dB)

$$u_{P_o} = \frac{\ln 10}{20} P_o u_{L_{po}}, \quad (\text{A9.19})$$

es decir

$$u_{P_o} = 0,36841. \quad (\text{A9.20})$$

Finalmente, aplicamos la ecuación (2.77), que en este caso adopta la forma:

$$u_{P_c}^2 = \left( \frac{\partial P_c}{\partial P_o} \right)^2 u_{P_o}^2 + \left( \frac{\partial P_c}{\partial P_{10}} \right)^2 u_{P_{10}}^2 + \left( \frac{\partial P_c}{\partial P_{20}} \right)^2 u_{P_{20}}^2 + \left( \frac{\partial P_c}{\partial P_{30}} \right)^2 u_{P_{30}}^2. \quad (\text{A9.21})$$

Resulta

$$u_{P_c} = 1,037 \text{ Pa}.$$

Es interesante verificar que la incertidumbre en el valor corregido es mayor (casi el triple) que la incertidumbre de la medición original (A9.20) sin corregir. Esto se debe a que en la búsqueda de los parámetros de un modelo como el (A9.3) incorporamos nuevas componentes de incertidumbre. Si la alinealidad no es muy marcada, no conviene realizar la corrección.

## Apéndice 10

### Teorema del muestreo

Supongamos que queremos muestrear una señal  $v(t)$  que varía en forma arbitraria con el tiempo con la sola restricción de algún límite en el ancho de banda ocupado. Un ejemplo de ello son las señales de audio, cuya máxima frecuencia útil está en el orden de los 20 kHz.

Una consideración central en estos casos es la elección de la tasa de muestreo, es decir la cantidad de muestras a tomar por unidad de tiempo. Veremos que si se desea posteriormente reconstruir la señal o procesarla en forma asimilable a un procesador analógico (amplificación, filtrado, compresión, etc.) será preciso muestrearla a una tasa mayor o igual que el doble de la máxima frecuencia presente en la señal.

Para ello consideremos la transformada de Fourier  $V(\omega)$  de  $v(t)$ :

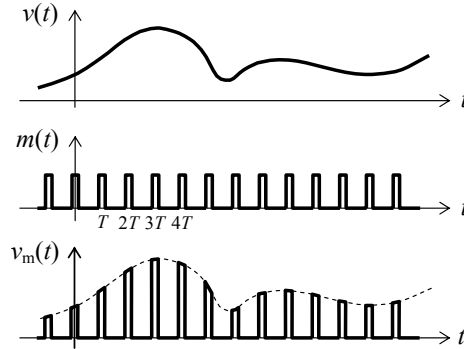
$$V(\omega) = \int_{-\infty}^{+\infty} v(t) e^{-j\omega t} dt. \quad (\text{A10.1})$$

Se cumple que

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} V(\omega) e^{j\omega t} dt. \quad (\text{A10.2})$$

Supongamos ahora que multiplicamos dicha señal por una señal de muestreo periódica  $m(t)$  dada por (ver figura A10.1)

$$m(t) = \begin{cases} 1/\tau & \text{si } -\tau/2 < t < \tau/2 \\ 0 & \text{si } \tau/2 < t < T - \tau/2 \\ m(t-T) & \text{para todo otro } t \end{cases} \quad (\text{A10.3})$$



**Figura A10.1.** Muestreo de una señal variable en el tiempo por multiplicación por un tren de pulsos de periodo  $T$  y ancho  $\tau$ .

El resultado es la señal muestreada

$$v_m(t) = v(t) \cdot m(t). \quad (\text{A10.4})$$

Aunque las señales pueden ser descriptas tanto desde el punto de vista de su evolución temporal como del espectro, frecuentemente la descripción espectral nos proporciona una idea más clara del efecto de un determinado proceso sobre la señal. Es precisamente el caso del proceso de muestreo, y por ello nos proponemos calcular el espectro de  $v_m(t)$ . Calculemos primero el espectro de la señal de muestreo  $m(t)$ . Por ser un tren de pulsos periódico, admite un desarrollo en serie de Fourier. Conviene la representación en la forma compleja:

$$m(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_m t} \quad (\text{A10.5})$$

donde  $\omega_m = 2\pi f_m = 2\pi/T$ , y

$$c_n = \frac{1}{T} \int_{-T/2}^{+T/2} m(t) e^{-jn\omega_m t} dt = \frac{1}{T} \int_{-T/2}^{+T/2} \frac{1}{\tau} e^{-jn\omega_m t} dt.$$

es decir,

$$c_n = \frac{1}{T} \frac{\text{sen } n\omega_m \tau / 2}{n\omega_m \tau / 2}. \quad (\text{A10.6})$$

Estos coeficientes siguen una evolución con  $n$  según la función  $\text{sinc}(x) = \text{sen } x / x$ . Para  $\tau \rightarrow 0$ , los coeficientes tienden a ser todos iguales a  $1/T$ . Multiplicando la expresión anterior de  $m(t)$  por  $v(t)$  se obtiene

$$v_m(t) = \sum_{n=-\infty}^{\infty} c_n v(t) e^{jn\omega_m t}, \quad (\text{A10.7})$$

de donde su transformada de Fourier resulta

$$\begin{aligned} V_m(\omega) &= \int_{-\infty}^{+\infty} \left( \sum_{n=-\infty}^{\infty} c_n v(t) e^{jn\omega_m t} \right) e^{-jn\omega t} dt = \\ &= \sum_{n=-\infty}^{\infty} c_n \int_{-\infty}^{\infty} v(t) e^{-j(\omega - n\omega_m)t} dt = \\ &= \sum_{n=-\infty}^{\infty} c_n V(\omega - n\omega_m) \end{aligned} \quad (\text{A10.8})$$

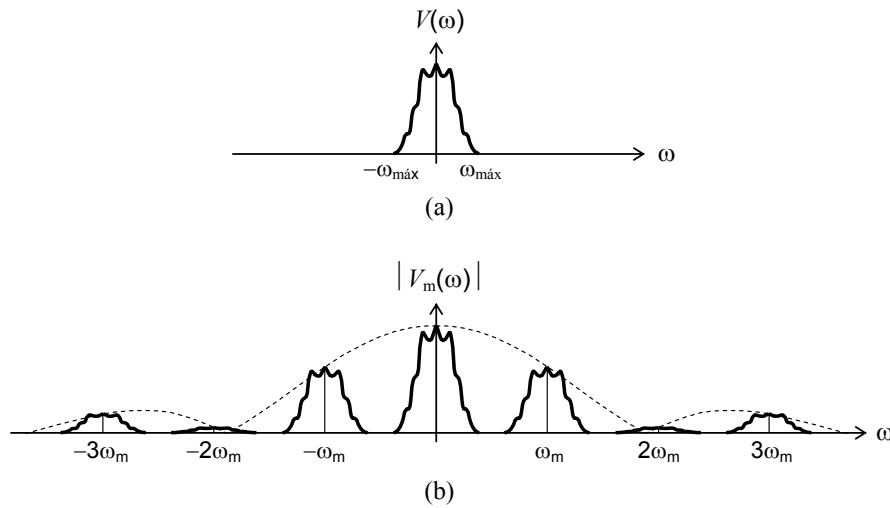


Este resultado es muy interesante, ya que revela que el espectro de la señal muestreada  $v_m(t)$  está constituido por infinitud de “copias” del espectro bilateral de la señal original distribuidas periódicamente en el eje de frecuencias con “periodo” igual a la frecuencia de muestreo. Las “copias” están ponderadas por los coeficientes  $c_n$  del espectro de la señal de muestreo (ver figura A10.2). Para el caso ideal en que  $\tau \rightarrow 0$ , dichas copias tienen todas la misma amplitud.

En el ejemplo de la figura A10.2 el espectro de la señal se encuentra limitado a una banda  $\pm \omega_{\text{máx}}$ , habiéndose elegido la tasa (angular) de muestreo  $\omega_m$  de manera que

$$\omega_m > 2 \omega_{\text{máx}}. \quad (\text{A10.9})$$

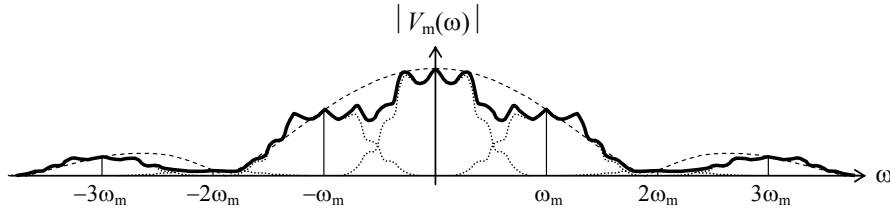
Si se cumple esta condición, denominada *condición de Nyquist*, las “copias” del espectro original se encontrarán separadas, y entonces será cuestión relativamente sencilla recuperar la señal original mediante un filtro pasabajos suficientemente selectivo, vale decir, que sea capaz de eliminar las copias centradas en  $\pm \omega_m$ ,  $\pm 2\omega_m$ ,  $\pm 3\omega_m$ , etc. Este resultado se conoce como el *Teorema del Muestreo*.



**Figura A10.2.** (a) Espectro de una señal, limitado a la banda de frecuencias  $[-\omega_{\text{máx}}, \omega_{\text{máx}}]$ . (b) Espectro de la señal muestreada con frecuencia de muestreo  $\omega_m > 2\omega_{\text{máx}}$ .

Si, en cambio, no se cumpliera la condición de Nyquist, como en el ejemplo de la figura A10.3, las copias se superpondrían entre sí, y ya no sería posible recuperar la banda base (es decir, la correspondiente al espectro de la señal original), ya que sobre ella aparecerían residuos de las otras bandas. Si se pretendiera recuperar la señal por medio de un filtro pasabajos, aparecerían componentes de frecuencia correspondientes al extremo del espectro de la señal “reflejadas” sobre la frecuencia  $\omega_m/2$ . Por ejemplo, una frecuencia  $\omega_m/2 + \Delta\omega$  se reflejaría como la frecuencia  $\omega_m/2 - \Delta\omega$ . Estas frecuencias se denominan *frecuencias alias* de las correspondientes frecuencias originales, y el fenómeno se conoce con el término inglés *aliasing*. No existe ningún procedimiento que

permita eliminar estas frecuencias alias sin eliminar simultáneamente otras frecuencias que podrían haber sido muestreadas genuinamente. En el ejemplo anterior, si elimináramos la frecuencia alias  $\omega_m/2 - \Delta\omega$ , también se eliminarían aquellas partes de la señal original que tuvieran precisamente esa frecuencia.



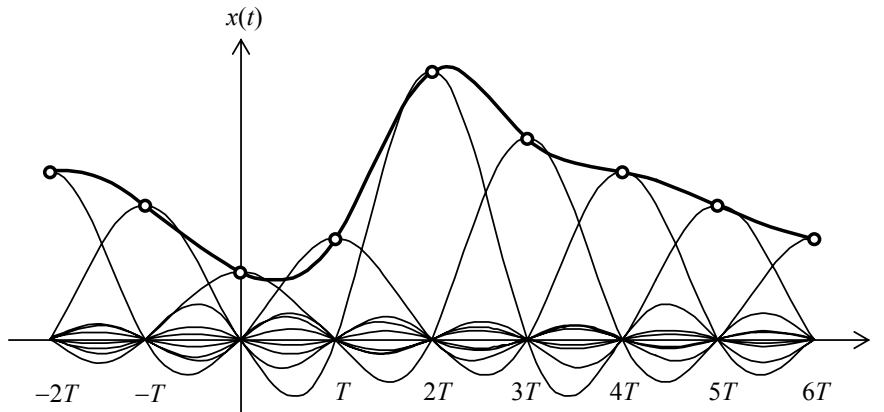
**Figura A10.3.** Espectro de una señal muestreada con frecuencia de muestreo  $\omega_m < 2\omega_{\text{máx.}}$ .

La frecuencia  $f_m/2$  se denomina *frecuencia de Nyquist*, y el teorema del muestreo puede plantearse alternativamente de esta manera: sólo podrán recuperarse perfectamente las señales muestreadas a una frecuencia de muestreo  $f_m$  si su espectro no excede la frecuencia de Nyquist.

Es interesante señalar que si la tasa de muestreo cumple con el teorema del muestreo existe una fórmula de reconstrucción perfecta de una señal muestreada, dada por

$$x(t) = \sum_{k=-\infty}^{\infty} x(kT) \frac{\text{sen } \pi f_m t}{\pi f_m t}. \quad (\text{A10.10})$$

En la figura A10.4 se muestra la reconstrucción basada en esta fórmula, denominada también *interpolación sinc* porque utiliza la función sinc para interpolar entre los instantes de las muestras.

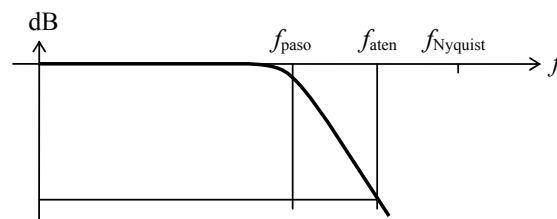


**Figura A10.4.** Interpolación sinc para reconstruir una señal muestreada en cumplimiento del Teorema del muestreo.

La interpolación sinc no es muy eficiente, ya que si bien la serie de la ecuación (A10.10) es convergente, converge muy lentamente, por lo que se necesita gran cantidad de términos para lograr un error aceptable. En algunos casos se utiliza una sinc truncada, es decir, en lugar de calcular cada sinc en toda la duración de la señal, se utilizan sólo valores correspondientes a unos cuantos ciclos antes y después de su valor central. Se suele utilizar en osciloscopios digitales para presentar el oscilograma de la señal en la pantalla. No tiene, en cambio, aplicaciones serias en audio digital, donde se suele utilizar filtros de reconstrucción mucho más eficientes computacionalmente.

Volviendo a la condición de Nyquist, obsérvese que no es suficiente con que las frecuencias útiles no vayan más allá de la frecuencia de Nyquist, ya que el fenómeno de *aliasing* no distingue las señales útiles de las que no lo son. Así, por ejemplo, las señales de audio no contienen frecuencias útiles por encima de los 20 kHz, por lo cual, idealmente, debería ser posible muestrearlas (sin mayor acondicionamiento previo) a una tasa de muestreo de 40 kHz. Ello no es así, sin embargo, ya que dichas señales contienen ruido de frecuencias superiores a 20 kHz (en algunos casos hasta de varios cientos de kHz). Dicho ruido es inaudible, pero al muestrearlo a una tasa de muestreo insuficiente se crearán frecuencias alias dentro de la banda útil que sí son audibles. Un tono de 38 kHz proveniente de la subportadora de una señal de FM estereofónica, por ejemplo, excede en 18 kHz la frecuencia de Nyquist correspondiente a este caso; por consiguiente, se reflejará como un alias de  $20 \text{ kHz} - 18 \text{ kHz}$ , es decir, un tono de 2 kHz, que es absolutamente audible.

La solución al problema planteado consiste en intercalar antes del muestreo un filtro pasabajos denominado *filtro antialias*, que reduce el ruido fuera de la banda útil a niveles despreciables (ver figura A10.5). Dado que todo filtro real tiene una transición gradual entre la frecuencia de paso y la de atenuación (ver figura A10.4), la frecuencia de muestreo deberá ser bastante mayor que el doble de la máxima frecuencia útil. Por ejemplo, en audio digital se toma como mínimo  $f_m = 44,1 \text{ kHz}$  en lugar de 40 kHz. La frecuencia de Nyquist vale, en este caso, 22,05 kHz, lo cual permite un margen de unos 2 kHz para la transición del filtro. Esta transición es muy rápida, lo cual obliga usar filtros de elevado orden, que producen transitorios indeseables. Por eso es que cuando no hay restricciones de memoria, se prefiere utilizar  $f_m = 48 \text{ kHz}$  o incluso  $f_m = 96 \text{ kHz}$ , lo cual permite usar filtros más simples con menos transitorios.



**Figura A10.5.** Respuesta en frecuencia de un filtro pasabajos anti-alias. La frecuencia de atenuación debe ser menor que la frecuencia de Nyquist.

La atenuación del filtro antialias por encima de la frecuencia de Nyquist debe ser suficiente para garantizar que el ruido creado por las frecuencias alias sea aceptable. En

el caso típico en que las muestras son ulteriormente digitalizadas, bastaría con que dicho ruido resulte menor que 0,5 LSB, ya que en ese caso no podrá ser discriminado por el conversor analógico digital.

Como comentario final, el teorema del muestreo puede generalizarse para señales de tipo pasabanda. En este caso se puede demostrar (Proakis et al., 1998) que si el espectro de una señal está concentrado en el rango  $f_{\min} < f < f_{\max}$  y, por lo tanto, su ancho de banda es

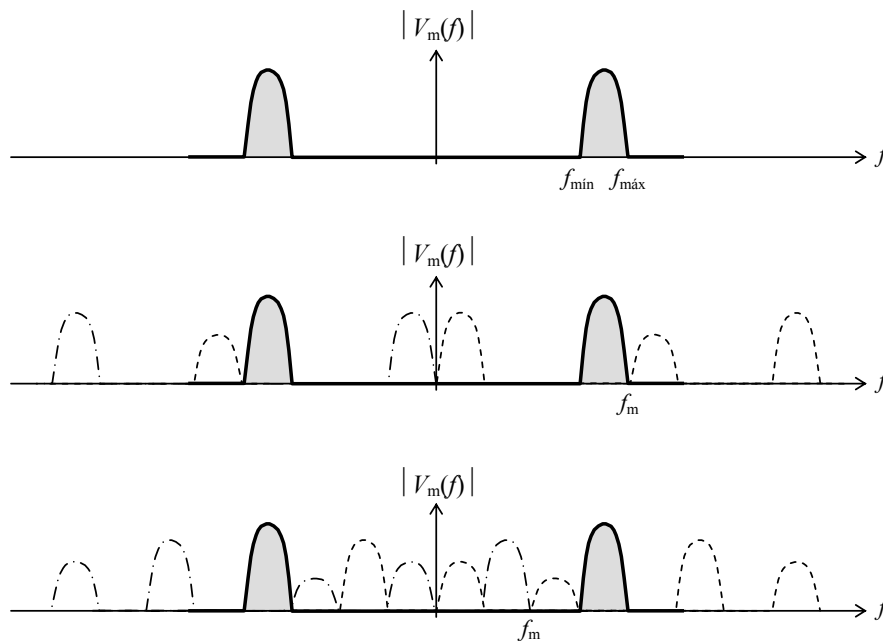
$$B = f_{\max} - f_{\min}, \quad (\text{A10.11})$$

entonces puede muestrearse satisfactoriamente (es decir sin que aparezcan alias) con una tasa de muestreo  $f_m$  tal que

$$2B \leq f_m < 4B, \quad (\text{A10.12})$$

aunque no se cumpla que  $f_{\max} > f_m/2$ , ya que las frecuencias alias caerán siempre fuera del rango  $f_{\min} < f < f_{\max}$  y, por consiguiente, podrán eliminarse por filtrado.

Si, por ejemplo, una señal está comprendida entre 15 kHz y 20 kHz, es decir que  $B = 5$  kHz, será válido muestrearla con cualquier tasa de muestreo entre 10 kHz y 20 kHz, a pesar de estar por debajo del doble de la máxima frecuencia, es decir, por debajo de 40 kHz (e incluso por debajo de la propia frecuencia máxima).



**Figura A10.6.** Teorema del muestreo para señales pasabanda. Arriba, espectro bilateral de una señal pasabanda de ancho de banda  $B$ . Centro, espectro de la señal muestreada a una tasa  $f_m = 4B$ . Abajo: espectro de la señal muestreada a una tasa  $f_m = 2B$ .

## Apéndice 11

### Estructura de un archivo FLAC

#### A11.1 Archivo y flujo de datos

Si bien aquí nos referimos principalmente a *archivos* FLAC, es interesante destacar que el formato es igualmente aplicable a un *flujo de datos* (*stream*). La diferencia entre un archivo y un flujo de datos es que el archivo es estático, es decir, datos asociados físicamente con un almacenamiento, mientras que un flujo de datos está asociado a una transmisión, en general dentro de una red (por ejemplo Internet o una intranet)

De hecho, el formato FLAC está pensado para permitir la transmisión en tiempo real (es decir a una tasa igual o superior a la que requiere la reproducción) a través de redes. El formato es, incluso, tolerante a la pérdida de cierta información general, ya que permite la decodificación aun a partir de un punto intermedio (Raynaudo et al., 2013).

#### A11.2 Estructura de datos

La estructura de datos de audio de un archivo FLAC (Coalson, 2008; Roveri, 2011) se muestra en la Figura A11.1. Está compuesta por cuadros (*frames*) y subcuadros (*subframes*), que codifican respectivamente los bloques (*blocks*) y subbloques (*subblocks*) de información PCM de audio. Un bloque contiene una cantidad  $N$  de muestras multicanal (en un caso típico, muestras estereofónicas) y se divide en subbloques, cada uno de los cuales contiene las  $N$  muestras de un canal. En el codificador de referencia,<sup>127</sup> el tamaño  $N$  de cada bloque se determina a partir de la tasa de muestreo. Por ejemplo, para  $F_s = 44100$  Hz,  $N \leq 4608 = 2^9 \times 3^2$ .

Cada cuadro contiene un encabezamiento (*header*) con metadatos y una serie de subcuadros. Cada subcuadro tiene su propio encabezamiento y la información codificada correspondiente a un subbloque. A diferencia de lo que ocurre en un archivo WAV, donde los canales se encuentran entrelazados, aquí los subcuadros aparecen en sucesión. Esto simplifica la decodificación.

El archivo comienza con el identificador “fLaC” (es decir, 0x6C4C6143 en código ASCII hexadecimal) seguido por uno o más bloques de metadatos, seguido por uno o más cuadros.

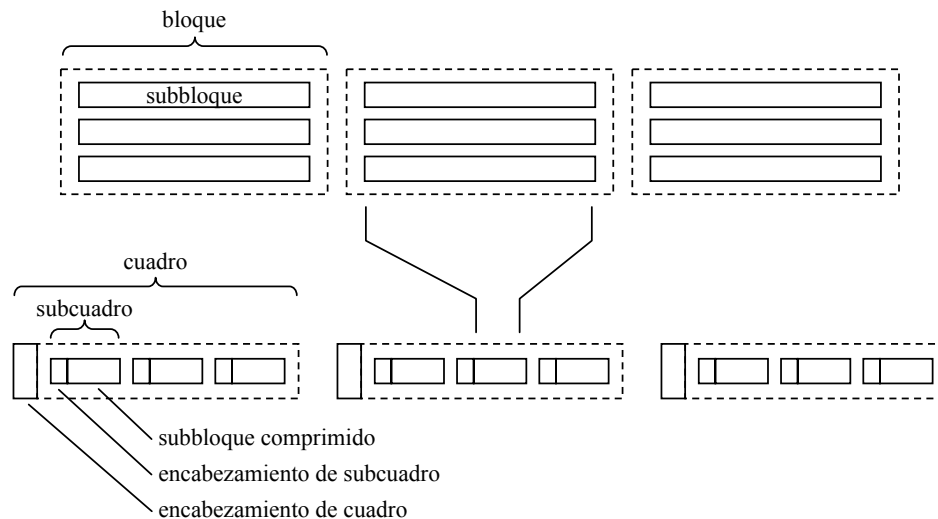
Todos los bloques de metadatos tienen un encabezamiento formado por 4 bytes (figura A11.2). El primer bit del primer byte es 1 si es el último bloque de metadatos y 0 en caso contrario. Los restantes 7 bits contienen un código entre 0 y 127 que indica el tipo de metadatos. En la tabla A11.1 se muestran los códigos de los bloques normalizados. Los tres bytes que siguen contienen la longitud del resto del bloque de metadatos en formato *big endian* (es decir, el byte más significativo está primero, como en la escritura normal).

El primer bloque de metadatos (obligatorio), se denomina *streaminfo* y se refiere a la información sobre el flujo de datos (ver la tabla A11.2). Se identifica con el número 0.

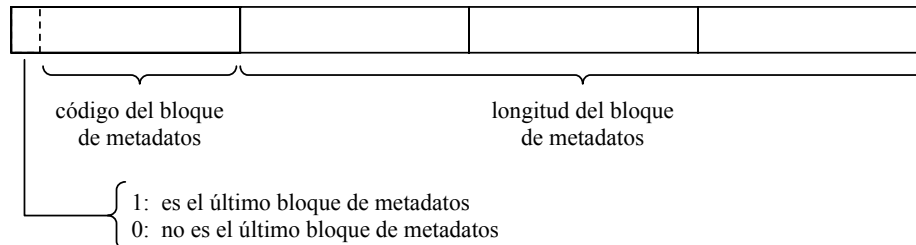
Según se puede apreciar, los campos no necesariamente tienen una cantidad entera de bytes, aunque el bloque completo tiene 34 bytes. Los dos primeros campos son especificaciones definidas a priori por el codificador. Los dos siguientes indican los tamaños mínimo y máximo de los cuadros codificados a lo largo de todo el archivo o flujo de

<sup>127</sup> El proyecto FLAC (<http://flac.sourceforge.net>) proporciona una especificación de formato y un codec de referencia, si bien por tratarse de un formato libre y abierto cualquier desarrollador puede crear su propio codec.

datos. Los valores dependen de la redundancia de los diferentes bloques de señal. Así, los bloques más redundantes tendrán menor longitud de cuadro que los más aleatorios. Los tres siguientes proporcionan las características de la señal y su muestreo: Tasa de muestreo, cantidad de canales y cantidad de bits por muestra. El octavo campo registra el número de muestras de todo el archivo o flujo de datos.



**Figura A11.1.** Ejemplo de estructura de un archivo de sonido con formato FLAC. Arriba, la señal PCM correspondiente a tres canales de audio antes de comprimir. Abajo, la señal comprimida con el agregado de los encabezamientos de cuadro y subcuadro.



**Figura A11.2.** Encabezamiento de cada bloque de metadatos.

Los últimos 16 bytes (128 bit) del encabezamiento de cuadro constituyen una firma digital<sup>128</sup> MD5 calculada sobre los datos originales antes de su compresión. Es el resultado de la aplicación de un algoritmo tipo *hash*<sup>129</sup> y su objetivo es permitir al decodificador verificar la integridad de los datos de audio recuperados. Para ello se vuelve a

<sup>128</sup> Véase, por ejemplo, <http://en.wikipedia.org/wiki/MD5>, <http://tools.ietf.org/html/rfc1321>

<sup>129</sup> Un algoritmo *hash* permite obtener un número de longitud fija (*código hash*) a partir de una cantidad arbitraria de bits. El más simple posible es el que obtiene un bit de paridad, es decir, un bit que es 1 si la cantidad de bits con valor 1 es par y 0 en caso contrario. Se utilizan como código de detección de errores y también en criptografía.

aplicar el algoritmo luego de la decodificación y se compara el resultado con la firma que figura en el encabezamiento. Si son iguales, se da por válida.

**Tabla A11.1.** Identificadores numéricos de los bloques de metadatos.

N°	Bloque
0	STREAMINFO
1	RELLENO (Padding)
2	APLICACIÓN
3	TABLA DE BÚSQUEDA (Seek Table)
4	COMENTARIO EN ESTILO VORBIS
5	HOJA DE MARCAS (Cue Sheet)
6	IMAGEN (Picture)
7 a 126	Reservado para futuros metadatos
127	Inválido

**Tabla A11.2.** Estructura de datos del bloque de metadatos *streaminfo*.

Bits	Descripción
16	Tamaño mínimo de los bloques (en muestras) <sup>1</sup>
16	Tamaño máximo de los bloques (en muestras) <sup>2</sup>
24	Tamaño mínimo de los cuadros (en bytes) <sup>3</sup>
24	Tamaño máximo de los cuadros (en bytes) <sup>3</sup>
20	Tasa de muestreo (en Hz) <sup>4</sup>
3	Cantidad de canales menos 1
5	Cantidad de bits por muestra <sup>5</sup> menos 1
36	Cantidad total de muestras <sup>6</sup>
128	Firma digital MD5 de los datos de audio sin comprimir

<sup>1</sup> El mínimo valor admisible es 16

<sup>2</sup> El máximo valor admisible es 65536

<sup>3</sup> Un valor 0 indica que es desconocido

<sup>4</sup> El valor no debe ser mayor que 655 350 Hz

<sup>5</sup> El codec de referencia permite hasta 24 bit

<sup>6</sup> Una muestra multicanal se considera una sola muestra

Aunque se han detectado casos donde dos conjuntos diferentes de datos dan origen a la misma firma digital (situación conocida como *colisión*) son casos extremadamente raros, por lo cual es muy improbable dar por válido un resultado corrupto.

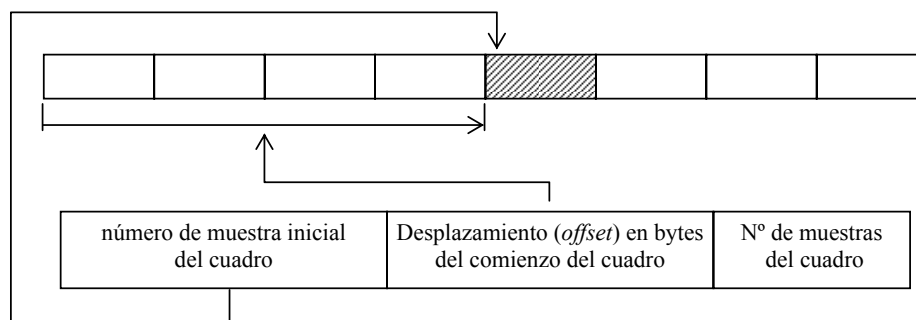
La validación por MD5 es conveniente pero no es una función imprescindible del decodificador, sobre todo en la manipulación de archivos locales mediante un hardware altamente confiable.

El segundo bloque de metadatos (opcional), denominado *relleno (padding)*, se prevé para reservar espacio en un archivo para uso futuro sin tener que modificar el tamaño del mismo. Consiste simplemente en una cantidad entera de bytes nulos. El uso podría ser agregar ulteriormente un bloque de metadatos no previsto durante la codificación, o para cualquier otra información particular, aunque no tenga un formato estándar. Este bloque podría ser útil en aplicaciones metrológicas, para almacenar información

contextual, por ejemplo datos GPS, fecha, descripción, observaciones, operador técnico, etc. Podría parecer que el reservar un espacio ocioso que tal vez nunca se llegue a usar va contra la filosofía de la compresión lo cual de alguna manera es cierto. Sin embargo, la cantidad de espacio requerido en general es despreciable frente a los datos de audio, incluso después de la compresión.

El tercer bloque (opcional), *aplicación*, involucra información sobre el software que originó el archivo, o bien del codec de terceros que dicho software utiliza, por ejemplo el codec de referencia del proyecto FLAC. La estructura y la longitud de la información la establece el fabricante del software. Los cuatro primeros bytes contienen un número de registro del software, que se gestiona en el sitio web del proyecto FLAC.

El cuarto bloque (opcional), *tabla de búsqueda* (*seek table*), resuelve el problema de la búsqueda eficiente de muestras específicas en un contexto en el que cada muestra tiene una duración diferente. La búsqueda directa es muy lenta, dado que requiere básicamente contar las muestras una por una. En lugar de ello se incorpora una serie de *puntos de búsqueda* (*seek points*) que permiten localizar cuadros específicos. La tabla de búsqueda contiene 18 bytes por cada punto de búsqueda. Los primeros 8 bytes indican el número correlativo de muestra de la primera muestra que aparece en el cuadro objeto de la búsqueda. Los siguientes 8 bytes indican el desplazamiento (*offset*) en bytes entre el primer byte del encabezamiento del primer cuadro y el primer byte del encabezamiento del cuadro objeto de la búsqueda. Los últimos dos bytes indican la cantidad de muestras de dicho cuadro. Dado que un cuadro contiene relativamente pocas muestras, la búsqueda final dentro del cuadro es rápida. La figura A11.3 muestra el mecanismo de direccionamiento del punto de búsqueda. Si, por ejemplo, buscamos la  $m$ -ésima muestra, bastará encontrar el punto de búsqueda más cercano. Ese punto corresponde al comienzo de algún cuadro, del cual por el segundo argumento sabemos en qué byte comienza. A partir de allí se inicia una búsqueda secuencial, pero partiendo de una posición mucho más cercana.



**Figura A11.3.** Arriba, una serie de cuadros. Abajo, información asociada a un punto de búsqueda.

También es posible reservar espacio para insertar un futuro punto de búsqueda. Ello se logra colocando, al final, uno o más puntos de búsqueda comodines, caracterizados por un valor 0xFFFFFFFFFFFFFFFF como número de muestra inicial.

El quinto bloque (opcional), *comentario en estilo Vorbis*, incluye una serie de campos con una estructura débil y no completamente normalizados. El estilo de comen-



tarios Vorbis es el que se utiliza en el formato comprimido Vorbis-OGG. La estructura de este bloque de metadatos se indica en la tabla A11.3.

**Tabla A11.3.** Estructura de datos del bloque de metadatos de *comentarios en estilo Vorbis*.

Bytes	Tipo	Contenido
4	Little endian	Longitud $N$ de una cadena de caracteres provista por el fabricante del codec
$N$	ASCII / UTF8 <sup>130</sup>	Cadena de $N$ caracteres provista por el fabricante del codec
4	Little endian	Cantidad $L$ de comentarios del usuario
4	Little endian	Longitud $M_1$ del comentario 1 del usuario
$M_1$	ASCII / UTF8	Comentario 1 del usuario ( $M_1$ caracteres)
...	...	...
4	Little endian	Longitud $M_L$ del comentario $L$ del usuario
$M_L$	ASCII / UTF8	Comentario $L$ del usuario ( $M_L$ caracteres)

La cadena provista por el fabricante del códec contiene en general información sobre el nombre del códec, versión y fecha. Por ejemplo, para el códec de referencia, corresponde a los  $N = 32$  caracteres que forman la cadena

“reference libFLAC 1.2.1 20070917”

Los comentarios del usuario deben tener una mínima estructura del tipo:

[Nombre del campo]=[Comentario]

donde [Nombre del campo] es una cadena ASCII insensible a las mayúsculas (*case insensitive*) que representa una descripción en inglés del contenido del campo, y [Comentario] es una cadena UTF-8 en cualquier idioma compatible, que representa el comentario en sí. En la tabla A11.4 se dan ejemplos de nombres de campo frecuentemente utilizados.

Estos campos están evidentemente pensados para la industria discográfica musical. Si bien en la especificación de Vorbis<sup>131</sup> se incluye una descripción y ejemplos, la misma no es normativa y no se establece ninguna característica obligatoria del contenido de los campos. De hecho se podrían añadir comentarios ad-hoc de mayor interés para la documentación de un registro con fines metrológicos. Para dar un ejemplo, podríamos tener un comentario como:

TITLE=Paisaje sonoro matinal de Plaza López

que iría precedido por el número 0x2B000000, que en formato little endian representa al decimal 43, es decir, la cantidad de bytes que utiliza el comentario (incluido el nombre del campo).

<sup>130</sup> A diferencia de ASCII, UTF-8 es un código para representar caracteres que permite utilizar los valores por encima de 127, incluso más de un byte, ofreciendo soporte para diversos idiomas.

<sup>131</sup> [http://xiph.org/vorbis/doc/Vorbis\\_I\\_spec.html](http://xiph.org/vorbis/doc/Vorbis_I_spec.html)

**Tabla A11.4.** Nombres de campo usuales.

TITLE	ORGANIZATION
VERSION	DESCRIPTION
ALBUM	GENRE
TRACKNUMBER	DATE
ARTIST	LOCATION
PERFORMER	CONTACT
COPYRIGHT	ISRC
LICENSE	

El siguiente bloque de metadatos (opcional) es la *hoja de marcas* (*cue sheet*), también llamada *hoja cue*. En principio está pensado para el caso en que el archivo flac se utiliza para comprimir el contenido de un disco compacto de audio (CD-DA), aunque admite casos más generales. Contiene información correspondiente a la ubicación de las *pistas* (*tracks*) y, dentro de éstas, los *puntos índice*. Las pistas son grandes divisiones correspondientes en general a piezas de música individuales, mientras que los puntos índice permiten ubicar secciones dentro de una pieza extensa. La hoja de marcas tiene una estructura como se muestra en las tablas A11.5 a A11.7.

**Tabla A11.5.** Estructura de datos de la hoja de marcas.

Longitud	Tipo	Contenido
128 byte	ASCII	Número de catálogo de medios (MCN)
8 byte	Big-endian hex	Número de muestras iniciales del CD
1 bit	bit	1 para CD, 0 en caso contrario
7 bit	bit	Reservado. Todos los bits son 0
258 byte	bit	Reservado. Todos los bits son 0
1 byte	hex	Cantidad $N$ de pistas ( <i>tracks</i> )
Variable		Pista 1 de la hoja de marcas
...	...	...
Variable		Pista $N$ de la hoja de marcas

**Tabla A11.6.** Estructura de datos de una pista de la hoja de marcas.

Longitud	Tipo	Contenido
8 byte	Big-endian hex	Desplazamiento en muestras del primer punto índice de la pista relativo al comienzo del flujo de audio
1 byte	hex	Número $n$ de pista
8 byte	ASCII	International Standard Recording Code
1 bit	bit	Tipo de pista. 0 para audio, 1 para el resto
1 bit	bit	0: con re-énfasis; 1: sin preénfasis
6 bit	bit	Reservado. Todos los bits son 0
13 byte	bit	Reservado. Todos los bits son 0
1 byte	hex	Cantidad $M$ de puntos índice en la pista
12 byte		Punto índice 1
...	...	...
12 byte		Punto índice $M$

NOTAS: 1) El número de catálogo de medios para un CD es un número de 13 cifras. Los restantes 115 bytes deben ser nulos. 2) Las muestras iniciales corresponden a la pista 0, que contiene al menos 2 s de silencio. 3) El International Standard Recording Code es un número de 12 dígitos asignado a cada pista por la Federación Internacional de la Industria Fonográfica (International Federation of the Phonographic Industry).

**Tabla A11.7.** Estructura de datos de un punto índice de pista.

Longitud	Tipo	Contenido
8 byte	Big-endian hex	Offset en muestras del punto índice, relativo al offset de la pista (para un CD es divisible por 588)
1 byte	hex	Número $L$ de punto índice
3 byte	bit	Reservado. Todos los bits son 0

El último bloque de metadatos corresponde a imágenes, entre ellas íconos, cubierta, fotografías de los autores e intérpretes, de instancias durante la grabación, etc. Por ser demasiado específico y de poca utilidad para nuestros propósitos, no profundizaremos su descripción.

Una vez finalizados los bloques de metadatos, comienzan a aparecer los cuadros y subcuadros que contienen encabezamientos y datos de audio comprimidos.

El encabezamiento de un cuadro se inicia con un código de sincronismo, que es invariablemente “111111111110”, lo cual permite comenzar a decodificar el audio de un archivo FLAC incluso en el curso de un flujo de datos del cual no se conoce el comienzo.<sup>132</sup> Al código de sincronismo le siguen dos bits reservados. El primero es 0 y el segundo vale 0 para un flujo de datos cuyo tamaño de bloque es fijo (la estrategia adoptada por el códec de referencia) y 1 para tamaño de bloque variable.

Luego siguen 4 bit que indican el tamaño  $B$  del bloque en muestras. La idea es usar pocos bits para los tamaños más frecuentes, pero permitir otros tamaños más generales, en cuyo caso se remite al final del encabezamiento del cuadro (tabla A11.8). Para una tasa de muestreo de 44,1 kHz el códec de referencia usa 0101, es decir,  $B = 4608$ .

**Tabla A11.8.** Códigos del tamaño del bloque en el encabezamiento de un cuadro.

Código	Tamaño $B$ del bloque en muestras
0000	Reservado
0001	192
$0010 \leq n \leq 0101$	$576 \times 2^{n-2}$
0110	Leer 8 bits al final del encabezamiento y sumar 1 ( $1 \leq B \leq 256$ )
0111	Leer 16 bits al final del encabezamiento y sumar 1 ( $1 \leq B \leq 65536$ )
$1000 \leq n \leq 1111$	$256 \times 2^{n-8}$

<sup>132</sup> Esta combinación de bits podría aparecer también como sucesión de códigos de Golomb-Rice, por ejemplo, para un divisor 4, el final de un número impar, luego cuatro veces el 3 seguido por un número mayor de 4 presentan el mismo patrón

Los 4 bits siguientes indican la tasa de muestreo. Si bien ésta es constante para todo el archivo y ya había sido declarada en el bloque de metadatos *streaminfo* al principio del archivo (ver tabla A11.2), la información se reitera en cada cuadro para permitir la decodificación de un flujo de datos aun si se han perdido los metadatos. Pero dado que esta información se repite en todos los cuadros, se utilizan pocos bits para las tasas de muestreo más habituales. En la tabla A11.9 se muestra la interpretación de cada código.

**Tabla A11.9.** Códigos de la tasa de muestreo en el encabezamiento de un cuadro

Código	Tasa de muestreo $F_s$ [kHz]
0000	Se obtiene del bloque <i>streaminfo</i>
0001	88,2
0010	176,4
0011	192
0100	8
0101	16
0110	22,05
0111	24
1000	32
1001	44,1
1010	48
1011	96
1100	Valor de 8 bit en kHz de $F_s$ al final del encabezamiento
1101	Valor de 16 bit en Hz de $F_s$ al final del encabezamiento
1110	Valor de 16 bit en decenas de Hz de $F_s$ al final del encabezamiento
1111	Inválido

Los próximos 4 bits representan la asignación de canales según se muestra en la tabla A11.10. Entre 0000 y 0111 corresponde a la cantidad de canales  $C$  menos 1. Los tres valores siguientes indican, para el caso estéreo, la forma en que se ha realizado la decorrelación entre canales, información necesaria para reconstruir los canales originales.

**Tabla A11.10.** Códigos de asignación de canales en el encabezamiento de un cuadro.

Código	Contenido	
0000 - 0111	Cantidad de canales menos 1	
1000	Canal 0: izquierdo	Canal 1: diferencia
1001	Canal 0: derecho	Canal 1: diferencia
1010	Canal 0: semisuma	Canal 1: diferencia
1011 - 1111	Reservado	

A continuación hay 3 bits referidos a la resolución en bit de la muestra. Igual que para el caso de la tasa de muestreo, se ha reducido la cantidad de bits para las resolucio-

nes más frecuentemente utilizadas (tabla A11.11). Luego aparece un bit 0, reservándose el 1 para un uso futuro no explicitado

**Tabla A11.11.** Códigos de resolución de la muestra en el encabezamiento de un cuadro.

Código	Resolución $n_b$ de la muestra en bit
000	Se obtiene del bloque <i>streaminfo</i>
001	8
010	12
011	Reservado
100	16
101	20
110	24
111	Reservado

Sigue luego un valor que indica el número de cuadro  $N$  si los bloques son de longitud constante  $L$  o el número de la primera muestra  $n_N$  del cuadro si los bloques son de longitud variable  $L_k$ . En el primer caso, el número de la primera muestra podría calcularse simplemente como

$$n_N = (N - 1)L + 1. \quad (\text{A11.1})$$

En el segundo caso habría que sumar las longitudes de los  $N - 1$  bloques anteriores, lo cual es un poco más complicado:

$$n_N = \sum_{k=1}^{N-1} L_k + 1 \quad (\text{A11.2})$$

Dado que el número de cuadros es mucho menor que el de muestras, su representación requiere menos bits y es, por lo tanto, más económica.

El número de cuadro o de muestra inicial del cuadro (según el caso) se codifica en el archivo con un método similar al que utiliza el código UTF-8 para codificar los caracteres UNICODE. Se trata esencialmente de un código de longitud variable y consiste en distribuir de la siguiente forma los bits de la representación binaria del número a codificar. Si el número es menor o igual a 127, se representa normalmente con 8 bit (de los cuales el más significativo es 0). Si es mayor que 127, se utilizan dos o más bytes donde los bytes menos significativos empiezan el prefijo 10 seguido por 6 de los bits originales que aun no se han distribuido en bytes menores. El byte más significativo comienza por tantos 1's como el número de bytes totales a utilizar (es decir, un código unario) seguido por uno o más 0's, seguidos por el resto de los bits aun no distribuidos. Así, por ejemplo, el número decimal 11521, que en binario es

101101 00000001,

quedará representado como

**11100010 10110100 10000001** ,

donde se han remarcado en negrita los prefijos y se han subrayado los bits originales. Observemos que la decodificación es muy sencilla, ya que del primer prefijo se sabe la cantidad de bytes, y luego basta eliminar los prefijos.

Si el código del tamaño de bloque (ver tabla A11.8) era 0110 ó 0111, entonces los próximos 8 bit ó 16 bit, respectivamente, representan el tamaño del bloque en muestras. Análogamente, si el código de la tasa de muestreo (ver tabla A11.9) era 1100, 1101 ó 1110, entonces los subsiguientes 8 bit ó 16 bit, respectivamente, representan los valores de la tasa de muestreo en kHz ( $F_s \leq 128$  kHz), en Hz ( $F_s \leq 65535$  Hz) o en decenas de Hz ( $F_s \leq 655350$  Hz).

Por último, el encabezamiento de cada cuadro termina con un byte que contiene el *código de comprobación de redundancia cíclica CRC-8 (cyclic redundancy code)* correspondiente a todos los bits anteriores del encabezamiento, incluido el código de sincronismo inicial (Williams, 1993).

Después del encabezamiento del cuadro (ver figura A11.1) aparecen los subcuadros, cada uno de los cuales a su vez se inicia con un encabezamiento propio. Este encabezamiento de subcuadro comienza con un bit 0 lo cual evita que el subcuadro pueda interpretarse como un nuevo cuadro (que comenzaría con el código de sincronismo), luego de lo cual aparece un código de 6 bit que establece qué tipo de modelo predictivo utiliza el presente subcuadro. Los valores se indican en la tabla A11.12.

**Tabla A11.12.** Códigos asociados a los modelos predictivos de un subcuadro.

Código	Modelo predictivo del subcuadro
000000	Constante
000001	Verbatim (idéntico)
00001x	Reservado
0001xx	Reservado
001000	Polinomial de orden 0
001001	Polinomial de orden 1
001010	Polinomial de orden 2
001011	Polinomial de orden 3
001100	Polinomial de orden 4
001101 - 001111	Reservado
01xxxx	Reservado
1xxxxx	LPC de orden xxxxx + 1

Luego sigue un bit que vale 1 si en el subbloque original hay bits desperdiciados por muestra y 0 si no los hay. Se consideran bits desperdiciados los bits menos significativos que son sistemáticamente 0 a lo largo del subbloque. En el primer caso se incluye la cantidad  $k$  de bits desperdiciados por muestra, expresada en código unario como  $k - 1$ . Por ejemplo, si se desperdician 4 bits (lo que equivale a que todas las muestras son múltiplos de  $2^4 = 16$ ) tendríamos la cadena de bits 10001, donde el primer 1 indica que hay bits desperdiciados, los siguientes tres 0's indican que los bits desperdiciados son  $3 + 1 = 4$ , y el último 1 es el símbolo de terminación del código unario.

Los bits desperdiciados se utilizan en algunos precodificadores con pérdida, como lossyWAV, que determinan si se puede eliminar información de los bits más significativos sin un efecto audible. La salida de estos precodificadores mantiene la cantidad original de bits pero reemplaza por 0 uno o más bits menos significativos.

La información siguiente depende del modelo predictivo utilizado en el presente subcuadro (la estrategia de predicción puede cambiar de un subbloque a otro). La estructura general es una lista de parámetros del modelo seguida por una serie de residuos.

### A11.2.1 Modelización constante

En este caso, dado que se conoce el tamaño  $B$  del bloque en muestras, basta proporcionar simplemente el valor PCM del valor constante sin compresión alguna.<sup>133</sup> Dado que este método de predicción es seleccionado únicamente si el codificador detectó que la señal es constante a lo largo del subbloque, los residuos serán siempre 0 y no es necesario incluirlos. Es esencialmente un código tipo *run-length*, en el que se especifica un valor y el número de veces que se repite.

### A11.2.2 Modelización mediante la propia señal (verbatim)

En este caso siguen los  $n_b B$  bits correspondientes a  $B$  muestras en el subbloque y  $n_b$  bits por muestra. Obsérvese que en este caso también los residuos serán siempre 0. Como modelo es muy ineficiente, ya que requiere tantos parámetros como muestras hay en el subbloque. Sólo se justifica cuando la señal tiene máxima entropía (es decir, carece de redundancia) y por lo tanto cualquier otro modelo producirá residuos prácticamente imposibles de comprimir.

### A11.2.3 Modelización polinomial

En principio, este modelo sólo necesitaría como parámetro el orden  $P$  del modelo polinomial (tabla 3.8), que ya está incluido en el encabezamiento (tabla 3.20) porque una vez definido el orden, los coeficientes son fijos (tabla 3.8). Sin embargo, dado que para calcular el valor predicho se requieren  $P$  muestras anteriores, las primeras  $P$  muestras no se pueden obtener ya que no se dispone de muestras previas al comienzo del subbloque. Por ello se deben proporcionar como parámetros del modelo las  $P$  muestras iniciales en formato PCM no comprimido. La predicción comenzará a partir de la muestra  $P + 1$ .

Es interesante notar que el modelo de orden 0 no es equivalente al modelo constante, como a primera vista podría parecer, ya que en este último no se calculan residuos dado que se parte de la premisa de que la señal era constante y por lo tanto el modelo predice exactamente los valores de la señal. En el modelo polinomial de orden 0, el residuo coincide con la señal. Un ejemplo de aplicación sería un caso de un silencio imperfecto, es decir un ruido de muy bajo nivel. En este caso es muy probable que sea una señal de alta entropía, es decir, muy poco redundante, lo cual significa que cualquier modelo dará residuos comparables con la propia señal, de allí que se prefiera el modelo más simple, el de orden 0.

Luego de los parámetros del modelo (las  $P$  muestras iniciales), aparecen los residuos codificados mediante Golomb-Rice según se describe más adelante.

### A11.2.4 Modelización LPC

Al igual que en el caso anterior, el modelo LPC comienza con una cantidad de muestras iniciales igual al orden  $P$  del predictor LPC. Los coeficientes del modelo LPC

<sup>133</sup> Intentar comprimir este valor no se justifica ya que brindaría sólo una mejora despreciable

ideales (es decir, los que minimizan el error cuadrático de los residuos) tienen potencialmente una precisión infinita, por lo que será necesario cuantizarlos.<sup>134</sup>

Los 4 bits que siguen contienen el valor  $n_{LPC} - 1$ , donde  $n_{LPC}$  es la cantidad de bits que se asignan a los coeficientes. Así, si  $n_{LPC} = 8$ , significa que los coeficientes estarán dados con una resolución de 8 bit. No se admite la combinación 1111, de modo que la máxima resolución, que corresponde al código 1110, es de 15 bit.

Los 5 bits siguientes indican, en formato complemento a 2,<sup>135</sup> la cantidad  $m$  de bits que se debe desplazar hacia la derecha (si es positivo) o hacia la izquierda (si es negativo) la representación de los coeficientes LPC. Dicho desplazamiento equivale a dividir por  $2^m$ . Así, por ejemplo, si un coeficiente LPC se representa con el binario de 8 bit  $a_h = 01110101$  (decimal 117) y corresponde desplazarlo 4 bit (es decir, dividirlo por  $2^4 = 16$ ), entonces el valor real del coeficiente será el binario 0111,0101, es decir, decimal 7,3125.

Luego siguen los  $P$  coeficientes, cada uno representado con  $n_{LPC}$  bits, totalizando una cantidad de bits igual a  $n_{LPC} P$ .

Una vez establecidos los parámetros del modelo aparecen los residuos con codificación de Golomb-Rice.

### A11.2.5 Residuos

En los casos de los modelos polinomial y LPC, el formato FLAC contiene una lista de los residuos (error de predicción) para cada muestra. El código de Golomb-Rice utilizado requiere un parámetro, el logaritmo en base 2 del divisor, o *parámetro de Rice*. FLAC permite representarlo con 4 bits o con 5 bits, para lo cual el formato prevé los siguientes 2 bits: 00 corresponde a 4 bits y 01 a 5 bits, mientras que los códigos 10 y 11 están reservados.

El parámetro de Rice puede ser único para todo el subbloque o, por el contrario el bloque puede estar subdividido en  $Q = 2^q$  particiones iguales y especificarse para cada una un parámetro de Rice diferente. Los 4 bits siguientes especifican el *orden de la partición*, es decir el exponente  $q$ .

El formato de cada partición incluye, según el caso, 4 ó 5 bits que determinan el parámetro de Rice y luego una cantidad  $n$  de residuos dada por

$$n = \begin{cases} B - P & \text{si hay una partición} \\ B/2^q - P & \text{si hay } 2^q \text{ particiones y es la primera} \\ B/2^q & \text{si hay } 2^q \text{ particiones y no es la primera} \end{cases} \quad (\text{A11.3})$$

donde  $B$  es la cantidad de muestras del bloque y  $P$  el orden del modelo. El hecho de restar  $P$  muestras a la primera (o única) partición se debe a que esas muestras corresponden a los valores iniciales que se incluyeron entre los parámetros del modelo.

<sup>134</sup> Esto hará que el modelo no cumpla con el mínimo error cuadrático, aunque se puede lograr un acercamiento aceptable, desde ya muy superior al que se logra con una aproximación polinomial cuyos coeficientes son fijos.

<sup>135</sup> El complemento a 2 es un formato que permite una representación de los números negativos. Para  $n$  bits permite representar números entre  $-2^{n-1}$  y  $2^{n-1} - 1$ . Los números positivos se representan en binario convencional, siendo 0 su bit más significativo. El código de un número negativo se obtiene complementando cada bit, sumando 1 e ignorando el acarreo en caso de existir. Así, con 5 bit se pueden representar números entre -16 y 15. Por ejemplo, el 6 se representa como 00110 y el -6 como 11001 + 1 = 11010



En el caso de representación con 4 bits, 0000 corresponde a un divisor igual a 1 (se aplicaría a una sucesión de residuos muy pequeños). El valor 1111 es considerado un *código de escape*, que implica que los residuos no están codificados. Los próximos 5 bits establecen, en ese caso, la resolución en bits de la muestra. Análogamente en el caso de que el parámetro de Rice esté representado con 5 bits, en cuyo caso el código de escape es 11111.

Si bien la ecuación (A11.3) indica la cantidad de residuos, no es posible a priori saber cuántos bits ocupan, ya que el código es de longitud variable. La única forma es decodificar las sucesivas muestras hasta completar la cantidad  $n$  dada en (A11.3).

### **A11.3 Detección de errores**

Cada cuadro finaliza con un código de redundancia cíclica CRC-16 de la totalidad de los datos del cuadro desde el código de sincronismo hasta el bit inmediatamente anterior al CRC

## Apéndice 12

### Definición de tipo de documento (DTD) para proyectos de Audacity

#### Introducción

Los archivos de proyecto .aup de Audacity son un tipo particular de documentos de texto que utilizan una estructura genérica basada en XML, un metalenguaje<sup>136</sup> de marcas (*extensible markup language*) que permite construir lenguajes de marcas específicos para diversas aplicaciones (Bray, 2008). Un lenguaje de marcas es, en general, un lenguaje que permite incluir texto, explicitando, mediante unos marcadores que delimitan fragmentos del texto, diversas categorizaciones de dichos fragmentos. Éstas pueden abarcar cuestiones estructurales, semánticas, de formato de texto, etc. La mayor utilidad de los lenguajes de marcas es permitir el análisis sintáctico del texto (en inglés, *parsing*) en forma automática y de una manera sencilla, así como extraer información ya clasificada por categorías.

Los marcadores son de la forma `<xxx>` y `</xxx>` cuando delimitan un texto no vacío o `<xxx/>` cuando corresponden a un texto vacío (hay casos donde es interesante que aparezca un marcador no asociado a un texto específico). Aquí xxx es el nombre asociado al marcador. A su vez puede haber marcadores anidados, correspondientes a subcategorías. Cada marcador puede tener asociados atributos y cada atributo toma un valor (puede ser numérico, textual o alfanumérico) que modifica al marcador.

#### Definiciones de tipo de documento (DTD)

Una definición de tipo de documento (*document type definition*, DTD) es una declaración de la sintaxis de un tipo particular de documentos XML (Miloslav, s/f), en este caso los correspondientes a archivos de proyecto de Audacity.

Contiene una serie de líneas encerradas entre delimitadores de tres tipos. Las encerradas entre `<!--` y `-->` son comentarios; las encerradas entre `<!ELEMENT` y `>` definen los *elementos* es decir, componentes principales de la sintaxis; y las delimitadas por `<!ATTLIST` y `>` son atributos que modifican y personalizan los elementos.

Un elemento puede tener elementos anidados (hijos), en ese caso los que están en el nivel inmediato inferior se delimitan con paréntesis y dentro de los paréntesis se separan con comas. La siguiente lista muestra los recursos para combinar elementos:

- xxx\* significa que puede haber 0 o más elementos del tipo xxx
- xxx+ significa que puede haber 1 o más elementos del tipo xxx
- xxx? indica que xxx es un elemento opcional (puede aparecer 0 ó 1 vez)
- , delimita elementos dentro de una sucesión ordenada
- | delimita elementos alternativos excluyentes

Así, `xxx (yyy, (zzz | www)*)` significa que un elemento tipo xxx contiene, anidados un elemento tipo yyy seguido de cualquier combinación y en cualquier orden de elementos tipo zzz y www.

La siguiente lista indica algunas palabras clave sobre el contenido de elementos y atributos:

---

<sup>136</sup> Un metalenguaje es un lenguaje cuyo objeto es referirse a otro u otros lenguajes.

#PCDATA indica un texto arbitrario (incluye caracteres reservados como < y >)  
 CDATA indica un texto formado por caracteres  
 #REQUIRED indica que es requerido  
 #IMPLIED indica que es un atributo opcional  
 EMPTY es un elemento que sólo puede contener atributos, no texto

### DTD para archivos de proyecto de Audacity

A continuación se detalla el contenido de la definición de tipo de documento (DTD) para los archivos de proyecto de Audacity (cuya extensión es .aup), disponible en <http://audacity.sourceforge.net/xml/audacityproject-1.3.0.dtd>.<sup>137</sup> En negrita y tipografía grande se transcriben las diferentes entradas, que en el original se agrupan por elemento. En tipografía pequeña e indentada, se agregan una descripción y comentarios.

**<!--DTD for Audacity Project files, as created by Audacity version 1.3.0 -->**

**<!--For more information: <http://audacity.sourceforge.net/xml/> -->**

**<!ELEMENT project (tags, (wavetrack | labeltrack | timetrack)\*)>**

El elemento project contiene un elemento tags (ver más abajo) y puede contener 0 o más elementos a seleccionar entre wavetrack, labeltrack o timetrack, en cualquier orden  
 Sus atributos son:

**<!ATTLIST project projname CDATA #REQUIRED>**

Nombre del proyecto (es el nombre del archivo seguido por \_data)

**<!ATTLIST project version CDATA #REQUIRED>**

Versión del DTD

**<!ATTLIST project audacityversion CDATA #REQUIRED>**

Versión de Audacity con la que se editó

**<!ATTLIST project sel0 CDATA #REQUIRED>**

Inicio de la última selección guardada, en segundos

**<!ATTLIST project sel1 CDATA #REQUIRED>**

Fin de la última selección guardada, en segundos

**<!ATTLIST project vpos CDATA #REQUIRED>**

Posición vertical ¿?

**<!ATTLIST project h CDATA #REQUIRED>**

Posición horizontal ¿?

**<!ATTLIST project zoom CDATA #REQUIRED>**

Zoom horizontal respecto a una ventana de aproximadamente 15 s

**<!ATTLIST project rate CDATA #REQUIRED>**

Tasa de muestreo en Hz

<sup>137</sup> La versión declarada se encuentra desactualizada. Ver al final del apéndice

**<!ELEMENT tags EMPTY>**

El elemento tags contiene metadatos típicos de los registros musicales.  
Sus atributos son:

**<!ATTLIST tags title CDATA #REQUIRED>**

Título

**<!ATTLIST tags artist CDATA #REQUIRED>**

Artista

**<!ATTLIST tags album CDATA #REQUIRED>**

Álbum

**<!ATTLIST tags track CDATA #REQUIRED>**

Pista

**<!ATTLIST tags year CDATA #REQUIRED>**

Año

**<!ATTLIST tags genre CDATA #REQUIRED>**

Género

**<!ATTLIST tags comments CDATA #REQUIRED>**

Comentarios

**<!ATTLIST tags id3v2 (0|1) #REQUIRED>**

ID3v2 es un “estándar de facto” para identificar registros musicales y sus metadatos, que incluyen datos como el autor, intérprete, título, etc. El estándar se puede descargar de <http://id3.org/id3v2.3.0>

**<!ELEMENT labeltrack (label\*)>**

El elemento labeltrack (Pista de etiquetas) puede contener 0 o más etiquetas (labels)  
Sus atributos son:

**<!ATTLIST labeltrack name CDATA #REQUIRED>**

Nombre de la pista de etiquetas

**<!ATTLIST labeltrack numlabels CDATA #REQUIRED>**

Cantidad de etiquetas

**<!ELEMENT label EMPTY>**

Cada elemento etiqueta se refiere a un texto, en general corto, asociado a un intervalo de tiempo (que puede ser de duración 0). No tiene contenido, estando su información expresada en sus atributos, que son:

**<!ATTLIST label t CDATA #REQUIRED>**

Instante inicial en segundos

**<!ATTLIST label t1 CDATA #REQUIRED>**

Instante final en segundos

**<!ATTLIST label title CDATA #REQUIRED>**

Título (el texto de la etiqueta)

**<!ELEMENT timetrack (envelope)>**

El elemento timetrack (Pista de tiempo) contiene una envolvente (envelope) formada por puntos de control y valores asociados a cada uno de ellos (ver más abajo).  
Sus atributos son:

**<!ATTLIST timetrack name CDATA #REQUIRED>**

Nombre del time track

**<!ATTLIST timetrack channel CDATA #REQUIRED>**

Canal al que se aplica

**<!ATTLIST timetrack offset CDATA #REQUIRED>**

Desplazamiento del timetrack desde el instante inicial

**<!ELEMENT wavetrack (waveclip\*)>**

Cada wavetrack contiene 0 o más waveclips. Un waveclip es un fragmento de audio.  
Un elemento wavetrack lleva los siguientes atributos:

**<!ATTLIST wavetrack name CDATA #REQUIRED>**

Nombre que aparece en el panel de control del track

**<!ATTLIST wavetrack channel CDATA #REQUIRED>**

Canal asignado

**<!ATTLIST wavetrack linked CDATA #REQUIRED>**

¿?

**<!ATTLIST wavetrack offset CDATA #REQUIRED>**

Desplazamiento en segundos del comienzo del wavetrack (aparentemente no se usa)

**<!ATTLIST wavetrack rate CDATA #REQUIRED>**

Tasa de muestreo del wavetrack (podría ser diferente entre los distintos tracks)

**<!ATTLIST wavetrack gain CDATA #REQUIRED>**

Ganancia a aplicar al wavetrack, incluyendo todos los waveclips que contenga

**<!ATTLIST wavetrack pan CDATA #REQUIRED>**

Paneo a aplicar al wavetrack, incluyendo todos los waveclips que contenga

**<!ATTLIST wavetrack mute CDATA "0">**

Sordina (0 es el valor por defecto, correspondiente a sordina desactivada)

**<!ATTLIST wavetrack solo CDATA "0">**

Solo (0 es el valor por defecto, correspondiente a solo desactivado)

**<!ELEMENT waveclip (sequence, envelope)>**

Cada wavclip contiene una secuencia formada por un cierto número (declarado) de muestras, a su vez subdividida en waveblocks, es decir segmentos de  $2^{20}$  bytes (aproximadamente 1 Mb), y una envolvente (envelope) formada por puntos de control y valores asociados a cada uno de ellos  
Un waveclip posee un único atributo:

**<!ATTLIST waveclip offset CDATA #REQUIRED>**

Desplazamiento en segundos desde el comienzo del track

**<!ELEMENT sequence (waveblock\*)>**

Una secuencia está formada por 0 o más waveblocks (segmentos de  $2^{20}$  bytes)  
Tiene los siguientes atributos:

**<!ATTLIST sequence maxsamples CDATA #REQUIRED>**

Máximo número de muestras de los waveblocks que la forman. Es igual a  $2^{20}/\text{bpm}$ , donde bpm es la cantidad de bytes por muestra

**<!ATTLIST sequence sampleformat CDATA #REQUIRED>**

Cantidad de muestras más 11 palabras de 4 bytes correspondientes al encabezamiento

**<!ATTLIST sequence numsamples CDATA #REQUIRED>**

Cantidad de muestras de la secuencia

**<!ELEMENT waveblock (simpleblockfile | silentblockfile | legacyblockfile | pcmaliasblockfile)>**

Un waveblock está formado por uno y sólo uno de 4 tipos de bloques.  
Tiene un solo atributo:

**<!ATTLIST waveblock start CDATA #REQUIRED>**

Número de muestra en la que comienza el bloque. Si hay más de un bloque son múltiplos de  $2^{20}/\text{bpm}$ . Por ejemplo, si  $\text{bpm} = 2$  (equivalente a una resolución de 16 bit), entonces será 524288

**<!ELEMENT simpleblockfile EMPTY>**

El waveblock más común es de tipo simpleblockfile. No tiene contenido más que los atributos  
Tiene los siguientes atributos:

**<!ATTLIST simpleblockfile filename CDATA #REQUIRED>**

Nombre del archivo. Es de la forma "e000075b.au"

**<!ATTLIST simpleblockfile len CDATA #REQUIRED>**

Longitud en muestras del waveblock

**<!ATTLIST simpleblockfile min CDATA #REQUIRED>**

Mínimo valor decimal normalizado entre -1 y 1 de todas las muestras del waveblock

**<!ATTLIST simpleblockfile max CDATA #REQUIRED>**

Máximo valor decimal normalizado entre -1 y 1 de todas las muestras del waveblock

**<!ATTLIST simpleblockfile rms CDATA #REQUIRED>**

Valor eficaz de las muestras normalizadas entre -1 y 1

**<!ELEMENT silentblockfile EMPTY>**

El elemento silentblockfile representa un bloque silencioso. No tiene contenido más que su único atributo:

**<!ATTLIST silentblockfile len CDATA #REQUIRED>**

Longitud en muestras del silencio

**<!ELEMENT legacyblockfile EMPTY>**

El elemento `legacyblockfile` representa un bloque en un formato anterior (correspondiente a las versiones de Audacity 0.98 a 1.0 y 1.1.0 a 1.1.2. Es posible cargar, editar y volver a guardar un bloque en ese formato pero no crear uno nuevo.

(Más información en [http://audacity.sourceforge.net/dox/\\_legacy\\_block\\_file\\_8h\\_source.html](http://audacity.sourceforge.net/dox/_legacy_block_file_8h_source.html)).

Tiene los siguientes atributos:

**<!ATTLIST legacyblockfile name CDATA #REQUIRED>**

Nombre del archivo

**<!ATTLIST legacyblockfile len CDATA #REQUIRED>**

Longitud del archivo en muestras ( $2^{20}$ )

**<!ATTLIST legacyblockfile summarylen CDATA #REQUIRED>**

Longitud del sumario incluido en el formato

**<!ATTLIST legacyblockfile norms CDATA "0">**

Variable booleana que vale 1 si no hay valor eficaz (RMS)

**<!ELEMENT pcmaliasblockfile EMPTY>**

El elemento `pcmaliasblockfile` representa audio almacenado en un archivo “alias” externo al directorio de datos del proyecto (lo cual sucede cuando se importa un archivo sin guardar una copia).

Tiene los siguientes atributos:

**<!ATTLIST pcmaliasblockfile summaryfile CDATA #REQUIRED>**

Nombre de un archivo de extensión `.auf` que contiene un sumario del archivo alias. Éste consiste en un submuestreo por un factor 1/85 de los datos de audio que corresponderían si la señal original se hubiera copiado al directorio de datos del proyecto.

NOTA: Si el archivo original es largo, se lo subdivide virtualmente en fragmentos de hasta  $2^{20}$  bytes. El sumario contiene, por lo tanto, hasta  $2^{20} / 85$  muestras, precedidas por un breve encabezamiento (ver sección 3.16).

**<!ATTLIST pcmaliasblockfile aliasfile CDATA #REQUIRED>**

Ruta completa y nombre del archivo alias

**<!ATTLIST pcmaliasblockfile aliasstart CDATA #REQUIRED>**

Número de muestra dentro del archivo alias donde se inicia el fragmento de audio al que hace alusión el archivo `.auf` de sumario. Para el primer archivo de sumario es 0, luego son múltiplos de  $2^{20}/4$ .

**<!ATTLIST pcmaliasblockfile aliaslen CDATA #REQUIRED>**

Cantidad de muestras que contiene el fragmento de audio correspondiente al archivo de sumario. Si hay más de un archivo de sumario, los primeros tienen  $2^{20}/4$  muestras. El último puede contener menos.

**<!ATTLIST pcmaliasblockfile aliaschannel CDATA #REQUIRED>**

Canal correspondiente al archivo alias.

**<!ATTLIST pcmaliasblockfile min CDATA #REQUIRED>**

Mínimo valor decimal normalizado entre -1 y 1 del segmento de audio correspondiente al archivo de sumario.

**<!ATTLIST pcmaliasblockfile max CDATA #REQUIRED>**

Máximo valor decimal normalizado entre -1 y 1 del segmento de audio correspondiente al archivo de sumario.

**<!ATTLIST pcmaliasblockfile rms CDATA #REQUIRED>**

Valor eficaz (RMS) decimal normalizado entre -1 y 1 del segmento de audio correspondiente al archivo de sumario.

**<!ELEMENT envelope (controlpoint\*)>**

Cada waveclip puede llevar, además de los archivos de audio que lo forman, una envolvente, definida por 0 o más puntos de control. La envolvente es una función en general lentamente variable entre 0 y 1 por la que quedarán multiplicadas las sucesivas muestras del waveclip. Tiene un solo atributo:

**<!ATTLIST envelope numpoints CDATA #REQUIRED>**

Número de puntos de control de la envolvente

**<!ELEMENT controlpoint EMPTY>**

Un punto de control es un par ordenado (tiempo en s, valor normalizado entre -1 y 1) que pertenece a la envolvente. La envolvente se completa interpolando exponencialmente entre puntos sucesivos. No contiene más datos que los definidos por sus atributos

**<!ATTLIST controlpoint t CDATA #REQUIRED>**

Instante de tiempo en segundos

**<!ATTLIST controlpoint val CDATA #REQUIRED>**

Valor de la envolvente en el instante de tiempo especificado.

NOTA: El DTD anterior, declarado en los propios archivos .aup, está desactualizado. En efecto, la versión vigente de Audacity (2.0.3 en julio de 2013) genera archivos de proyecto .aup que no cumplen con dicho DTD, válido para la versión 1.3 (Audacity team, s/f). Dos posibles versiones son las siguientes:

**<!ELEMENT tags (tag\*)>**

**<!ELEMENT tag EMPTY>**

**<!ATTLIST tag name CDATA #REQUIRED>**

**<!ATTLIST tag value CDATA #REQUIRED>**

**<!ELEMENT tags (tag\*)>**

**<!ELEMENT tag EMPTY>**

**<!ATTLIST tag (TITLE | ARTIST | ALBUM | TRACK | YEAR | GENRE | COMMENTS | ID3V2)>**

**<!ATTLIST tag value CDATA #REQUIRED>**

En la primera, más general, los nombres de los tags son arbitrarios; en la segunda son los correspondientes al uso en la industria discográfica.



## Apéndice 13

### Breve descripción de Scilab

Scilab es un software libre de cálculo matricial, en el sentido de que puede operar directamente con matrices y vectores, tanto en lo que respecta a las operaciones clásicas de la teoría de matrices (suma, producto, transposición, cálculo de determinantes, inversión de matrices cuadradas, etc.) como a operaciones en masa sobre un conjunto de datos (por ejemplo, aplicar una función o procedimiento a todas las componentes de una matriz). Esto lo hace particularmente atractivo para el procesamiento de señales digitales ya que con una sola instrucción se puede aplicar un procedimiento iterativo. Puede descargarse de <https://www.scilab.org/> y posee una licencia CeCILL compatible con la GPL (*General Public License*, Licencia Pública General) de la Free Software Foundation.<sup>138</sup> Se pueden descargar los instaladores para Linux, Windows o Mac OS X en <http://www.scilab.org/download>.

Hay muchas formas de crear matrices en Scilab. La más simple es usando el operador de asignación, "=", encerrando entre corchetes una lista de números separados por comas para delimitar las componentes de una fila y puntos y coma para delimitar las filas:

```
a = [1, 2; 3, 4];
```

El punto y coma que cierra la línea de código es opcional. Si no se lo coloca Scilab presenta el resultado en pantalla, lo cual es interesante para la depuración pero puede ser molesto en el caso de matrices de gran tamaño, ya que se gasta mucho tiempo en presentar valores cuya visualización es irrelevante.

También puede usarse la coma como operador de concatenación. Si *b* y *c* son vectores fila, entonces

```
a = [b, c];
```

es un vector fila que contiene todas las componentes de *b* seguidas por las de *c*; y si *b* y *c* son vectores columna,

```
a = [b; c];
```

es un vector columna que contiene todas las componentes de *b* seguidas por las de *c*. El operador ":" permite crear un vector de números que se van incrementando de *a* una unidad entre dos valores. Por ejemplo,

```
a = [1:m];
```

es un vector fila con componentes 1, 2, ..., *m*. Intercalando un número *h* (no necesariamente entero) en la siguiente forma

```
a = [n:h:m];
```

---

<sup>138</sup> En términos generales esto implica la libertad de usar el software para cualquier propósito, ya sea en forma personal, académica o comercial, para modificarlo, para copiarlo y compartirlo, para compartir las modificaciones realizadas y para publicar dichas modificaciones siempre que se lo haga bajo la misma licencia GPL. Pueden obtenerse más detalles en <http://www.gnu.org/copyleft/gpl.html>

se crea un vector de números separados por **h**, es decir: **n**, **n + h**, **n + 2\*h**, ..., **n + r\*h**, donde **r** es tal que **n + r\*h** sea el máximo posible menor o igual que **m**.

Es posible crear matrices de **n** filas y **m** columnas formadas por ceros mediante **zeros(n,m)**, y por unos mediante **ones(n,m)**.

El tamaño (cantidad de filas y columnas) puede conocerse invocando la función **size(a)**, y si se desea la cantidad de componentes de un vector, puede invocarse **length(a)**.

El valor de la componente **n** de un vector **x** se invoca escribiendo **x(n)** y si se desea crear un subvector **y** formado por las componentes **n**, **n+1**, ..., **m** basta escribir

```
y = x(n:m) ;
```

En el caso de matrices, la componente en la fila **n** y columna **m** se invoca con **x(n,m)**. También se puede invocar la fila **n** o la columna **m** completas escribiendo, respectivamente,

```
y = a(n, :) ;  
y = a(:, m) ;
```

En cualquiera de los dos casos se puede reemplazar “:” por un intervalo:

```
y = a(n, m1:m2) ;  
y = a(n1:n2, m) ;  
y = a(n1:n2, m1:m2) ;
```

Una prima o apóstrofe “'” a la derecha del símbolo de una matriz actúa como operador de transposición. Así,

```
y = y' ;
```

convierte el vector columna en vector fila.

En el caso del procesamiento de señales de audio puede crearse un vector columna (matriz  $n \times 1$ ) a partir de un archivo de sonido mediante el procedimiento **wavread**. Así,

```
[x, Fs, Nbits] = wavread('ruta_y_nombre_de_archivo') ;
```

crea un vector **x** que contiene las muestras de audio normalizadas entre  $-1$  y  $+1$  (es decir que al mínimo valor binario se le asigna el valor  $-1$  y al máximo,  $+1$ ). Los parámetros **Fs** y **Nbits** proporcionan la tasa de muestreo en Hz y el número de bits (8 ó 16). La función **sound** permite reproducir un sonido:

```
sound(x, Fs, Nbits)
```

Los operadores de multiplicación, división y potenciación, **\***, **/**, **^**, actúan matricialmente o, si son precedidos por un punto, **.\***, **./**, **.^**, lo hacen componente por componente. Las funciones como el logaritmo natural (**log**), logaritmo decimal (**log10**), seno (**sin**), coseno (**cos**) o tangente (**tan**) actúan por componentes.

Scilab puede trabajar en forma simbólica con polinomios y funciones racionales. Así, la función **poly** define un polinomio a partir de sus ceros

```
p = poly([c1, c2], 'x') ;
```

o bien a partir de sus coeficientes

```
p = poly([a0, a1, a2], 'x', 'coeff');
```

La función **roots** obtiene las raíces de un polinomio:

```
z = roots(p);
```

El resultado es un vector columna de números complejos.

Se puede operar con polinomios como si fueran números. Un cociente de polinomios proporciona una función racional. Scilab reduce numerador y denominador removiendo factores comunes, siempre y cuando sean exactamente iguales. Una forma cómoda de generar polinomios es crear un polinomio elemental

```
x = poly(0, 'x');
```

equivalente a la variable  $x$  y luego operar. Por ejemplo,

```
p = 1 - 2*x + x^2;
```

obtiene el polinomio  $1 - 2x + x^2$ .

Para evaluar un polinomio (o función racional)  $p$  en un valor del argumento se aplica la función **horner**. Así,

```
q = horner(p, 5);
```

evalúa el polinomio (o función racional)  $p$  en  $x = 5$ . **horner** también funciona si el segundo argumento es una matriz, en cuyo caso el polinomio (o función racional) se evalúa en cada componente de la matriz. También es aplicable a un argumento que a su vez es otro polinomio (o función racional).

Se puede definir en Scilab una función de transferencia **H** a partir de dos polinomios **p** (numerador) y **q** (denominador) mediante la función **syslin**:

```
H = syslin('d', p, q);
```

donde **'d'** indica que se trata de una función de transferencia discreta. Para funciones de transferencia en tiempo continuo se utilizaría **'c'**.

La transferencia en tiempo discreto puede aplicarse a una señal  $x$  mediante la función **flts**:

```
y = flts(x, H);
```

Puede realizarse el análisis de espectro de una señal mediante el uso de la transformada discreta de Fourier, implementada mediante el algoritmo conocido como *Transformada Rápida de Fourier*, FFT. Así, si  $x$  es un vector,

```
y = fft(x, -1);
```

es su transformada discreta de Fourier. El parámetro  $-1$  indica el signo del exponente de las exponenciales que se utilizan en la transformada. Si en su lugar se utiliza  $1$  y se divide el resultado por la longitud de  $x$  se obtiene la transformada rápida inversa de Fourier. Si se desea restringir el número de puntos a los  $N$  primeros, por ejemplo, basta poner

```
y = fft(x(1:N), -1);
```

Es posible graficar una señal mediante el comando `plot2d`. Aunque tiene muchas opciones, se obtiene una gráfica sencilla mediante

```
plot2d(x, y)
```

donde `x` e `y` son vectores de igual longitud que contienen los valores de la abscisa y la ordenada. `plot2d` produce gráficas segmento-lineales, pero a partir de varias decenas de puntos, la imagen aparece como una curva suave. Si `y1` e `y2` son vectores columna de igual longitud que `x`, se pueden graficar simultáneamente `y1` e `y2` versus `x` mediante

```
plot2d(x, [y1, y2])
```

Pueden abrirse nuevas ventanas gráficas, para lo cual se utiliza

```
xset('window', i);
```

donde `i` es el número de ventana. Si no se ha abierto ninguna ventana antes de la primera invocación de `plot2d`, se utiliza la ventana 0. Dentro de una misma ventana gráfica pueden incluirse varios gráficos diferentes mediante el comando `subplot`. Así,

```
subplot(m,n,p)
```

divide la ventana en `m` filas y `n` columnas de gráficas y ubica la próxima gráfica a trazar en la celda `p`-ésima, contando primero por filas y después por columnas.

Como en todos los lenguajes de programación, Scilab permite escribir estructuras condicionales y de control. La estructura `for` permite repetir una cantidad especificada de veces ciertas instrucciones. Por ejemplo

```
a = [];  
for i=1:5  
    a = [a, i*ones(1,i)];  
end
```

crea el vector `a = [1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5]`. La estructura `if` permite realizar o no una serie de operaciones según cierta condición:

```
if i < 5 then  
    a = 1;  
elseif i < 10 then  
    a = 2;  
else  
    a = 3;  
end
```

La estructura `while` repite una serie de operaciones mientras una expresión lógica dada resulte verdadera. Las instrucciones

```
i = 6;  
a = [];  
while (5 < i) & (i < 10)  
    a = [a, i];  
    i = i + 1;  
end
```

crean el vector [6, 7, 8, 9]. Si el valor inicial de *i* hubiera sido  $\leq 5$  ó  $\geq 10$ , el lazo no se habría ejecutado.

Por último, la estructura **select** permite realizar diversas operaciones según el valor de una variable. En la siguiente estructura,

```
select a
case 1
    b = 1;
case 2
    b = 3;
case 3
    b = 0;
end
```

si *a* = 2 se asigna a *b* el valor 3.

Las funciones proporcionadas por Scilab cubren un amplio espectro de aplicaciones. Sin embargo, frecuentemente surge la necesidad de introducir funciones específicas no implementadas. Es posible definir nuevas funciones en Scilab mediante la sintaxis

```
function [y1, ..., yn] = nombre(x1, ..., xm)
    instrucciones
endfunction
```

donde **x1, ..., xm** son variables de entrada e **y1, ..., yn** son variables de salida, **nombre** es el nombre de la función e **instrucciones** es una secuencia de instrucciones en las que se definen los valores de *y1, ..., yn*.

La sintaxis anterior puede ingresarse directamente *en línea*, es decir en la ventana principal de Scilab o en cualquier editor de texto (por ejemplo SciNotes, el editor interno de Scilab). En el primer caso, una vez definida la función puede invocarse directamente mediante

```
nombre(x1, ..., xm);
```

donde las variables de entrada **x1, ..., xm** deben haber sido asignadas previamente. En el segundo caso, una vez guardado el archivo con extensión .sci, es necesario cargarlo, para lo cual se usa la expresión

```
exec ruta
```

donde *ruta* es la ruta completa del archivo .sci. Una vez cargada puede invocarse igual que en el caso anterior.

Si se desea que la función esté permanentemente disponible para Scilab será preciso incluirla en una biblioteca. Por ejemplo, puede usarse la instrucción:

```
genlib('worklib', 'SCI/macros/work/', %f, %t)
```

para generar la biblioteca **worklib** (incluyendo su estructura)<sup>139</sup> a partir de las funciones ubicadas como archivos **.sci** en el directorio **SCI/macros/work/**.<sup>140</sup> La función **genlib** compila cada función en la forma de un archivo **.bin** y crea dos archivos, **names** y **lib**; **names** contiene los nombres de las funciones y **lib** es un archivo binario asociado a la biblioteca.

Para que la biblioteca así generada esté disponible cada vez que se inicia Scilab puede colocarse la instrucción

```
worklib = lib('SCI/macros/work/')
```

en un archivo **.scilab** ubicado en el directorio **SCIHOME**.<sup>141</sup> Este archivo se ejecuta durante la inicialización declarando dicha biblioteca.

En ocasiones es necesario cargar en una variable el contenido de un archivo, no necesariamente de sonido. En ese caso primero se debe abrir el archivo mediante la función **mopen** con la sintaxis

```
[fd, err] = mopen(ruta_y_nombre, modo, swap)
```

Aquí **fd** es el identificador o descriptor numérico del archivo, **err** es un código de error (0 significa que no hubo error), **modo** es una cadena numérica que establece el tipo de apertura, por ejemplo:

```
'r':  lectura
'rb' lectura de datos binarios
'rt': lectura de texto
'w':  escritura
'r+': lectura y escritura.
```

Si **swap** es 1 (el valor por defecto), cuando se detectan números en formato little-endian se produce un reordenamiento automático de los bytes para recuperar el orden normal.. Si es 0 no hace nada

Luego de abrir el archivo se puede realizar la lectura mediante las funciones **mget**, **mgetl** o **mgeti**. La función **mget** lee bytes o palabras numéricas

```
x = mget(n, tipo, fd)
```

Aquí **n** es la cantidad de ítems a leer, **tipo** indica cómo se interpretan los datos:

```
'd':  doble
'f'   float
'l':  long
'i':  int
's':  short
'c':  caracter
```

---

<sup>139</sup> Los símbolos **%f** y **%t** son las constantes lógicas *false* (falso) y *true* (verdadero)

<sup>140</sup> **SCI** es una variable de entorno que contiene una cadena de caracteres que indica el directorio donde se ha instalado Scilab.

<sup>141</sup> La variable de entorno **SCIHOME** contiene una cadena de caracteres que indica donde se encuentran algunos archivos de inicialización del usuario.

Todos estos pueden ir seguidos de

‘u’: sin signo  
‘l’ little endian  
‘b’: big endian

**fd** es el identificador de archivo. Si es  $-1$  se refiere al último archivo abierto, igual que si se omite .

Por ejemplo:

```
x = mget(10, 'ulf', 2)
```

leerá 10 valores float little endian sin signo del archivo abierto con identificador 2

La función **mgetl** lee líneas de texto:

```
texto = mgetl(file_desc,m)
```

donde **texto** es la variable que contendrá el texto, **file\_desc** puede ser la ruta y nombre del archivo o el identificador **fd** devuelto por **mopen**

## Apéndice 14

### Transformada rápida de Fourier (FFT)

La transformada discreta de Fourier está dada por

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{j \frac{2\pi k n}{N}} \quad (\text{A14.1})$$

Supondremos que el núcleo de la transformada  $e^{j2\pi kn/N}$  se puede calcular y almacenar en una tabla previamente (son sólo  $N$  valores, dado que la exponencial compleja es una función periódica con periodo  $N$ ). Entonces el cálculo de las  $N$  componentes del espectro DFT, es decir  $\{X_n\}_{n=0, \dots, N-1}$ , requiere realizar  $N^2$  multiplicaciones y  $N^2$  sumas, es decir  $2N^2$  operaciones de punto flotante. Para  $N$  grande, este proceso se vuelve computacionalmente muy costoso. Por ejemplo, si  $N = 4096$ , una resolución apropiada para representar todos los rasgos espectrales que percibe el oído humano, resultan 33 millones de operaciones por cada transformada. Esta cantidad de cálculos es prohibitiva en muchos sistemas.

En 1965, Cooley y Tuckey introdujeron un algoritmo que permite reducir el orden  $N^2$  del problema a un orden  $N \log_2 N$  (Cooley, Tuckey, 1965).<sup>142</sup> La idea básica es descomponer el problema y reducirlo al cálculo de dos transformadas de  $N/2$  puntos, con lo cual se realizan  $2 \times 2(N/2)^2 = N^2$  operaciones de punto flotante en lugar de  $2N^2$ . Si  $N$  es una potencia de 2, es posible aplicar recursivamente este procedimiento, con lo cual se logra una sustancial reducción de la cantidad de operaciones.

Para verlo, llamemos  $W_N = e^{-j2\pi/N}$ . Entonces podemos reescribir

$$X(n) = \sum_{k=0}^{N-1} x(k) W_N^{k n} \quad (\text{A14.2})$$

Podemos considerar separadamente las muestras pares e impares de la señal:

$$\begin{aligned} X(n) &= \sum_{h=0}^{\frac{N}{2}-1} \left( x(2h) W_N^{2h n} + x(2h+1) W_N^{(2h+1)n} \right) = \\ &= \sum_{h=0}^{\frac{N}{2}-1} \left( x(2h) W_N^{2h n} + W_N^n x(2h+1) W_N^{2h n} \right) \end{aligned} \quad (\text{A14.3})$$

<sup>142</sup> El algoritmo había sido en realidad inventado por Gauss en 1805 y publicado póstumamente en 1866, para simplificar los cálculos para la determinación de trayectorias de asteroides (Gauss, 1866; Wikipedia, 2013g; Heideman, et al. 1984)



Observemos que

$$W_N^{2hn} = \left(W_N^2\right)^{hn}$$

y

$$W_N^2 = \left(e^{-j2\pi/N}\right)^2 = e^{-j2\pi/(N/2)} = W_{N/2}$$

de donde

$$X(n) = \underbrace{\sum_{h=0}^{\frac{N}{2}-1} x(2h) W_{N/2}^{hn}}_{Y_n} + W_N^n \underbrace{\sum_{h=0}^{\frac{N}{2}-1} x(2h+1) W_{N/2}^{hn}}_{Z_n} \quad (\text{A14.4})$$

$Y_n$  y  $Z_n$  son, respectivamente, las transformadas discretas de Fourier de las componentes pares e impares de  $x_h$ . Podemos subdividir el cálculo de (A14.4) en los valores de  $n$  menores que  $N/2$  y los mayores o iguales a  $N/2$ . En el primer caso,

$$X(n) = Y(n) + W_N^n Z(n).$$

En el segundo caso, con un pequeño cambio de variable tenemos

$$X(n + N/2) = Y(n + N/2) + W_N^{n+N/2} Z(n + N/2)$$

Ahora, dado que  $\{x(2h)\}$  y  $\{x(2h+1)\}$  tienen sólo  $N/2$  puntos, sus transformadas son periódicas de periodo  $N/2$ , es decir,

$$Y(n + N/2) = Y(n)$$

$$Z(n + N/2) = Z(n)$$

por lo que

$$X(n + N/2) = Y(n) + W_N^{n+N/2} Z(n),$$

expresión válida para  $n = 1, \dots, N/2 - 1$ . Pero además es fácil ver que  $W_N^{N/2} = -1$ , por lo que llegamos a las relaciones siguientes

$$\begin{cases} X(n) = Y(n) + W_N^n Z(n) \\ X(n + N/2) = Y(n) - W_N^n Z(n) \end{cases} \quad (\text{A14.5})$$

Basta calcular  $Y_n$  y  $Z_n$  para  $n = 0, \dots, N/2 - 1$  para poder obtener  $X_n$  para  $n = 0, \dots, N - 1$ . Para ello se requieren  $N$  sumas y  $N/2$  multiplicaciones.

Para obtener  $Y(n)$  y  $Z(n)$  procedemos de igual manera, a cada una de ellas la subdividimos en los términos pares e impares. Por cada una realizaremos  $N/2$  sumas y  $N/4$  multiplicaciones, pero como ahora son dos las transformadas a obtener, en total tendremos  $N$  sumas y  $N/2$  multiplicaciones.

El proceso se repite de modo que en cada paso hay que calcular el doble de transformadas de la mitad de puntos. Esto sigue hasta que finalmente, en  $\log_2 N$  pasos se llega a tener que calcular  $N$  transformadas de 1 punto cada una. Pero la transformada de una señal de longitud 1 es ella misma, por lo tanto no hace falta calcular nada más. En resumen, habremos realizado  $N \log_2 N$  sumas y  $N/2 \log_2 N$  multiplicaciones. Para el caso mencionado anteriormente de  $N = 4096$ , será un total de 73728 operaciones de punto flotante, que es alrededor de 455 veces menos operaciones que haciendo los cálculos por aplicación directa de la fórmula.

¿A qué se debe un ahorro de cómputo tan espectacular? Conceptualmente, la aplicación directa de (A14.1) implica calcular muchas veces lo mismo, ya que si bien el exponente adopta  $N^2$  valores, la exponencial es periódica y sólo toma  $N$  valores distintos. El procedimiento indicado reordena los cálculos de modo de aprovechar esta situación.

El único inconveniente es que para cada paso es necesario realizar una intercalación de las componentes pares e impares, lo cual complica el direccionamiento de las componentes. Se puede verificar que es equivalente a contar desde 1 hasta  $N$  invirtiendo el orden de los bits de la numeración binaria normal (el MSB pasa a ser el LSB y viceversa). Por ejemplo, para  $N = 16$  tendríamos:

```

0000
1000
0100
1100
0010
1010
...
0111
1111

```

El algoritmo bosquejado es un ejemplo de una amplia clase de algoritmos conocidos como *Transformada Rápida de Fourier*, FFT (*Fast Fourier Transform*). Todos los paquetes de software matemático, así como los procesadores digitales de señal (DSP), incorporan algoritmos FFT altamente optimizados, incluyendo no sólo este caso particular sino también casos donde  $N$  no es potencia de 2 sino múltiplo de factores primos, donde si bien no se logra una reducción tan grande de cálculos, sí se logran reducciones significativas. Por esta razón es muy poco frecuente tener que implementar un algoritmo FFT desde cero. Aunque sin el objetivo de lograr el mejor código, es igualmente una práctica de programación recomendable escribir un código que lo realice.

## Apéndice 15

### Igualdad de Parseval y simetría de la DFT

La Transformada Discreta de Fourier de una señal discreta  $x(k)$  está dada por

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-\frac{j2\pi kn}{N}} \quad (\text{A14.1})$$

mientras que la Transformada Discreta de Fourier Inversa, que permite recuperar la señal a partir de la transformada, es

$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} X(n) e^{\frac{j2\pi kn}{N}}. \quad (\text{A15.2})$$

Las versiones normalmente utilizadas por los programas de cálculo matemático empiezan los índices en 1 en lugar de en 0. Algunos autores ubican el factor  $1/N$  en la transformada directa en lugar de la transformada inversa.

Podemos obtener una relación con el valor eficaz de la señal, dado por

$$X_{\text{ef}} = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} x^2(k)}, \quad (\text{A15.3})$$

mediante el siguiente razonamiento:

$$\begin{aligned} X_{\text{ef}}^2 &= \frac{1}{N} \sum_{k=0}^{N-1} x^2(k) = \frac{1}{N} \sum_{k=0}^{N-1} \left( x(k) \frac{1}{N} \sum_{n=0}^{N-1} X(n) e^{\frac{j2\pi kn}{N}} \right) = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left( \frac{1}{N} \sum_{n=0}^{N-1} x(k) X(n) e^{\frac{j2\pi kn}{N}} \right) = \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \left( \frac{1}{N} \sum_{k=0}^{N-1} x(k) X(n) e^{\frac{j2\pi kn}{N}} \right) = \\ &= \frac{1}{N} \sum_{n=0}^{N-1} X(n) \left( \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{\frac{j2\pi kn}{N}} \right) \end{aligned}$$

Si la señal  $x(n)$  es real, la operación de conjugación de cada término es equivalente a cambiar el signo de las exponenciales complejas, por lo cual

$$X_{\text{ef}}^2 = \frac{1}{N} \sum_{n=0}^{N-1} X(n) \left( \frac{1}{N} \sum_{k=0}^{N-1} x(k) e^{-\frac{j2\pi kn}{N}} \right) = \frac{1}{N^2} \sum_{n=0}^{N-1} X(n) \overline{X(n)}.$$

Resulta,

$$X_{\text{ef}}^2 = \sum_{n=0}^{N-1} \left| \frac{X(n)}{N} \right|^2, \quad (\text{A15.4})$$

denominada igualdad de Parseval.

También es posible establecer ciertas condiciones de simetría conjugada entre los coeficientes de Fourier para el caso de señales reales. Partiendo de la definición de la transformada, conjugamos y operamos

$$\begin{aligned} \overline{X(n)} &= \sum_{k=0}^{N-1} x(k) e^{-\frac{j2\pi kn}{N}} = \sum_{k=0}^{N-1} x(k) e^{-\frac{j2\pi (-n)k}{N}} = \\ &= \sum_{k=0}^{N-1} x(k) e^{-\frac{j2\pi (N-n)k}{N}} \end{aligned}$$

El último paso es posible dado que  $e^{j2\pi K} = 1$  para cualquier  $K$  entero. El último miembro puede interpretarse como la transformada en  $N - n$ , por lo cual:

$$X(N - n) = \overline{X(n)}. \quad (\text{A15.5})$$

Particularmente, esto implica que

$$|X(N - n)| = |X(n)|,$$

por lo que podemos escribir

$$X_{\text{ef}}^2 = \left| \frac{X(0)}{N} \right|^2 + \sum_{n=1}^{N/2-1} 2 \left| \frac{X(n)}{N} \right|^2 + \left| \frac{X(N/2)}{N} \right|^2. \quad (\text{A15.6})$$

Podemos reinterpretar los términos de la sumatoria agrupando en la definición de la transformada inversa los términos conjugados:

$$\begin{aligned} x(k) &= \frac{X(0)}{N} + \sum_{n=1}^{N/2-1} \left( \frac{X(n)}{N} e^{\frac{j2\pi nk}{N}} + \frac{\overline{X(n)}}{N} e^{\frac{j2\pi (N-n)k}{N}} \right) + \\ &+ \frac{X(N/2)}{N} e^{\frac{j2\pi (N/2)k}{N}} \end{aligned}$$

Si expresamos  $X(k)$  como

$$X(n) = |X(n)| e^{j\varphi},$$

los términos entre paréntesis pueden a su vez reescribirse de la siguiente forma:

$$\begin{aligned} \frac{|X(n)|}{N} \left( e^{j\left(\frac{2\pi nk}{N} + \varphi\right)} + e^{-j\left(\frac{2\pi nk}{N} + \varphi\right)} \right) &= \\ &= \frac{2|X(n)|}{N} \cos\left(\frac{2\pi nk}{N} + \varphi\right) \end{aligned}$$

que, reemplazando, conduce a

$$x(k) = \frac{X(0)}{N} + \sum_{n=1}^{N/2-1} \frac{2|X(n)|}{N} \cos\left(\frac{2\pi nk}{N} + \varphi\right) + \frac{X(N/2)}{N} e^{\frac{j2\pi(N/2)k}{N}}.$$

Comparando con (A14.6) vemos que  $2|X(k)/N|^2$  es el valor eficaz al cuadrado de una señal senoidal real. Esto permite introducir el espectro real:

$$X_r(n) = \begin{cases} \frac{X(0)}{N} & n = 0 \\ \frac{2|X(n)|}{N} & n = 1, \dots, N/2-1. \\ \frac{X(N/2)}{N} & n = N/2 \end{cases} \quad (\text{A15.7})$$

El primer término es una componente de continua cuyo valor eficaz coincide con su valor absoluto. Los términos que siguen tienen un valor eficaz igual a la amplitud sobre raíz de 2. El último término ( $n = N/2$ ), pese a estar escrito como exponencial compleja, se trata de una componente que alterna los valores  $\pm X(N/2)/N$  por lo que su valor eficaz coincide con el valor absoluto, que es constante. Este término corresponde a la frecuencia de Nyquist, es decir, la máxima que no genera frecuencias alias.

## Apéndice 16

### Espectro de las ventanas de análisis

Al aplicar las ventanas utilizadas en análisis de espectro (por ejemplo las de Hann y Blackman), dado que multiplican en el dominio tiempo a la señal a enventanar,

$$x_w(t) = w(t) x(t), \quad (\text{A16.1})$$

el espectro de la señal enventanada es la convolución de los espectros de Fourier de la señal y de la ventana:

$$X_w(\omega) = W(\omega) * X(\omega). \quad (\text{A16.2})$$

Para comprender lo que esto implica, si suponemos que la señal es la espectralmente más simple, una exponencial compleja,  $x(t) = e^{j\omega t}$ , cuya transformada de Fourier posee una sola línea espectral no nula, de acuerdo a la definición de convolución el espectro estará formado por el espectro de la ventana  $W(\omega)$  desplazado a la posición de dicha línea espectral. En otras palabras, el espectro original de la señal quedará reemplazado por un espectro que abarca el rango completo de frecuencias, creando una distorsión.

Esto se conoce como *derrame espectral*. La consecuencia del derrame espectral es la aparición de componentes espurias asimilables a un ruido (aunque en realidad constituyen una distorsión de la señal).

Si ahora vamos agregando otras componentes espectrales a la señal, cada una de ellas sumará una copia del espectro de la ventana (ponderada por la amplitud de la componente agregada y desplazada en frecuencia) contribuyendo a incrementar el ruido espectral aportado por las otras.

La ventana ideal, es decir, una que no produjera ninguna alteración del espectro, es aquella en la cual no existan otras componentes espectrales fuera de las aportadas por la señal. Es fácil ver que su espectro debería ser una delta de Dirac, que corresponde en realidad a una ventana constante para todo  $t$ . Tal ventana no serviría en la práctica pues implicaría muestrear desde  $t = -\infty$  hasta  $t = \infty$ , por lo cual en realidad no cumpliría su propósito. No obstante, la forma ideal del espectro de tal ventana nos sirve de orientación acerca de qué sería deseable obtener como espectro de una ventana: algo que concentre la mayor parte de su espectro en frecuencia 0 y que el inevitable derrame espectral decaiga rápidamente hacia las altas frecuencias.

De allí que sea importante conocer el espectro de las ventanas, ya que nos da el criterio de aceptabilidad y nos muestra sus limitaciones. Consideremos la ventana de Hann en tiempo continuo de duración  $T$ . y para simplificar el análisis de espectro la supondremos centrada en 0:

$$w_{\text{hann}}(t) = 0,5 \left( 1 - \cos \frac{2\pi}{T} \left( t - \frac{T}{2} \right) \right) r(t), \quad (\text{A16.3})$$

donde  $r(t)$  es la ventana rectangular entre  $-T/2$  y  $T/2$ . Para calcular el espectro tengamos en cuenta que el coseno puede escribirse como suma de dos exponenciales complejas:

$$w_{\text{hann}}(t) = \left( 0,5 - 0,25 e^{j\frac{2\pi}{T}\left(t - \frac{T}{2}\right)} - 0,25 e^{-j\frac{2\pi}{T}\left(t - \frac{T}{2}\right)} \right) r(t), \quad (\text{A16.4})$$

es decir

$$w_{\text{hann}}(t) = \left( 0,5 + 0,25 e^{j\frac{2\pi}{T}t} + 0,25 e^{-j\frac{2\pi}{T}t} \right) r(t), \quad (\text{A16.4})$$

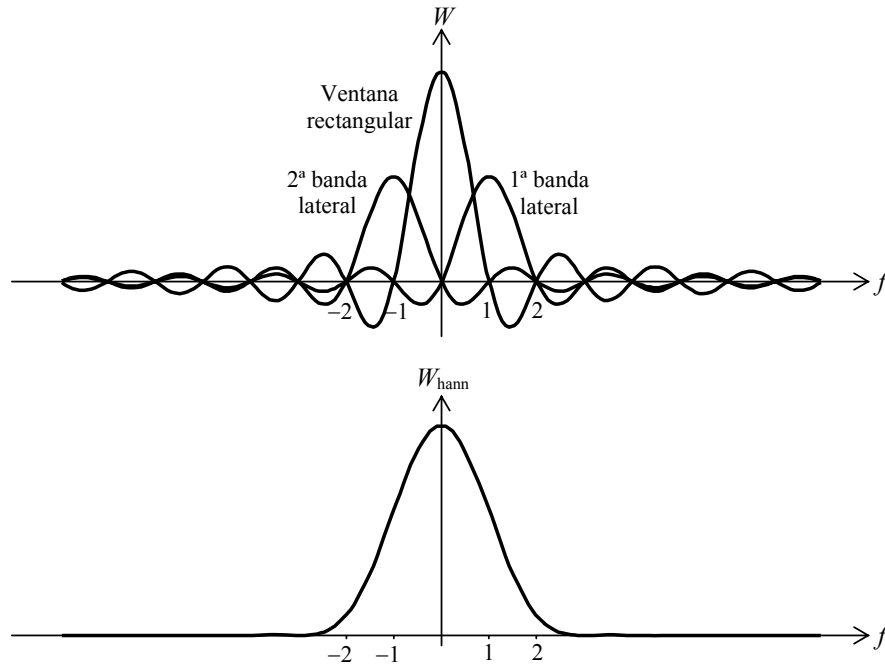
Cada exponencial tiene una única línea espectral por lo que simplemente se traslada el espectro de la ventana rectangular que vale:

$$R(\omega) = T \frac{\text{sen}(\omega T / 2)}{\omega T / 2}. \quad (\text{A16.5})$$

El espectro de la ventana de Hann es, entonces

$$W_{\text{hann}}(\omega) = 0,5R(\omega) + 0,25 R(\omega - 2\pi/T) + 0,25 R(\omega + 2\pi/T), \quad (\text{A16.6})$$

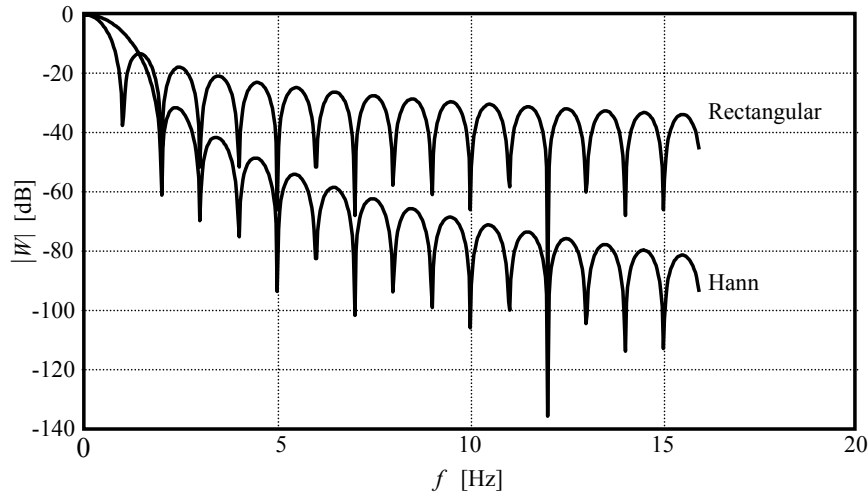
En la figura A16.1 se muestran las tres componentes por separado (donde hemos llamado *bandas laterales* a los espectros desplazados de la ventana rectangular) y la suma de ellas.



**Figura A16.1.** Arriba componentes espectrales que conforman el espectro de la ventana de Hann para  $T = 1$  s. Abajo espectro de la ventana de Hann resultante.

Se ha aprovechado el hecho de que para alta frecuencia la función sinc (es decir  $\text{sen}(\pi x)/\pi x$ ) disminuye lentamente, de modo que un pseudo ciclo difiere poco del siguiente. Por consiguiente, superponiendo dos negativos a uno positivo del doble de amplitud se produce una cancelación casi total.

En la figura A16.2 se muestra una comparación de las ventanas rectangular y de Hann en un diagrama con escala vertical en dB.



**Figura A16.2.** Comparación entre el espectro de las ventanas rectangular y de Hann ( $T = 1$  s).

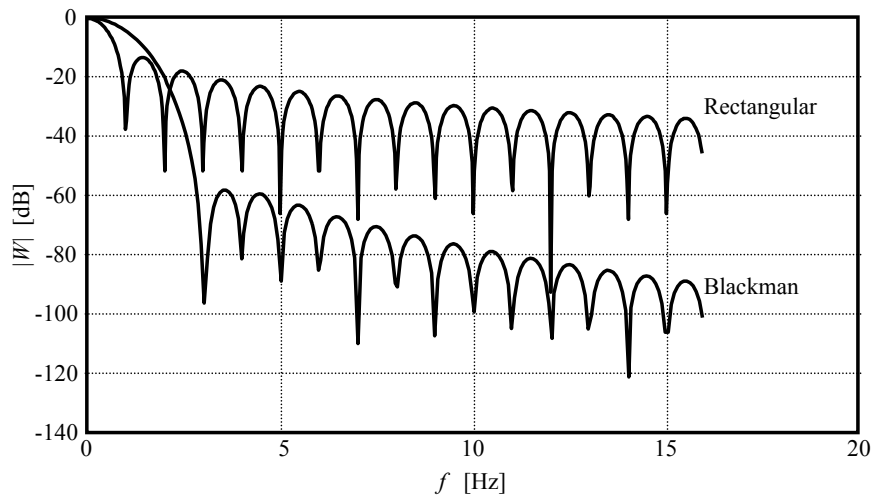
Se observa un ensanchamiento del lóbulo de baja frecuencia (llamado también lóbulo central en el espectro bilateral. Esto implica que resultará más difícil separar componentes tonales muy próximas que en el caso de la ventana rectangular. A cambio de ello el primer lóbulo lateral bajó de  $-13$  dB, correspondiente a la ventana rectangular a  $-33$  dB. De todas maneras en realidad este lóbulo debería compararse con el segundo lóbulo lateral de la rectangular, que es de  $18$  dB.

Con un razonamiento similar se llega al espectro de la ventana de Blackman, expresado en este caso mediante una ventana rectangular y dos pares de bandas laterales:

$$\begin{aligned}
 W_{\text{blackman}}(\omega) = & 0,42R(\omega) + 0,25 R(\omega - 2\pi/T) + 0,25 R(\omega + 2\pi/T) + \\
 & + 0,04 R(\omega - 4\pi/T) + 0,04 R(\omega + 2\pi/T)
 \end{aligned}
 \tag{A16.7}$$

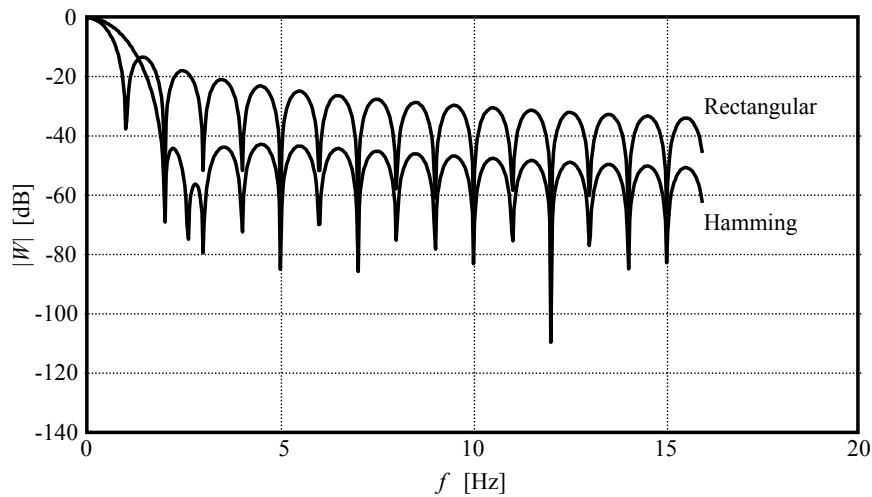
En la figura A16.3 se muestra la comparación entre las ventanas rectangular y de Blackman. Se observa una caída muy importante en frecuencias intermedias con lo cual el ruido de banda ancha causado por el derrame espectral tiene menor importancia que en el caso de la ventana de Hann y, por supuesto, que la ventana rectangular. Así, el primer lóbulo lateral es de  $-58$  dB que corresponde al tercer lóbulo lateral de la ventana rectangular, de  $-21$  dB.





**Figura A16.3.** Comparación entre el espectro de las ventanas rectangular y de Blackman ( $T = 1$  s).

La ventana de Hamming tiene una forma analítica similar a la ventana de Hann, pero a diferencia de ésta, no se propone anular la señal en los extremos de la ventana, razón por la que en alta frecuencia la caída es sumamente lenta, como en la rectangular. La virtud que presenta es que reduce mucho la zona del primer lóbulo lateral, haciendo que el ruido espectral sea menor de  $-40$  dB fuera del lóbulo central, poseyendo una pendiente general que lo torna de tipo browniano, es decir concentrado principalmente en baja frecuencia.



**Figura A16.4.** Comparación entre el espectro de las ventanas rectangular y de Hamming ( $T = 1$  s).

**Ancho de banda equivalente de las ventanas**

Se define el ancho de banda efectivo de ruido de una ventana  $w(t)$  como (Gade et al, 1987a y 1987b)

$$B_{ef} = \frac{\int_{-\infty}^{\infty} |W(2\pi f)|^2 df}{\max |W(2\pi f)|^2} \quad (A16.7)$$

Podemos aplicar el teorema de Parseval (Miyara et al., 2003) para transformar la integral en el dominio espectral en una integral en el dominio temporal:

$$\int_{-\infty}^{\infty} |W(2\pi f)|^2 df = \frac{1}{2\pi} \int_{-\infty}^{\infty} |W(\omega)|^2 d\omega = \int_0^T w^2(t) dt$$

Teniendo en cuenta que el desplazamiento de las bandas laterales es un número entero de veces la frecuencia  $1/T$  (ver figura A16.1 y ecuaciones (A16.6) y (A16.7)), resulta que el máximo se da en el origen y es igual al coeficiente del término independiente. Tenemos, entonces, para una ventana de Hann,

$$B_{ef, Hann} = \frac{\int_0^T w^2(t) df}{0,5^2 T^2} = \frac{1}{T^2} \int_0^T \left(1 - \cos \frac{2\pi}{T} t\right)^2 dt = \frac{3}{2T} \quad (A16.8)$$

es decir,

$$B_{ef, Hann} = \frac{3}{2} \frac{F_s}{N}. \quad (A16.9)$$

Este ensanchamiento con respecto a la resolución de análisis se debe al derrame espectral. Análogamente, para una ventana de Blackman,

$$\begin{aligned} B_{ef, Blackman} &= \frac{\int_0^T w^2(t) df}{0,42^2 T^2} = \\ &= \frac{1}{0,42^2 T^2} \int_0^T \left(0,42 - 0,5 \cos \frac{2\pi}{T} t + 0,08 \cos \frac{4\pi}{T} t\right)^2 dt = \\ &= \frac{1}{T} \left(1 + \frac{1}{2} \left(\frac{0,5}{0,42}\right)^2 + \frac{1}{2} \left(\frac{0,08}{0,42}\right)^2\right) = \frac{1,727}{T} \\ B_{ef, Blackman} &= 1,727 \frac{F_s}{N} \end{aligned} \quad (A16.10)$$

## Índice alfabético

Acoplador acústico	175	B*-Tree	143
Acústica arquitectónica	178	B+ tree	141
Adaptive Transform Audio Coding	135	Banco de filtros	19
ADTL	101	Banda	18
Ajuste de nivel	150	Banda crítica	232
Ajuste por tonalidad	234	Banda de octava	19, 224, 243
Algoritmo de Levinson-Durbin	109	Banda de tercio de octava	19, 210, 224, 243
Aliasing	69, 309	Banda lateral	347
Allocation File	143	Bandas de porcentaje constante	223
Análisis de espectro	17, 158, 160, 206	BD	137
análisis espectral	206	BD-R	137
Análisis sintáctico	326	BD-RE	137
Analizador de ancho de banda constante	17	Beamforming	124, 174
Analizador de espectro	17	Bit	72
Analizador de líneas	17	Bit menos significativo	74
Analizador de octava	17	Blackman, lóbulo central	223
Analizador de porcentaje constante	17	Blackman, ventana	348
Analizador de tercio de octava	17	Blackman, ventana	220, 221, 247
Ancho de banda	19	Blackman-Harris, ventana	220, 221
Ancho de banda constante	206	Blocks	313
Ancho de banda crítica	232	Bloque streaminfo	320
Ancho de banda equivalente	350	Bloques	313
Ancho de banda efectivo	232, 350	Blu-ray Disc	137
Android	142	Buffer	123
Anotaciones	96, 99	Bundle	143
Apreciación	40	Butterworth, filtro	208, 245
Aproximación de Butterworth	242	C, lenguaje	148
Árbol B*	143	Cabezal	119, 120, 121
Árbol B+	141	Calibración	197, 270
Archivo WAV	100	Cámara anecoica	166
archivos malos	142	Campo aleatorio	167
Ardour	148	Campo de presión	166
ASCII	101, 144	Campo difuso	167, 177
Associated data list	101	Campo difuso, corrección	170
ATRAC	135	Campo estacionario	167
Atributos	111	Campo libre	166, 168, 176
Atributos residentes.	141	Campo sonoro	165
au, formato	117	Canales estéreo	110
Audacity	148, 196, 241	Cantidad	262
Audacity, dependencia	151	Carga computacional	236
Audacity, etiquetas	153	Cauer, filtro	208
Audacity, grabación	150	CD	125, 126
Audacity, proyecto	111	CD-DA	126
Audacity, Proyecto	151	CD-R	125, 133, 144
Audacity, Selección	152	CD-ROM	125, 132
Audibilidad tonal	233	CD-RW	126
Audio digital	67	CD-WO	133
Auditoría	141	CD-WR	134
auf, formato	118	CeCILL	196
aup	111	CeCILL, licencia	333
aup, archivo	113, 117, 118	CF	93, 123
aup, formato	111	Cianinas	125
Auricular circumaural	175	Cilindro	121
Auricular HDA 200	175	Cinta de audio digital	138
Auriculares audiométricos	175	CIRC	129, 131

Clusters	138	Chi-cuadrada	283
CLV	126	Chunks	96
Codificador de referencia FLAC	313	DAT	92, 118, 138
Código abierto	148	Data chunk	97
Código de canal de grupo	131	dB	2
Código de Golomb-Rice	110	dBFS	174
Código de Golomb-Rice	105	decibel	2
Código de intercalación cruzada	129	Decorrelación intercanal	110
Código de larga distancia	137	Delta de Dirac	181
Código de longitud variable	104	Densidad de probabilidad	15, 265, 275
Código de predicción lineal	107	Densidad de probabilidad conjunta	48, 266
Código de redundancia cíclica	322	Densidad espectral	22, 226
Código de Reed-Solomon	129	Densidad espectral	35
Código de sincronismo	319	Dependencia, Audacity	151
Código unario	105	Derrame espectral	22, 36, 223
Coefficientes de Fourier	211	Derrame espectral.	218, 346
Colisión	315	Derrame temporal	229
Comentario en estilo Vorbis	316, 317	Descalibración	56
Compact disc	125	Desfragmentación	123, 125
Compact Flash	123	Desvío estándar	14, 266, 274
Complemento a 2	97	Desvío estándar muestral	45, 267, 275, 279
Compresión	95	Desvío estándar poblacional	275
Compresión run-length	107	Detección de errores	325
Condición de Nyquist	69, 72, 309	DFT	211, 213
Condiciones de referencia	43	Diafragma	163
Condiciones de repetibilidad	264	Diezmado	89
Condiciones de reproducibilidad	264	Digital Versatile Disc	137
Confiabilidad	6	Digitalización	72
Conformación de ruido	91	Dirección frontal	168
Constante de Boltzmann	172	Disco compacto	126
Constante de calibración	154	Disco Versátil Digital	137
Contraste de algoritmos	239	Disco Versátil Holográfico	137
Conversión analógica/digital	81	Discos duro	119
Conversión analógico/digital	73	Discos ópticos	125
Conversión digital / analógica	80	Discos rígido	119
Conversor analógico/digital	67, 72	Disk Operating System	139
Conversor de aproximaciones sucesivas	82	Dispersión estadística	14, 286, 287, 288, 291
Conversor flash	81	Distorsión	173
Convolución	190, 234, 346	Distorsión a bajos niveles	75
Convolución circular	234, 235	Distorsión total armónica	95
Convolución lineal	236, 237	Distribución acumulada de frecuencias	272
Cool Edit	147	Distribución de frecuencias	271
Cooley y Tuckey	213, 340	Distribución de Gauss	277
Corrección de errores	67, 129	Distribución de Laplace	104
Corrección por campo difuso	170	Distribución de Rayleigh	293
Correlación	236	Distribución estadística	14, 271
Cota del error	265	Distribución normal	6, 13, 14, 277
Covarianza	48, 266	Distribución <i>t</i> de Student	279
Covarianza muestral	48, 267	Distribución uniforme	278
CRC-16	325	Distribución de Laplace	105
CRC-8	322	Dither	75, 150
Creative Audigy	243	Dither gaussiano	78
Cross Interleaved Reed-Solomon Code	129	Dithering	75
Cuadros	313	DM	87
Cue	100	Document type definition	326
Cue chunk	100	DOS	139
Cue sheet	318	DPCM	84
Cues	96	DSP	206, 342
Chebyshev, filtro	208	DTD	326

DVD	125, 137	Etiquetas	96, 99
DVD-R	137, 144	Etiquetas, Audacity	153
DVD-RW	137, 144	Euler	196
ECMA-094	144	Eventos de edición	111
ECMA-167	144	Exactitud	40, 263
ECMA-377	138	Exportación	111
Ecuación característica	193	Expstudio Audio Editor	148
Ecuación diferencia	185	EXT4	142
Ecuación diferencia	184	Extended File System	142
Ecuación diferencia, solución	191	Extent	139
Edición de audio digital	147	Extents	141, 142, 144
Efecto Kerr	135	Factor de cobertura	47
<i>EFM</i>	125	fase lineal, filtro	227
Electret	165	Fast Fourier Transform	21
Elemento labeltrack	112	FAT	139, 140
Elemento project	112	FAT 32	139
Elemento tags	112	Fenómeno estacionario	13
Elemento timetrack	112	FFT 21, 142, 213, 214, 219, 220, 224, 226, 231,	
Elemento wavetrack	112	232, 237, 340, 342	
Elementos	111	FFT, análisis de espectro	223, 247
Elíptico, filtro	208	FGMOS	124, 125
Encabezamiento	103, 313, 314	FGMOS)	124
Encabezamiento de cuadro	319	File Allocation Table	139
Encabezamiento de subcuadro	322	Filtro A	157
Enmascaramiento espectral	136	Filtro antialias	71, 94, 311
Enmascaramiento temporal	136	Filtro C,	158
Entradas de directorio	139	Filtro de Chebychev	72
Entropía	107	Filtro de fase lineal	227
Envelope	332	Filtro digital	199
Enventanado	22	Filtro elíptico	72
Envolvente	293	Filtro FFT 156, 169, 227, 228, 229, 230, 233,	
Envolvente energética	2, 16, 199	242, 244, 248, 253	
Error aleatorio	265	Filtro FIR	201, 207, 210, 227
<i>Error de escala</i>	58	Filtro IIR	242, 245, 248
Error de linealidad	56, 62, 81	Filtro multibanda	253
Error de medición	265	Filtro pasabanda	208
Error de observación	265	Filtros de bandas críticas	238
Error de paralaje	56	Filtros de Chebychev	208
Error de ráfaga	129	Filtros digitales	19, 207
Error de salva	129	Filtros espejo en cuadratura	136
Error sistemático	56, 265	Filtros, plantillas	239
Error sistemático aditivo	56	Finite impulse response	207
Error sistemático multiplicativo	56, 58	FIR	184, 186, 207, 227
Error sistemático no lineal	301	Firma digital	314
Error sistemático relativo	58	FLAC	103
Errores aritméticos	150	FLAC, estructura de archivo	110
Escritura incremental	144	Flujo de datos	313
Espectro	206	Flujo del tránsito	7, 16
Espectro de banda angosta	232	Flujo vehicular	288
Espectro de bandas	223	Format chunk	97
Espectro de bandas de porcentaje constante	20	Formato PCM	95
Espectro de las ventanas	222	Formato RAW	95
Espectro de líneas	21, 22, 214	Formato RIFF	96
Espectrograma	158	Formato WAV	95, 96, 103
Esperanza matemática	266, 273	Formatos de archivos de sonido	95
Estándar primario	269	Fórmula de reconstrucción	310
Estimador insesgado	275, 283	Fracción de octava	17
Estructura de archivo FLAC	313	Fragmentos	96
Estructura de datos de audio	313	Frames	313

Frecuencia angular natural	27	IEC 61260	210, 239
Frecuencia central	18, 19	IEC 61672	10, 11, 94, 249, 252
Frecuencia de muestreo	68	IFFT	231
Frecuencia de Nyquist	69, 188, 216, 226, 310, 311	Igualdad de Parseval	35, 343, 344
Frecuencias alias	69, 309, 311	IIR	184, 199, 207, 208, 227
Frecuencias nominales	19	Impedancia acústica de radiación	167
Frecuencias normales de vibración	22	Importación	111
Frecuencias subarmónicas	218	Impulso discreto	181
Free Audio Editor	147	Impulso filtrado	182
Free Lossless Audio Compression	103	Incidencia tipo B	46
Free Software Foundation	148	Incertidumbre	1, 39, 44, 54, 254, 267
FSF	196	Incertidumbre combinada	297
Ftalocianinas	125	Incertidumbre combinada muestral	298
Fuentes omnidireccional	178	Incertidumbre en cadena	65
Función de densidad de probabilidad	276	Incertidumbre en condiciones ambientales	50
Función de transferencia	190, 238	Incertidumbre estándar	44, 267
Función de transferencia relativas a la cabeza	178	Incertidumbre estándar combinada	47, 48, 268
Función de transferencia $z$	193	Incertidumbre expandida	47, 269
Función error	15, 291	Incertidumbre tiempo-frecuencia	22
Función error inversa	291	Incertidumbre tipo A	45, 268
Función gamma	280	Incertidumbre tipo B	45, 268
Generación de señales	151	Incertidumbre y resolución	55
General Public License	142, 196	Incidencia frontal	168, 170
Goldwave	147	Índice de directividad	179
Golomb-Rice, código	105	Índice de reducción acústica	54
GPL	142, 148, 196	Inestabilidad	209
GPL, licencia	333	Inferencia	281
GPS	146, 154, 316	Inferencia estadística	271
Grabación zonificada	123	Infinite impulse response	207
Grabación, Audacity	150	Informe de medición	41
Grabador digital	249	Interoperabilidad	146
<i>granularidad</i>	143	Interpolación sinc	310
Guía para la expresión de la incertidumbre	262	Intervalo de confianza	281, 282
GUM	39, 263	IRAM 4113	249
Hamming, ventana	220, 349	IRAM 4113-2	232
Hann, solapamiento	223, 225	ISO 1996	249
Hann, ventana	220, 221, 222, 223, 228, 347	ISO 1996-2	232
Hanning (véase Hann)	220	ISO 7196	11
Hash	314	ISO-GUM	262
HCSD	93	Ji-cuadrada	283
HDA 200	175	Jitter	79
Header	313	Kaiser, ventana	221
HFS+	143	$L_{A90}$	16
Hierarchical file System	143	Label chunk	101
High Performance File System	140	Labeled text chunk	101
Hoja cue	318	Labels	96
Hoja de marcas	318	Labeltrack	112
Holographic Versatile Disc	137	labl	101, 103
HPFS	140	Laser-trimming	81
HRTF	178	Latencia	82
HVD	137	Levinson-Durbin	109
ID3	328, 332	Licencia General Pública	196
Identificador	96	Licencia GPL	148
IDFT	213	Licencia Pública General	142, 333
IEC 60537	10, 11	Linealidad	252
IEC 60651	10, 11	Linear predictive coding	107
IEC 61183	170	Líneas espectrales	232
		Linux	111, 142
		LIST	101

Lista de datos asociados	101	Micrófonos de campo de presión	167
Little-endian	96	Micrófonos de campo difuso	167
Lóbulo lateral	223	Micrófonos de campo libre	167
Lóbulo principal	222, 223	Micrófonos MEMS	174
Lóbulos laterales	222	Minidisc	92, 118, 134
Lógica binaria	72	Moda	273
long distance code	137	Modelización constante	107, 323
$L_p$	2	Modelización polinomial	107, 323
LPC	107	Modelización por predicción lineal	109
LSB	74	Modelización predictiva	107
LTI	183	Modelización verbatim	107, 323
ltx	101	Modelo de la medición	262
Mac OS X	111	Modelo de medición	42
Marcado	67	Modelo LPC	109, 323
Marcas	96, 99, 100	Modelo predictivo	103
Master File Table	141	Modulación de código de pulso	83
Matlab	196, 241	modulación de código de pulsos diferencial	84
Matriz de covarianza	53, 268	Modulación de densidad de pulsos	174
MD	134	modulación delta	87
MD5	314	Modulación ocho a catorce	125
MDCT	136	Modulación ocho a diez.	138
Media	265, 273	modulación sigma-delta	89
Media muestral	266, 279, 281	Monel	165
Mediana	273	MOSFET	124
Medición	262	Movimiento Browniano	172
Medición, método	262	Muestra	271, 274
Medición, modelo	262	Muestreo	67, 68
Medición, principio	262	Muestreo y retención	68
Medición, procedimiento	262	Multímetro digital	254
Mediciones acústicas	1	Music Editor Free	147
Medidas de dispersión	272	NAND, Memoria flash	125
Medidas de tendencia central	272	New Technology File System	140
Membrana	163	Newton-Raphson, método	238
Memoria flash	93, 123	Nivel de presión sonora	2
MEMS	174	Nivel equivalente	3, 4, 51
Mensurando	39, 262	Nivel ponderado A	11, 53
Mensurando, realización	262	Nivel ponderado C	11
Metadatos	99, 313, 314, 316, 320	Nivel sonoro continuo equivalente	3
Método de medición	41, 262	Niveles estadísticos	13
Método directo de medición	41	Niveles ponderados	7
Método indirecto de medición	41	Noise shaping	91
MFT	141	NOR, Memoria flash	125
Micrófono	1, 163	Note chunk	101
Micrófono capacitivo	163	Notes	96
Micrófono de campo difuso	170	NTFS	140, 142
Micrófono de campo libre	168, 170, 171	NTFS-3G	142
Micrófono MEMS, respuesta en frecuencia	174	Numeración binaria	72
Micrófono MEMS, sensibilidad	174	Nyquist, lenguaje	148
Micrófono, amortiguamiento	172	Octave	196, 241
Micrófono, compensación de respuesta	168	Offset	316
Micrófono, diagrama polar	170	Open source	148
Micrófono, distorsión	173	OS X	142, 143
Micrófono, distorsión	173	Oscilograma	117, 149
Micrófono, grilla protectora	168	Overlap-add	229
Micrófono, polarización	165	Oversampling	72
Micrófono, resistencia acústica	172	Palabra de sincronismo	130
Micrófono, respuesta direccional	170, 171	Palabra digital	72
Micrófono, respuesta en frecuencia	168	Parámetro de Rice	106, 324
Micrófono, ruido intrínseco	172	Parseval, teorema	350

Patrón de laboratorio	270	Realización de un mensurando	262
Patrón de medición	269	Realización del mensurando	40
Patrón de referencia	270	Recuantización	89
Patrón de trabajo	270	Red R-2R	80
Patrón primario	269	Reducción de ruido	67, 161
PCM	83	Redundancia	110
PDM	174	Reed-Solomon, código	129
Percentil	13, 291	Régimen permanente	30
Percentiles	291	Régimen transitorio	30
Perfil de ruido	162	Registro de aproximaciones sucesivas	82
Periodo de muestreo	68	Registro digital	67
Pista	121	Relación señal a ruido	74
Pista, Audacity	151	Repetibilidad	264
Pistas	120, 318	Repetición	264
Placa de sonido	241	Réplica	264
Placa interna	163	Reproducibilidad	264
Plantillas	26	residuos	104
Platos	120	Residuos	324
Población	271	Resolución	40, 55, 72, 263
Polarización	165	Resolución espectral	215, 226
Polinomio característico	191	Respaldo	67
Polos	193	Respuesta al impulso	183, 190, 191, 234
Ponderación A	11, 173, 200	Respuesta al impulso finita	207
Ponderación A, B C y D	10	Respuesta al impulso infinita	207
Ponderación A, B, C y D	8	Respuesta en frecuencia	26, 168, 184, 186, 191, 250
Ponderación C	11	Respuesta en frecuencia normalizada	29
Ponderación D	12	Respuesta forzada	30
Ponderación exponencial	2	Respuesta impulsiva	3
Ponderación G	8, 9, 11, 12	Respuesta lenta	3
Ponderación temporal	2, 16	Respuesta libre	30
Ponderación Z	241	Respuesta rápida	3
Porcentaje constante	206	Respuesta transitoria	30, 254, 295
Posicionamiento absoluto	121	Ruido 1/f	173
Posicionamiento relativo	121	Ruido blanco	6, 16, 286
Power Sound Editor	147	Ruido de cuantización	91, 94
Preamplificador	165	Ruido de fondo	13, 16
Preamplificador, distorsión	173	Ruido de tránsito	288
Precisión	40, 263	Ruido estacionario	286
Pre-eco	229	Ruido interno	252
Presión de referencia	2, 3	Ruido no estacionario	288
Presión sonora	1	Ruido rosa	173
Principio de medición	41, 262	Ruido térmico	6, 172
Probabilidad	275	Ruido tonal	231
Probabilidad de cobertura	47	Ruido, micrófono	173
procedimiento de medición	40	Sample and hold	68
Procedimiento de medición	41, 262	Scilab	142, 196, 197, 241, 333
Producto de convolución	182	SD	93, 123
Productos de intermodulación	77	SDHC	147
Project	111	SDHC	123
Promedio	273	SDM	89
Promedio energético	197	Sector	121
Protocolo de análisis	154	Sectores	138
Protocolo de medición	147	Secure Digital	123
Proyecto	111	Seek table	316
Proyecto, Audacity	151	Selección	152
proyectos en Audacity	111	Selección, Audacity	152
Pulso unitario	181	Sensibilidad	154
Puntos de búsqueda	316	Sensibilidad de referencia	176
Puntos índice	318		



Sensibilidad del micrófono	197	Transformación bilineal	195
Sensibilidad del transductor	98	Transformada coseno discreta modificada	136
Señal aleatoria	22	Transformada discreta de Fourier	136, 211, 213, 340
Señal de calibración	98, 156	Transformada Discreta de Fourier	343
Señal digital	67	Transformada discreta de Fourier inversa	213
Señal discreta	72, 181	Transformada rápida de Fourier	213, 340
Señal muestreada	308	Transformada Rápida de Fourier	21, 342
Series de Renard	19, 270	Transformada $z$	188
Servocontrol	121	Transformada $z$ de un retardo	189
Sistema discreto, respuesta en frecuencia	184, 186	Transformada $z$ inversa	193
Sistema estable	193	Transistor de efecto de campo	124
Sistema invariante en el tiempo	183	Traverso DAW	148
<i>Sistema lineal</i>	183	Trazabilidad	269
Sistemas de archivos	138	Tubo de impedancia	167
Sistemas discretos	183	Tubo de Kundt	167
sistemático no lineal	62	UDF	143
Slack	139	Umbral auditivo	2
Sobremuestreo	72, 89, 210	Umbral de enmascaramiento	233
Software libre	103, 148	Unicode	140, 143, 144, 321
Software matricial	142	Unidad de disco	119
Solapado y suma	229	Unidades de asignación	138
Solución homogénea	191, 193	Universal Disk Format	143
Solución particular	192	USB	119, 123
Sound Engine	148	UTF-8	317, 321
Sparse files	142	Valor eficaz	1, 197, 286, 288
Streaminfo	313	Valor esperado	265
Sub-bandas	135	Valor medio	273
Subblocks	313	Valor medio muestral	274
Subbloques	313	Valor medio poblacional	274
Subcódigo	130	Valor verdadero	39, 262
Subcuadros	313	Variable aleatoria	13, 271
Subframes	313	Variable aleatoria continua	275
Submuestreo	89, 209	Variable aleatoria discreta	275
Sumario	117, 149	Varianza	266, 274
Tabla de asignación virtual	144	Varianza muestral	283, 284
Tabla de búsqueda	316	Varianza poblacional	284
Tabla maestra de archivos	141	Velocidad lineal	126
Tags	112	Ventana de Blackman	247, 348
Tarjeta SD	139	Ventana de Hamming	349
Tarjeta SDHC	147	Ventana filtrante	230
Tasa de muestreo	7, 14, 15, 68, 307, 311, 314	Ventanas	220, 346
Tasa de muestreo de un proyecto	111	Ventanas, espectro	222
Temperatura de Curie	134	Ventanas, lóbulo principal	222
Teorema de muestreo	216	Ventanas, lóbulos laterales	222
Teorema del muestreo	69, 307, 309, 312	Virtual allocation table	144
Tercio de octava	19	Wave Editor	147
THD	253	Wave Pad	147
Tiempo de estabilización	5, 6, 7, 15, 16, 17, 290	Waveblock	114, 330
Tiempo de integración	286, 288	Waveclip	112, 329
Tiempo de reverberación	54	Wavesurfer	148
Tiempo discreto	67	Wavetrack	112, 329
Timetrack	112	Wavosaur	147
Tolerancia	26, 263	Wavread	196, 198
Tono de calibración	100, 154	Windows	111
Tracks	318	XML	111, 113, 154, 326
Trama	129	ZBR	123
Transductor	58	Zero-padding	210, 229
Transductores	163	Zoom H4	249



La Asociación de Acústicos Argentinos, AdAA, fue fundada por voluntad de un grupo de acústicos en 1976. Al año siguiente obtiene la personería jurídica como sociedad civil sin fines de lucro por Resolución 1704 de la Inspección General de Personas Jurídicas. En 1977 la AdAA es reconocida por la International Commission of Acoustics (ICA),

dependiente de la UNESCO, como representante de Argentina y a partir de 1996 es miembro fundador de la Federación Iberoamericana de Acústica (FIA) que congrega a las asociaciones de la Argentina, Brasil, Chile, España, México, Venezuela, Uruguay, Perú y Portugal. La AdAA mantiene contactos permanentes con Sociedades de Acústica de todo el mundo, con particular énfasis en América Latina y la península Ibérica, propiciando y favoreciendo el contacto entre laboratorios universitarios e investigadores argentinos y extranjeros. Abarca, entre otros, los campos de la Acústica, la Electroacústica, la Ingeniería de Audio y de Sonido, el Procesamiento de Señales, la Metrología Acústica, el Ruido Urbano y el Control de Ruido y Vibraciones. A lo largo de su actividad la Asociación ha sido organismo consultor para distintos proyectos legislativos, tanto nacionales como provinciales. La finalidad de la AdAA es favorecer el desarrollo de la Acústica y sus disciplinas asociadas. Para ello desde hace más de 30 años realiza congresos, reuniones y jornadas científicas y técnicas, promoviendo el intercambio de información con el exterior y el contacto entre estudiantes y profesionales en estas áreas del conocimiento. En 2010 ha sido designada como organizadora principal del Congreso Internacional de Acústica ICA 2016, a realizarse en Buenos Aires, en 2016. Dentro de su labor de difusión, recientemente ha comenzado a publicar y auspiciar la publicación de libros de temas relativos a la Acústica, organizando la primera feria del libro de Acústica de autores argentinos en la Semana del Sonido realizada en La Plata, en agosto de 2013.

Este edición de 200 ejemplares se terminó de imprimir el 17 de agosto de 2013 en los talleres gráficos de la Imprenta Editorial Magenta, Av. Pellegrini 358, Rosario, Argentina.

