

INFORME 953-2*

**CODIFICACIÓN DIGITAL PARA LA TRANSMISIÓN DE SEÑALES DE SONIDO
DE ALTA CALIDAD EN LA RADIODIFUSIÓN POR SATÉLITE
(ANCHURA DE BANDA NOMINAL DE 15 kHz)**

(Cuestión 2/10 y 11, Programas de Estudios 51C/10, 2F/10 y 11)

(1982-1986-1990)

1. Introducción

Los progresos tecnológicos permiten en la actualidad plantear la transmisión de señales de sonido codificadas digitalmente dirigidas al público. En particular se podrá introducir esta técnica en el servicio de radiodifusión por satélite. Ello exigirá la definición previa de las características de codificación, multiplexaje y modulación de las señales de sonido. El presente Informe se refiere sólo a la codificación de sonido digital; el multiplexaje, la estrategia para la protección de errores en los bits y los aspectos de la modulación se tratan, respectivamente, en los Informes 954 y 632. En el Informe 1073 figura información sobre los métodos de corrección u ocultación de errores.

En el Programa de Estudios 51C/10 se señala que la introducción de técnicas digitales en la transmisión del sonido permitirá mejorar la calidad de las señales transmitidas. La Recomendación 651 trata de la codificación MIC digital. En el Anexo II al presente Informe se dan los resultados de pruebas objetivas y subjetivas.

En el Informe 1119 se describen otros métodos de codificación con los que se obtiene una reducción muy sustancial de la velocidad binaria manteniendo al mismo tiempo un grado de calidad muy elevado.

Además de la mejora de la calidad, debe tenerse en cuenta lo siguiente:

- las fórmulas de transacción entre objetivos de calidad y de velocidad binaria pueden ser diferentes tratándose de servicios radiofónicos cuyas exigencias de calidad y limitaciones de planificación sean distintas, y pueden variar también en función de las necesidades de cada país;
- la utilización de una sola norma para cada aplicación es claramente ventajosa tanto para las entidades de radiodifusión y fabricantes de receptores como para el público.

2. Codificación MIC digital**2.1 Codificación lineal**

La señal audio se codifica en forma digital mediante un convertidor analógico/digital de alta precisión tras pasar por un filtro antipliegues adecuado y quizás por una red de preacentuación. La codificación uniforme resultante de este proceso conduce a una representación lineal con un mínimo de 14 bits por muestra. En este caso, se utiliza para cada muestra una codificación de complemento de 2. Cabe asociar un posible factor de escala a bloques de muestras sucesivas.

2.2 Codificación casi instantánea o de coma flotante

Una señal de sonido codificada linealmente de 16 bits se puede transmitir a 14 bits por muestra utilizando un sistema de codificación de coma flotante en el que el factor de escala se basa en 64 muestras consecutivas. Se puede utilizar o no preacentuación.

Cuando se necesita una velocidad binaria menor por canal audio, la compresión-expansión casi instantánea permite reducir el número de bits por muestra de 14 bits a 10 bits. Esta compresión-expansión se aplica a bloques de 32 muestras sucesivas con un factor de escala complementario en cinco segmentos. Se sigue utilizando para cada muestra la representación de complemento de 2. En este caso se recomienda el empleo de preacentuación para reducir el ruido modulado por el programa.

* Este Informe debe señalarse también a la atención de la CMTT.

2.3 Codificación diferencial

La codificación diferencial es otra técnica que permite conseguir una velocidad de transmisión de datos más baja por canal audio, aprovechando el hecho de que en la mayor parte del material radiofónico, la correlación entre dos muestras consecutivas es grande. En este caso, la diferencia entre muestras sucesivas será por lo general más pequeña que los propios valores de las muestras. Codificando la señal diferencial, en vez de las muestras originales, puede conseguirse una reducción de la velocidad de transmisión de datos para una determinada condición en materia de relación señal/ruido.

Las muestras diferenciales pueden codificarse instantáneamente, una a una, como en un codificador lineal, o por bloques de muestras sucesivas (con factores de escala complementarios), como en la compresión-expansión casi instantánea. Se ha propuesto [CCIR, 1986-90] utilizar esta combinación de codificación diferencial y compresión-expansión casi instantánea para reducir de 15 a 8 el número de bits por muestra. Se recomienda emplear preacentuación a fin de reducir el ruido radiofónico modulado.

2.4 *Acentuación MIC*

La calidad del sistema de transmisión, en particular las características de ruido y distorsión, dependen mucho de las estadísticas de la señal y de la ley de codificación utilizada. Además, puede considerarse que la respuesta de detección del ruido audible del sistema receptor, incluido el oído humano, no es uniforme. Teniendo en cuenta esos factores, puede mejorarse la característica adoptando una preacentuación adecuada antes del proceso de codificación, y la correspondiente desacentuación en el receptor después del proceso de decodificación. Para los sistemas MIC se han estudiado y evaluado exhaustivamente dos métodos de acentuación. El anexo I contiene información más detallada de los dos sistemas, uno basado en la Recomendación J.17 del CCITT y el otro basado en 50/15 μ s.

3. **Modulación delta adaptable (MDA)**

3.1 *Antecedentes*

El sistema de modulación delta adaptable (MDA) se basa en la modulación delta (el decodificador se define matemáticamente en el punto 3.2). La señal de audio se codifica en forma digital con un modulador delta simple, después de haber sido modificada por dos procesos:

- se hace pasar la señal de audio por una red de preacentuación variable que altera el espectro de audio, y
- se comprime el nivel de la señal de audio, según su pendiente. La compresión es «infinita» por encima del nivel de señal de entrada de 48 dB, o sea que la señal se comprime al mismo nivel para la conversión a forma digital. Se obtiene una carga óptima del canal digital para la mayoría de las señales audio.

El tren digital de bits que representa la señal audio modificada se transmite al receptor, junto con dos trenes de bits de muy baja velocidad binaria que contienen la información de señal de control, por lo que el decodificador puede invertir exactamente el tratamiento efectuado por el codificador. El decodificador recupera la señal de audio mediante la integración del tren de bits correspondiente, efectuando una expansión de gama dinámica basada en la señal de control de la compresión (pendiente), y desempeña una desacentuación espectral basada en la información contenida en la señal de control de acentuación.

Cabe observar los siguientes puntos en relación con la utilización de la modulación delta:

- a) la velocidad binaria de la modulación delta se puede reducir sin gran efecto en la calidad de la señal audio (la relación S/N se degrada en 9 dB cuando la velocidad binaria se reduce a la mitad);
- b) los errores reproducidos son tolerables;

- c) un convertidor digital a analógico para la modulación delta es muy sencillo y no exige componentes de precisión;
- d) la desacentuación ofrece un filtrado de salida adecuado y no se necesitan filtros de paso bajo de corte rápido.

Los puntos c) y d) influyen en los circuitos de decodificación necesarios.

3.2 Definición del decodificador MDA (fig.1)

3.2.1 Decodificador de audio

El decodificador de audio consiste en un integrador permeable alimentado con impulsos derivados del tren de bits de datos de audio, de modo que los unos y ceros de los datos, a una velocidad del orden de 250 kbit/s, hacen variar la salida positiva o negativamente por pasos iguales. El tamaño de los impulsos es linealmente proporcional a una señal de control aplicada, en una gama de unos 50 dB, y el integrador tiene una respuesta de frecuencia definida por:

$$(1 + sT_0)^{-1}$$

donde $T_0 = 0,5$ ms.

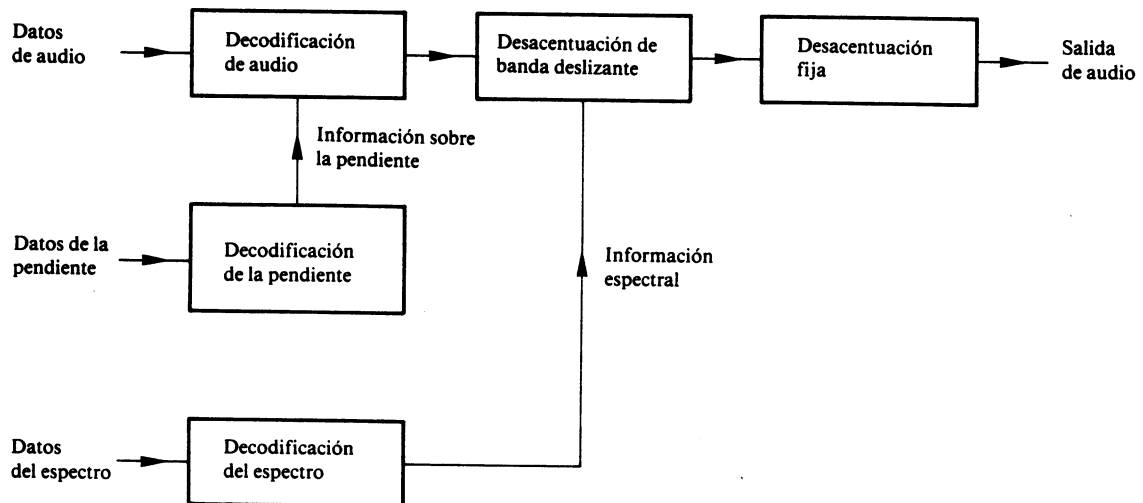


FIGURA 1 – Decodificador de modulación delta adaptable

3.2.2 Desacentuación de banda deslizando (fig. 2)

La desacentuación tiene la respuesta de frecuencia variable,

$$\left[\frac{10sT_1}{1+sT_1} + \frac{1+sT_2}{1+sT_3} \right]^{-1}$$

donde $T_2 = 5 \mu\text{s}$, $T_3 = 50 \mu\text{s}$ y T_1 es variable bajo el control de los datos de espectro.

Adviértase que el primer término dentro del corchete es una característica de paso alto a la frecuencia f_1 , con

$$f_1 = \frac{1}{2nT_1}$$

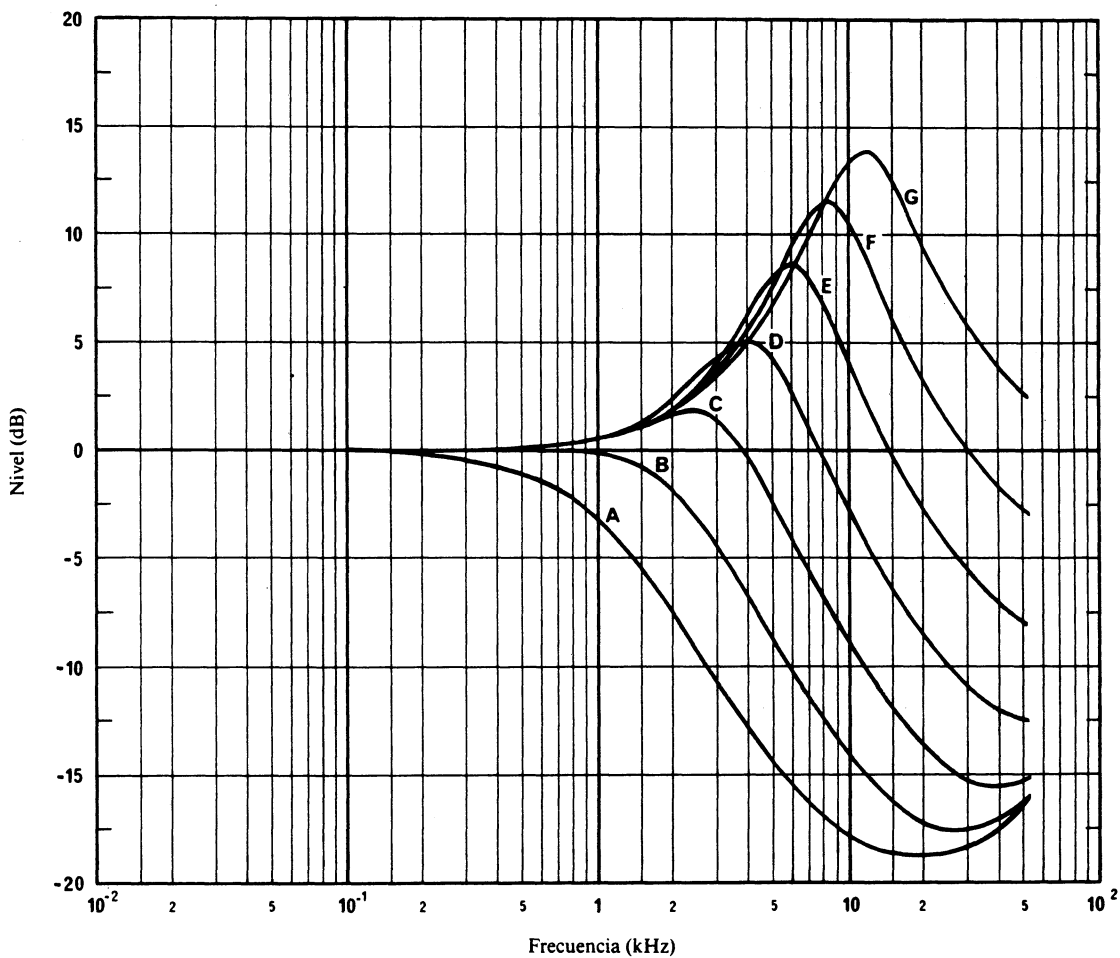


FIGURA 2 – Filtro de desacentuación de modulación delta adaptable (MDA)

3.2.3 Decodificador de espectro

La información de control de desacentuación está contenida en un tren de bits de datos a unos 8 kbit/s, que es típicamente, pero no necesariamente, un submúltiplo entero de la velocidad binaria de los datos de audio. El nivel medio V_m del tren de bits se obtiene alimentando los datos por un filtro de paso bajo de tres polos con la característica:

$$(1 + sT_4)^{-3}$$

donde $T_4 = 2$ ms.

Si la altura de los impulsos de datos es V_p , un parámetro de control x viene dado por:

$$x = \frac{V_m}{V_p}$$

La salida del filtro alimenta un potenciador para producir una señal de control que actúa sobre la acentuación variable. Las constantes del circuito se fijan de modo que proporcionen la relación:

$$f_1 = \frac{1}{2nT_1} = 4000 (2^{10x}) \quad \text{Hz}$$

Con esta definición, un cambio de 0,1 en x desplaza f_1 una octava.

3.2.4 Decodificador de pendiente

La señal de información de pendiente está contenida en otro tren de bits de datos a 8 kbit/s aproximadamente. Este tren de bits se convierte en una señal de control por medio de un filtro de paso bajo y una potenciación, exactamente igual que en el decodificador de espectro. La altura de los impulsos integrada en el decodificador de audio es proporcional a esta señal de control.

Como en el decodificador de espectro, el nivel medio V_m del tren de bits se calcula utilizando un filtro de paso bajo de tres polos con la característica:

$$(1 + sT_4)^{-3}$$

donde $T_4 = 2$ ms,

y se alimenta en un potenciador que suministra la señal V_{ss} de control del tamaño del paso o del impulso destinada a utilizarse en el decodificador de audio. Este potenciador tiene la característica:

$$V_{ss} = V_0 (2^{10y})$$

donde y es el nivel medio normalizado de los impulsos (como x en el caso anterior) y V_0 es un factor de proporcionalidad adaptado al decodificador de audio. Con esta definición, la altura de los impulsos varía en 6 dB por cada variación de 0,1 en y .

3.2.5 Desacentuación fija

El dispositivo de desacentuación fija es un filtro de paso bajo de un solo polo con la característica:

$$(1 + sT_5)^{-1}$$

donde $T_5 = 25$ μ s.

En los decodificadores de los usuarios no se necesita ningún filtrado ulterior de la salida.

REFERENCIAS BIBLIOGRÁFICAS

Documentos del CCIR:

[1986-90]: a. 10/52 (Japón).

BIBLIOGRAFÍA

- FIELDER, L. D. [septiembre de 1985] Pre- and post-emphasis techniques for audio recorders. *J. Audio Eng. Soc.*, Vol. 33, 9.
- FIELDER, L. D. [octubre de 1985] Audible modulation noise in floating point conversion systems. *J. Audio Eng. Soc.*, Vol. 33, 10.
- FIELDER, L. D. [octubre de 1985] Assessing the low level sound quality of digital audio systems. Presentada en el 19th Convention of the Audio Engineering Society, preimpresión N.º 2268.
- TODD, C. y GUNDRY, K. [marzo de 1984] A digital audio system for broadcast and prerecorded media. Audio Engineering Society, preimpresión N.º 2071, París, Francia.
- TODD, C. y GUNDRY, K. [1984] A digital audio system coding method for DBS, cable, and terrestrial broadcasting. Conference Publication 240, Ninth International Broadcasting Convention (IBC 84), Brighton, Reino Unido.

ANEXO I

TÉCNICAS DE ACENTUACIÓN PARA SEÑALES SONORAS MIC DE GRAN CALIDAD

En la técnica de codificación MIC digital del sonido han de utilizarse los métodos de preacentuación y desacentuación para diversos fines:

- en el caso del tratamiento de la señal de sonido digital, la acentuación disminuye la perceptibilidad subjetiva del ruido de cuantificación, en particular, en el caso del llamado ruido modulado por programa con codificación de compresión-expansión;
- en el caso del tratamiento de las señales de sonido por compresión-expansión y codificación lineal, la acentuación puede disminuir la degradación debida a errores en los bits en valores bajos de C/N .

Se han propuesto dos características distintas para ambos sistemas, de compresión-expansión y de codificación lineal. El primer sistema de preacentuación se basa en la Recomendación J.17 del CCITT (con una pérdida de inserción de 6,5 dB a 0,8 kHz), y el segundo en la característica de constante de tiempo 50/15 μ s. La fig. 1 de la Recomendación 651 muestra las características de ambos sistemas.

Los resultados de las investigaciones sobre codificación lineal [CCIR, 1982-86a] y sobre codificación lineal y con compresión-expansión [CCIR, 1982-86b y c] indican claramente las ventajas de utilizar preacentuación en la mayoría de los casos. Se ha observado que ambos sistemas ofrecen la ventaja de reducir el ruido de cuantificación, con una ligera ventaja en favor de la preacentuación 50/15 μ s en el caso del funcionamiento sin compresión-expansión. Respecto al riesgo de sobrecarga, los documentos [CCIR, 1982-86a, b y c] reflejan puntos de vista distintos. En cuanto a las degradaciones debidas a los errores binarios para niveles bajos de la relación C/N , no se han observado diferencias claras entre estos dos sistemas.

Los resultados de las evaluaciones teóricas [CCIR, 1982-86b] y de las pruebas subjetivas [CCIR, 1982-86d y e] efectuadas para sistemas con compresión-expansión, basándose en las características de preacentuación que indica la fig. 1 de la Recomendación 651, muestran que podría lograrse una reducción significativa de la capacidad subjetiva de audición del ruido modulado por el programa aplicando la preacentuación de la Recomendación J.17 del CCITT. No obstante, los resultados de otras pruebas subjetivas [CCIR, 1982-86f y g] no mostraron una diferencia significativa. En [CCIR, 1982-86f] se efectuó un ajuste de las pérdidas de inserción en cada sistema para evitar la sobrecarga.

Por lo que se refiere a la codificación digital de señales sonoras, se utiliza ampliamente la preacentuación de la Recomendación J.17 del CCITT (con una pérdida de inserción de 6,5 dB a 0,8 kHz) para fines de transmisión, tal como especifica la Recomendación J.41 del CCITT, mientras que el sistema de preacentuación 50/15 μ s se utiliza ampliamente en aplicaciones de usuario.

REFERENCIAS BIBLIOGRÁFICAS

Documentos del CCIR

- [1982-86]: a. 10/21 (UER); b. 10-11S/201 (UER); c. 10-11S/139 (Japón); d. 10/269 (Alemania (República Federal de)); e. 10-11S/206 (Francia); f. 10-11S/205 (Japón); g. 10-11S/207 (Canadá).

ANEXO II

**Resultados de pruebas subjetivas sobre modulación delta y otros
métodos de codificación para sonido de alta calidad**

1. Mediciones subjetivas efectuadas en Australia

La «Australian Broadcasting Corporation» (AuBC) ha realizado pruebas subjetivas del sistema de Modulación Delta Adaptable (MDA) [AuBC, 1985], de conformidad con la Recomendación 562 para relacionar la calidad audio subjetiva con la proporción de bits erróneos, con miras, en última instancia, a relacionar todo ello con la relación C/N .

El equipo MDA tenía una tarjeta para generación de errores aleatorios, que introducía errores en el tren de datos en una proporción conocida y determinable según la posición del conmutador. Se comprobó la exactitud de las posiciones del conmutador, con resultados positivos dentro de las variaciones estadísticas normales.

Se utilizó un entorno cuidadosamente controlado y medido para reproducir el material en secuencias conforme a la Recomendación 562.

Para evaluar la parte audio del sistema se eligieron siete secuencias de material procedentes de un disco compacto o de grabaciones originales MIC de 16 bits, que fueron grabados consecutivamente como MIC de 16 bits en un magnetoscopio, para reproducirlos posteriormente. En el cuadro I se enumeran las 7 secuencias de programa.

CUADRO I – *Secuencias de programa de referencia*

Material (y curva) N.º	Descripción
1	Sólo voz masculina
2	Palabra hablada masculina
3	Piano
4	Flauta y coro infantil
5	Palabra hablada femenina
6	Orquesta
7	Grupo moderno mixto

En la fig. 1 se indica que, para una BER dada, las degradaciones varían considerable a la vez que uniformemente, con el material del programa, pero que se obtiene una nota de degradación media de 3,5 para una BER de 10^{-3} aproximadamente.

Se pensó que una nota global de degradación de 3,5 para una BER de 10^{-3} equivalía a una relación C/N inferior a 9 dB. Sin embargo, esta última se degradaría al utilizar cuatro niveles de datos, como se proyecta hacer en Australia. No pudieron realizarse mediciones en este modo de operación, pero los resultados indican una degradación de 1 dB aproximadamente para los niveles exteriores de datos y de 1,8 dB aproximadamente para los niveles interiores. Esto puede influir en la elección de los programas que pueden transmitirse por los diferentes canales de datos. Sin embargo, la AuBC no ha podido disponer de facilidades que confirmen estos valores.

Todavía no se ha establecido una correlación independiente entre los valores de C/N y BER que se obtienen con receptores en condiciones reales.

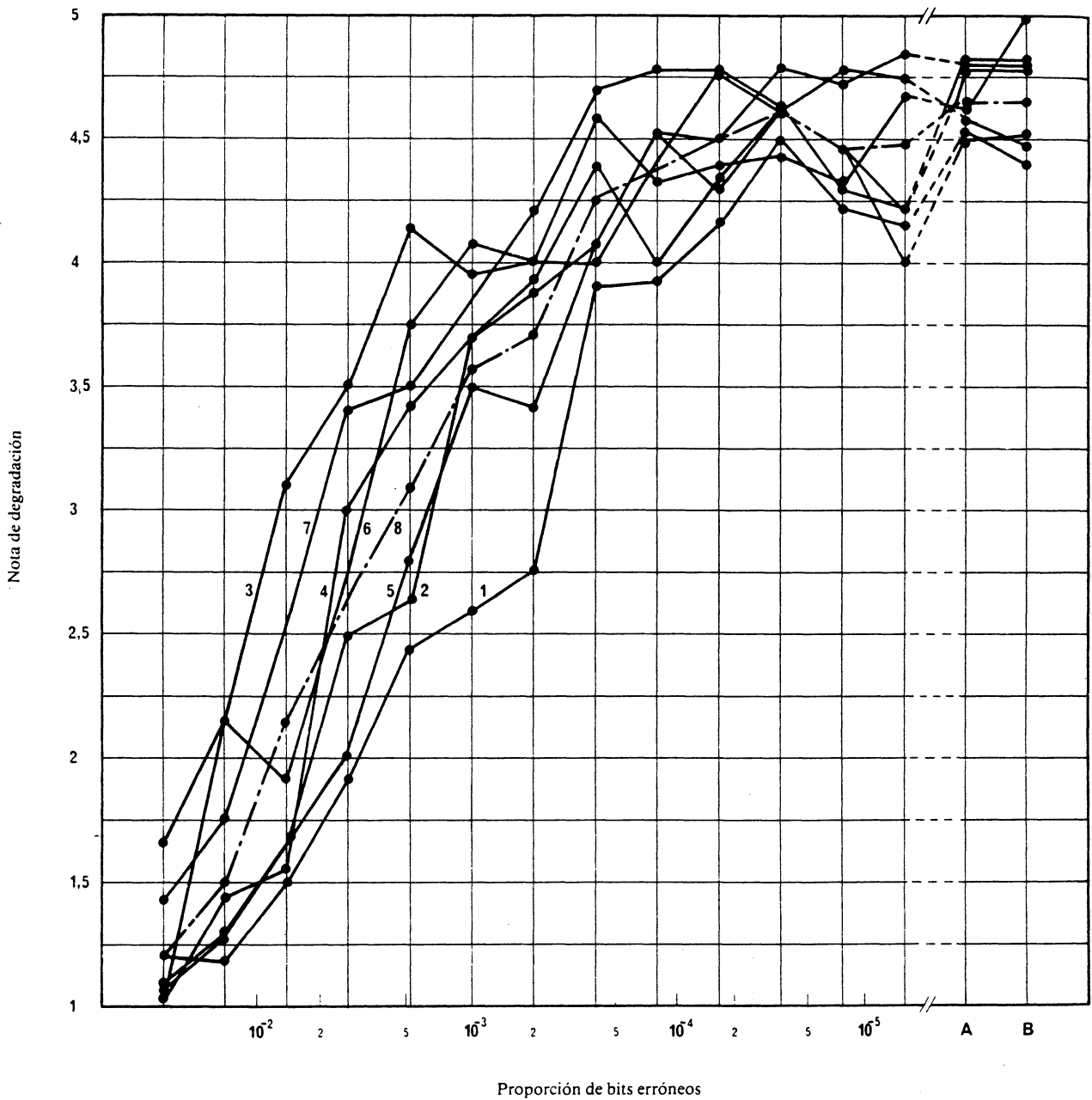


FIGURA 1 - Nota de degradación en función de la posición del conmutador para diferentes materiales de programa

Nota 1. - Las curvas 1-7 corresponden a las selecciones de programa indicadas en el cuadro I. La curva 8 indica el resultado global (media).

Nota 2. - En la posición A de la escala, la señal se transmitió por el sistema Dolby MDA sin errores.

Nota 3. - En la posición B de la escala, la señal proviene directamente de la fuente analógica (por ejemplo MIC en magnetoscopio).



2. Mediciones subjetivas efectuadas en Canadá

2.1 Evaluación subjetiva de una selección de códigos MDA, lineales y casi instantánea

Con objeto de identificar las ventajas relativas de los diversos sistemas de codificación propuestos, en Canadá se llevó a cabo un programa de pruebas.

El programa de pruebas abarcaba los siguientes sistemas de codificación:

- expansión * semiinstantánea 10-14 (NICAM 3) con preacentuación dada en la Recomendación J.17 del CCITT;
- expansión instantánea 10-14 (ley A) con preacentuación dada en la Recomendación J.17 del CCITT;
- codificación lineal de 14 bits con preacentuación dada en la Recomendación J.17 del CCITT;
- modulación delta adaptable (velocidad de muestreo = 330 kHz); y
- modulación delta adaptable (velocidad de muestreo = 204 kHz).

Los tres primeros sistemas de codificación se modelaron en un computador que funcionaba conjuntamente con un procesador vectorial.

La preacentuación aplicada a las señales se modificó por filtrado digital. La frecuencia de muestreo se cambió también a 32 kHz por proceso digital. Luego se procedió a la expansión y al truncado para generar el número adecuado de bits y obtener por último el tren digital compandido. Se efectuó la reformatización adecuada para almacenar la secuencia de prueba en un magnetoscopio.

Este proceso no se utilizó para el sistema MDA que no se podía modelar en el computador. Antes del proceso por MDA era necesaria la conversión a señal analógica. En la salida de la MDA, se reconvertía la señal analógica en 16 bits/muestra para inserción en la secuencia de prueba grabada.

El material utilizado era leído de discos láser compactos o generado por el propio computador:

- Cardozo, Pájaro Campana (Arpa india);
- Orff, Carmina Burana (canto);
- Gong sintético generado a 44 kHz simulando el sonido de un triángulo en las cuatro primeras notas de «Frère Jacques» (4,2 kHz, 4,7 kHz, 5,3 kHz y 4,2 kHz) (tiempo de ataque: $T = 3$ ms, caída: $T = 300$ ms);
- Gong sintético generado a 44 kHz simulando una guitarra baja en las cuatro primeras notas de «Frère Jacques» (65,4 Hz, 73,4 Hz, 82,4 Hz, y 65,4 Hz) (tiempo de ataque: $T = 3$ ms, caída: $T = 300$ ms).

Las pruebas se efectuaron según el método de comparación propuesto por el CCIR (Recomendación 562) en que la secuencia A-B se presenta dos veces y se dejan luego 15 s para la puntuación, de acuerdo también con la escala de comparación del CCIR. Se efectuaron mediciones audiométricas de cada sujeto antes de las pruebas subjetivas. Por término medio, se pusieron 3 oyentes en una sala de escucha bien calibrada. Participaron en las pruebas 25 oyentes en total.

Se repitieron todas las comparaciones en orden diverso para obtener una medida de la coherencia de los oyentes. Así, se eliminaron 4 oyentes del análisis de los resultados. Además, cuando dos notas del mismo oyente para la misma comparación diferían en más de dos grados, se descartaron.

A causa de vibraciones y matraqueos incontrolables provocados en la sala de escucha por el gong sintetizado de baja frecuencia y alto nivel, la secuencia de prueba fue también evaluada por 8 oyentes provistos de auriculares. Los resultados de este grupo se utilizan para el gong de baja frecuencia.

Además de la nota de opinión media y de la desviación típica obtenida a partir de los datos filtrados, se procedió a un estudio de la importancia estadística de los resultados utilizando la distribución de Student-t de una cola. Se obtuvo el margen de confianza en que, por término medio, la población normal preferirá el códec** «A» al códec «B».

Los resultados se resumen en el cuadro II. La nota media de opinión expresada en la escala -3, 0, +3, la desviación típica y el nivel de confianza en que «el códec A es mejor que el códec B» figuran en cada caso. En el cuadro, la expansión casi instantánea 10-14 según el Informe 953 se denomina NICAM 3.

* «codec» = «codificador-decodificador».

CUADRO II - Resultados de las pruebas subjetivas efectuadas en Canadá

Codec «A»	Codec «B»	Arpa	Canto	Gong alta frecuencia	Gong baja frecuencia
		$\bar{X} / S / \% \text{ Conf}$	$\bar{X} / S / \% \text{ Conf}$	$\bar{X} / S / \% \text{ Conf}$	$\bar{X} / S / \% \text{ Conf}$
NICAM 3, Rec. J.17 del CCITT	MDA (330 kHz)	0,1 0,7 88%	0,2 0,9 91%	-0,1 1,3 28%	1,1 0,8 99,95%
NICAM 3, Rec. J.17 del CCITT	MDA (204 kHz)	0,0 0,8 50%	0,2 0,9 92%	1,8 1,0 99,999%	1,4 0,5 99,995%
MDA (330 kHz)	Ley A, Rec. J.17 del CCITT	0,0 0,6 50%	-0,2 0,8 6%	0,9 1,4 99,99%	-1,2 0,5 0,01%
NICAM 3, Rec. J.17 del CCITT	Ley A, Rec. J.17 del CCITT	0,0 0,7 50%	-0,1 1,0 27%	0,8 1,1 99,975%	0,1 0,3 88%
NICAM 3, Rec. J.17 del CCITT	Lineal de 14 bits Rec. J.17 del CCITT	-0,1 0,7 18%	-0,1 0,8 19%	-22 0,8 0,001%	0,1 1,0 60%

\bar{X} : nota media de opinión en una escala de -3 a +3 según la Recomendación 562.

S : desviación típica.

% Conf : nivel de confianza estadística en que el codec «A» es preferido al codec «B».

Según se puede ver en el cuadro, no se halló diferencia importante entre estos dos sistemas de codificación cuando la MDA se muestreó a 330 kHz, salvo en el caso del gong de frecuencia baja crítica en el que se percibió una conformación del espectro de ruido, que se advertía sobre todo durante el tiempo de caída del gong. Cuando la MDA se muestreó a 204 kHz, también se pudo percibir una degradación de la reproducción del gong de alta frecuencia, con un resultado peor que el NICAM 3. En estos dos casos sin embargo, no se pudo percibir una diferencia importante entre los dos codecs con material de programa normal.

2.2 Evaluación de la calidad de un código diferencial casi instantáneo y un código MIC casi instantáneo

Se ha efectuado en Canadá una evaluación objetiva y subjetiva de dos sistemas de codificación propuestos en el Informe 1075 para la radiodifusión de TVAD por satélite. Los sistemas de codificación analizados fueron:

- el sistema MICD con compansión casi instantánea (NI-MICD) de 15 a 8, propuesto para el sistema MUSE*;
- el sistema con compresión-expansión casi instantánea (NICAM) de 14 a 10, de conformidad con la Recomendación 651.

* El esquema NI-MICD empleado en MUSE se basa en un factor de fuga de 0,9375, comparado con un factor de fuga de 0,975 empleado en las simulaciones aquí citadas. Las simulaciones con valores del factor de fuga entre 0,95 y 0,99 demostraron sólo diferencias menores en las relaciones señal/ruido.

2.2.1 Evaluación objetiva

Ambos sistemas de codificación se simularon por computador, de conformidad con las descripciones que aparecen en el Informe 953-1 para el sistema NICAM y en el Informe 1075 para el NI-MICD. En la simulación del sistema NI-MICD se utilizó un factor de fuga de 0,975. Las señales de prueba de audiofrecuencia aplicadas a cada codificador fueron las siguientes:

- a) señal sinusoidal de frecuencia variable;
- b) sonidos sintéticos de baja y alta frecuencia;
- c) sonido natural de arpa.

Las señales de prueba de entrada de los codificadores consistieron en muestras digitales de audio de 32 kHz, cuantificadas uniformemente con 15 bits/muestra en el caso del codificador NI-MICD y con 14 bits/muestra para el codificador NICAM. Para cada señal de muestra y cada codificador, se calcularon las relaciones señal/ruido en bloques sucesivos de 256 muestras (8 ms) y se representaron gráficamente en función del tiempo.

La señal sinusoidal de frecuencia variable consistió en una secuencia de 10 sinusoides, de amplitud constante y frecuencia variable de 30 a 15 360 Hz. En la Figura 2a se muestra un fragmento de esta señal de prueba y en las Figuras 2b y 2c están representadas las curvas de relación S/N obtenidas. Como puede verse en la Figura 2c, un sistema de cuantificación instantánea como el NICAM tiene una relación S/N que es independiente de la frecuencia. Un sistema de codificación diferencial como el NI-MICD produce mejores valores de relación S/N en frecuencias bajas (hasta 1 kHz), pero estos valores se deterioran en frecuencias más altas, hasta llegar a ser menores que los obtenidos con el sistema NICAM.

La señal de prueba de sonido sintético consistió en las cuatro primeras notas de "Frère Jacques" (Do-Re-Mi-Do) sintetizadas a baja frecuencia (60 - 80 Hz) y a alta frecuencia (4 - 6 Hz). En la Figura 3a se ve la forma resultante de la señal en el dominio del tiempo y en la Figura 3b, las curvas de relación S/N obtenidas para el sonido de baja frecuencia. La superioridad del codificador diferencial (NI-MICD) resulta evidente para una señal de baja frecuencia y alto nivel, a pesar de su menor número de bits por muestra (8 bits, frente a 10 en el NICAM). Ambos codificadores presentan relación S/N decreciente a medida que disminuye el nivel de la señal. Las curvas de S/N obtenidas con sonido sintético de alta frecuencia se muestran en la Figura 3c. Los valores de S/N obtenidos con el sistema NICAM son superiores, para señales de alto nivel, a los obtenidos con el NI-MICD.

El sonido natural de arpa, representado en la Figura 4a, muestra una gama dinámica muy amplia así como grandes variaciones de nivel en periodos breves. En las Figuras 4b y 4c pueden verse las curvas de S/N obtenidas con esta señal de prueba. Dichas curvas son comparables, quizás con una ligera ventaja para el sistema NI-MICD respecto del NICAM, cuando la señal de entrada alcanza las crestas de su gama dinámica.

2.2.2 Evaluación subjetiva

Se realizó una prueba de audición informal, con cascos de escucha, actuando de oyentes cuatro ingenieros de radiodifusión sonora y dos estudiantes de ingeniería. Las pruebas se efectuaron según el método de comparación propuesto por el CCIR (Recomendación 562), en el que la secuencia A-B se presenta dos veces al oyente y a continuación se le dejan unos 15 segundos para que dé su puntuación, entre -3 (codificador A mucho peor que codificador B) y +3 (codificador A mucho mejor que codificador B).

El material de la prueba consistió en fragmentos de un solo de arpa, de un solo de órgano, de un solo de voz femenina y de piezas ejecutadas por un conjunto de música de cámara, un cuarteto de jazz y una banda de música pop. En el material de prueba se incluyeron también los sonidos sintéticos de baja y alta frecuencia descritos en el punto 2.2.1. Tanto con el codificador NI-MICD como con el NICAM se utilizó la ley de preacentuación/desacentuación de 50/15 μ s.

La prueba de audición aunque informal, produjo resultados homogéneos. Los oyentes no encontraron diferencias significativas entre los codificadores NI-MICD y NICAM para los seis fragmentos de música natural. La nota media de opinión para esos seis fragmentos fue superior en 0,2 para el sistema NI-MICD, con una desviación típica de 0,8.

El grupo de oyentes mostró una clara preferencia por el codificador NI-MICD en el caso del sonido sintético de baja frecuencia (nota media de 2,2 con desviación típica de 0,7) y por el NICAM en el caso del sonido sintético de alta frecuencia (nota media de -1,8 con desviación típica de 0,7). Estos resultados concuerdan con los de la evaluación objetiva del punto 2.2.1, según el cual el codificador NI-MICD produjo mejores valores de relación S/N que el NICAM para señales de baja frecuencia, pero peores en alta frecuencia.

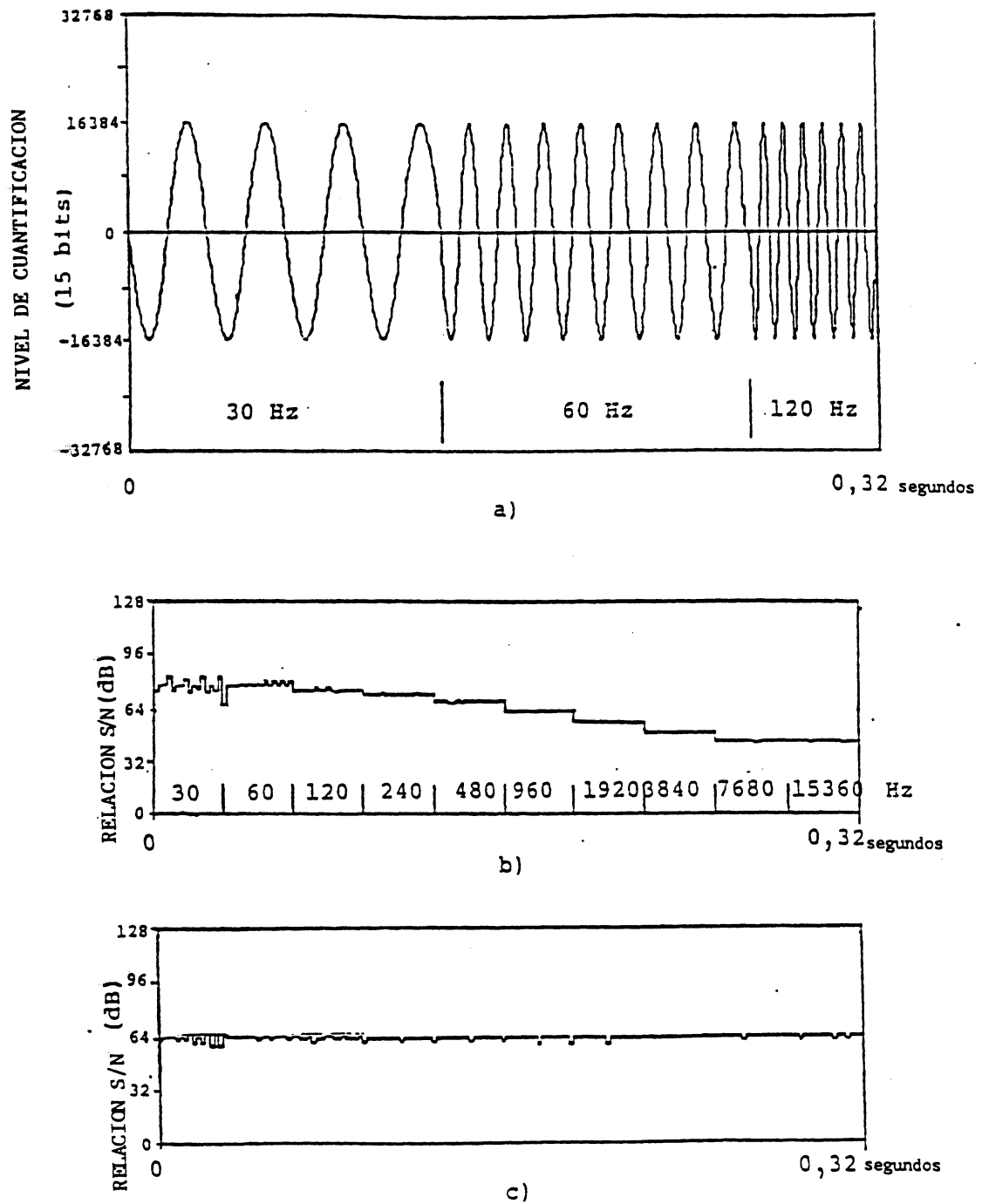


FIGURA 2

Señal sinusoidal de frecuencia variable

- a) fragmento de señal de prueba
- b) NI-MICD: relación S/N en función de la frecuencia de la onda sinusoidal
- c) NICAM : relación S/N en función de la frecuencia de la onda sinusoidal

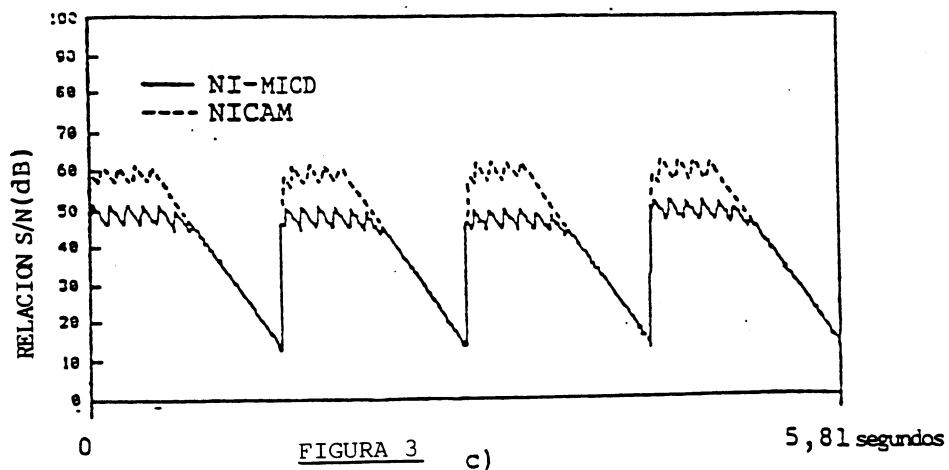
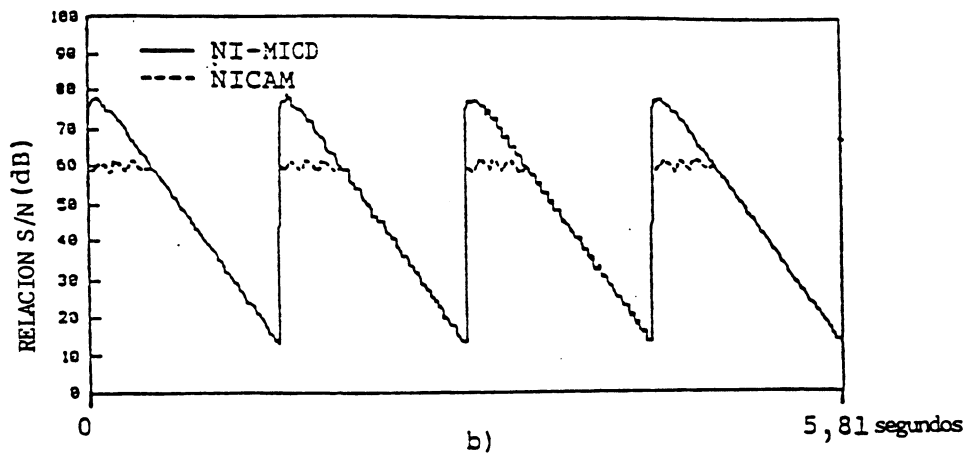
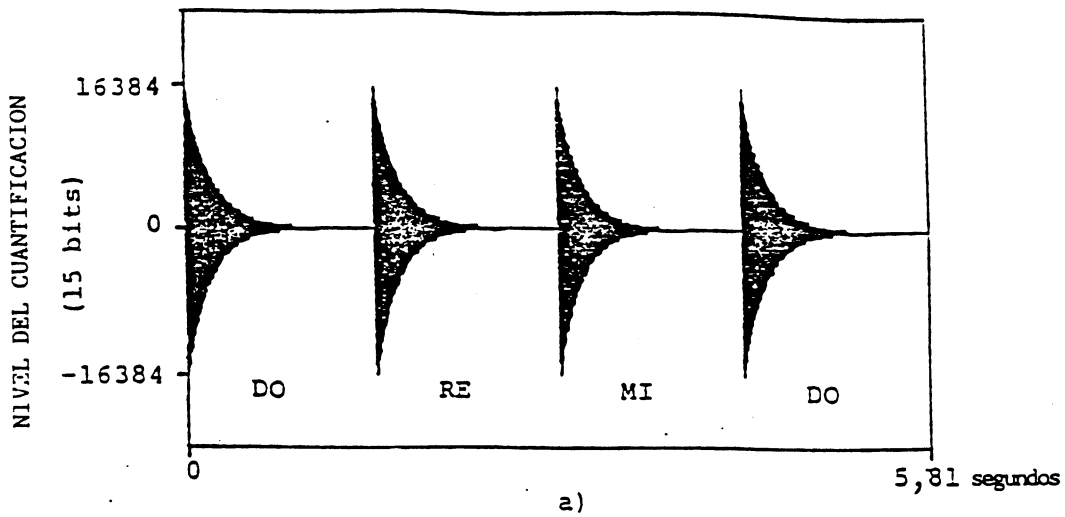
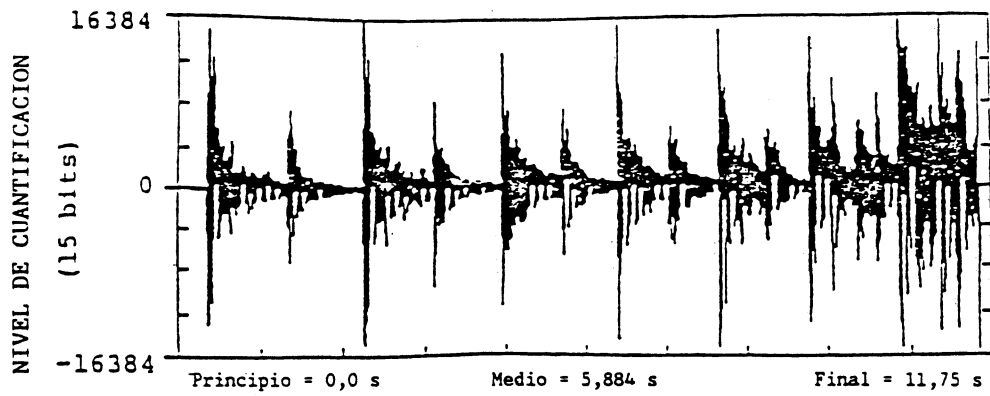


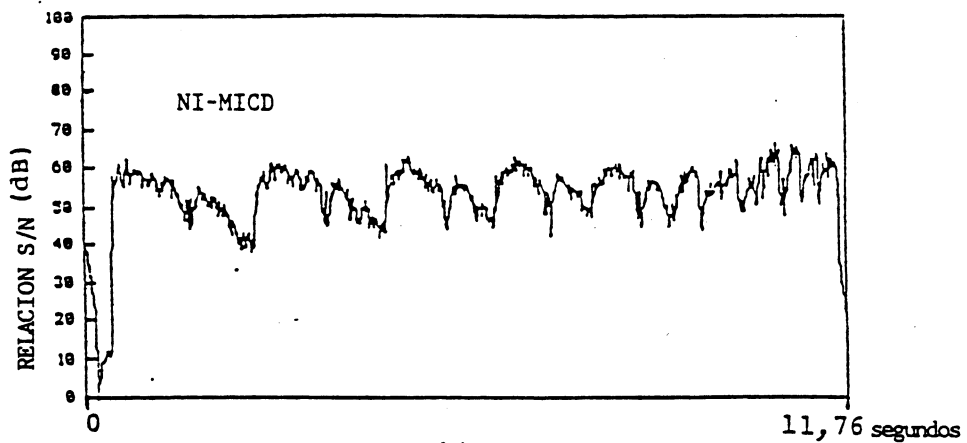
FIGURA 3 c)

Sonidos sintéticos de baja y alta frecuencia

- a) forma de onda de la señal de prueba
- b) curvas de relación S/N para sonido de baja frecuencia
- c) curvas de relación S/N para sonido de alta frecuencia



a)



b)

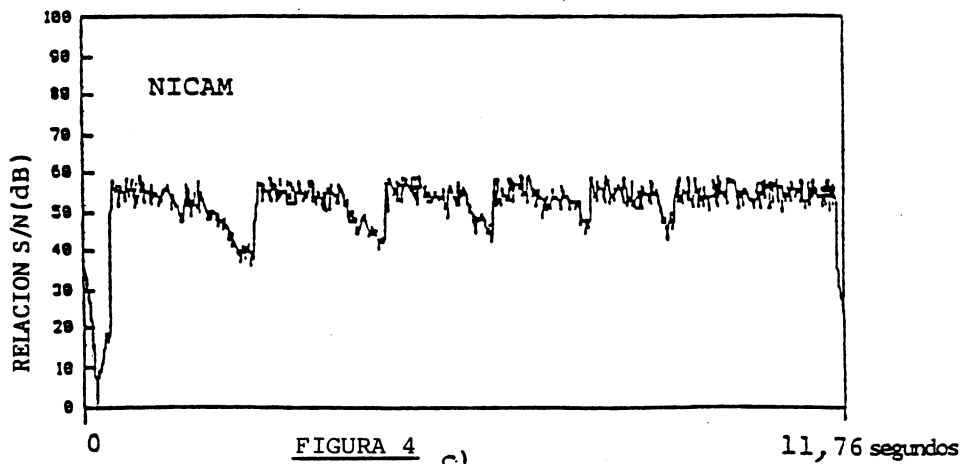


FIGURA 4 c)

Sonido natural de arpa

- a) forma de onda de la señal de prueba
- b) curva de relación S/N para el codificador NI-MICD
- c) curva de relación S/N para el codificador NICAM"

3. Mediciones subjetivas realizadas en los Estados Unidos de América

Se llevaron a cabo pruebas de escucha cuidadosamente controladas, en las que participaron oyentes entrenados, para evaluar el sistema de Modulación Delta Adaptable (MDA) descrito en el Informe 953 . En el experimento se probó y comparó el sistema MDA (velocidad de muestreo = 300 kHz) con respecto a toda una gama dinámica de discos compactos.

Para esas pruebas se contrataron veinticuatro oyentes bien dispuestos y altamente calificados, cuyo oído funcionaba normalmente desde el punto de vista clínico, según lo confirmó una evaluación audiométrica realizada por un audiólogo autorizado y diplomado. La edad de los oyentes era de 16 a 27 años para asegurar la probabilidad de un oído normal. Todos ellos eran oyentes experimentados o expertos a quienes se les indicó que debían "votar" por el sistema cuya calidad les resultara superior. No se les proporcionó más detalles acerca de lo que debían escuchar. Quince oyentes eran hombres y nueve eran mujeres y fueron remunerados convenientemente para estimular su buena disposición.

Las condiciones de escucha (entorno) se mantuvieron en estrecho acuerdo con las estipuladas en la Recomendación 562 de la Comisión de Estudio 10. Las pruebas se llevaron a cabo en un estudio de grabación con moqueta y tapicería, con un tiempo de reverberación a media frecuencia de aproximadamente 0,25 segundos y un tamaño de 16 x 22 pies. Los altavoces se separaron entre sí unos 10 pies. Las pruebas se efectuaron con un solo oyente por vez.

El nivel de escucha en el centro de la sala se fijó a 85 dB para los pasajes de música alta. Se eligió ese nivel porque a otros inferiores las diferencias entre los sistemas podrían no ser perceptibles.

Las 10 selecciones de DC, que se indican en el Cuadro V , fueron elegidas por la Unión Europea de Radiodifusión (UER) -Material de Evaluación de la Calidad del Sonido (MECS) y Discos Compactos (DC) y consideradas materiales críticos de prueba.

CUADRO III - Segmentos de programas

<p>Melodía electrónica - Frère Jacques ABBA - Música Pop Triángulos - Instrumento único Gran piano - Teclas aisladas y breve selección musical Campanas - Instrumento único Eddie Rabbitt - Música country Soprano - Solo Conjunto de instrumentos de viento - Breve selección musical Xilófono - Teclas aisladas y selección musical Hombre hablando en alemán</p>
--

Estas selecciones estereo utilizadas con fines de prueba presentaban amplias variaciones en lo que respecta a la gama dinámica, el espectro y las características temporales. Representan materiales de radiodifusión típicos y atípicos. Se eligieron por ser especialmente problemáticos para el tipo de tratamiento MDA en prueba, y por revelar distorsiones de codificación. Entre las deficiencias de calidad del sistema que podrían ser audibles figuran: linealidad A/D y D/A, distorsión de repliegue, reducción de la velocidad binaria y ruido modulado por la señal radiofónica.



Los extractos duraron entre 20 y 140 segundos y el tiempo total de las pruebas subjetivas fue de media hora.

Se utilizó el método psicofísico clásico de elección obligada entre dos posibilidades (Two Alternatives Forced-Choice (2 AFC)). (El método que figura en la Recomendación 562 de la Comisión de Estudio 10 no era apropiado, pues se trataba de detectar una señal y no de establecer una escala de calidad.) Las diez selecciones se presentaron cuatro veces cada una, en dos órdenes aleatorios. Los oyentes conmutaban hacia adelante y hacia atrás los dos sistemas de una pareja y luego votaban por el de "calidad sonora superior", inmediatamente después de finalizar cada selección. Los segmentos audio estaban separados entre sí aproximadamente cinco segundos.

En el Cuadro IV figuran los resultados iniciales de las pruebas, tales como, para cada selección, el número de juicios positivos emitidos y el porcentaje que ese número representa. De ese modo se puede examinar la influencia, en caso de que exista, de la selección del segmento audio en la elección de los oyentes, y realizar una comparación directa de los dos trayectos audio. Cuando no existe ninguna diferencia apreciable, los resultados se distribuyen de forma aleatoria con una mediana de $N/2$ (50%), y una desviación típica de $\sqrt{N/4}$ (donde N es el número de pruebas).

Cada oyente emitió 40 juicios, es decir, cuatro votos sobre cada una de las diez selecciones. Por lo tanto, el grupo de 24 oyentes emitió un total de 960 juicios. La selección aleatoria daría como resultado (por selección) 48 votos (50%) para el sonido DC, con una desviación típica de 4,9 votos (5,1%).

El resultado global de las pruebas iniciales (Cuadro IV) es de un 52,5% de preferencia para el MDA y un 47,5% para el DC, por lo que la diferencia, si la hay, es muy pequeña. Desde este punto de vista general se podría llegar a la conclusión de que el compresor-expansor es acústicamente transparente. Sin embargo, en cinco selecciones se manifestó una preferencia que difería en más de un 5% del 50% (45 a 55%): la melodía electrónica, 55,2% de preferencia para el DC; la ABBA, 65,6% de preferencia para el MDA; los triángulos, 61,5% para el DC; el solo de soprano, 59,4% para el MDA y el xilófono, 62,5% para el MDA; se advirtió asimismo que el juicio de 3 a 5 oyentes era la causa fundamental de este resultado.

Dado que unos pocos oyentes fueron capaces de detectar diferencias en estas selecciones críticas (aunque no existía preferencia alguna por uno de los trayectos audio), se llevó a cabo un estudio complementario para determinar si las pequeñas diferencias en el nivel sonoro podían ser la causa de los resultados obtenidos. La adaptación de los niveles de los dos trayectos (trayecto directo DC y trayecto DC vía códec MDA) se mejoró de aproximadamente 1,0 dB de diferencia a menos de 0,5 dB). Las condiciones de prueba se mantuvieron invariables. En el Cuadro V se indican los resultados.

Después de efectuar los ajustes adicionales para adaptar los niveles, a los oyentes expertos les resultó aún más difícil distinguir el trayecto del códec MDA del trayecto directo de la señal DC. Tres de los cinco segmentos radiofónicos se aproximaron en unos 10 puntos porcentuales a los verdaderos resultados aleatorios (50/50) y terminaron en 44, 50, y 52%. Los resultados correspondientes a las otras dos selecciones no cambiaron de forma apreciable de (2 a 3 puntos porcentuales).

Una prueba cuidadosamente controlada ha demostrado que oyentes experimentados y expertos no mostraron preferencia entre el sonido de material de radiodifusión crítico reproducido por una fuente directa de DC y el sonido de un DC escuchado a través del códec MDA en las condiciones de prueba ya descritas.

CUADRO IV

Resultados de la prueba de escucha experimental

Agosto de 1988

Votos a favor del DC

24 sujetos de prueba

96 juicios de elección obligada por selección

	Total de votos	Porcentaje
1 - Melodía electrónica # 7	53	55,2
2 - ABBA # 69	33	34,4
3 - Triángulos # 32	59	61,5
4 - Piano grande # 39	47	49,0
5 - Campanas # 34	48	50,0
6 - Eddie Rabbitt # 70	49	51,0
7 - Soprano # 44	39	40,6
8 - Conjunto de instrumentos de viento # 66	47	49,0
9 - Xilófono # 36	36	37,5
10 - Hombre hablando en alemán # 54	45	47,0
TOTALES	456	47,5

La desviación típica de 96 sucesos aleatorios = 4,9 (5,1%).

CUADRO V

Resultados de las pruebas de escucha complementarias

Septiembre de 1988

Votos a favor del DC

20 sujetos de prueba

160 juicios de elección obligada por selección

	Total de votos	Porcentaje
1 - Melodía electrónica # 7	86	53,8
2 - ABBA # 69	70	43,8
3 - Triángulos # 32	80	50,0
4 - Soprano # 44	83	52,0
5 - Xilófono # 36	65	40,6
TOTALES	384	48,0

La desviación típica de 160 sucesos aleatorios = 6,3 (4%).

REFERENCIA BIBLIOGRÁFICA