

# **UNIVERSIDAD NACIONAL DE INGENIERÍA**

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



## **ANÁLISIS DEL PROBLEMA DEL LOUDNESS EN EL AUDIO DE LA TELEVISIÓN DIGITAL TERRESTRE Y LAS ALTERNATIVAS DE SOLUCIÓN**

### **INFORME DE SUFICIENCIA**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:  
INGENIERO EN TELECOMUNICACIONES**

**PRESENTADO POR:  
CHRISTIAN JESÚS RAVICHAGUA INGA**

**PROMOCIÓN  
2010-I**

**LIMA - PERÚ  
2014**

**ANÁLISIS DEL PROBLEMA DEL LOUDNESS EN EL AUDIO  
DE LA TELEVISIÓN DIGITAL TERRESTRE Y LAS  
ALTERNATIVAS DE SOLUCIÓN**

**DEDICATORIA:**

El presente trabajo se lo dedico a mis padres, mi hermano y a mi esposa, que gracias a su incondicional apoyo pude concluir mi carrera.

### **AGRADECIMIENTOS:**

En primer lugar quiero agradecer a Dios por estar conmigo en cada paso que doy, por fortalecer mi corazón e iluminar mi mente, por haber puesto en mi camino a aquellas personas que han sido el soporte y compañía durante el periodo de estudio.

Quiero agradecer a mi familia, a mis padres, mi hermano y a mi esposa, por su apoyo y confianza en todo lo necesario para cumplir mis objetivos como persona y estudiante, por brindarme los recursos necesarios y estar a mi lado apoyándome y aconsejándome siempre.

Quiero agradecer a mis maestros quienes me han enseñado a ser mejor en la vida y a realizarme profesionalmente.

Un agradecimiento especial a mi asesor Ing. Marcial Antonio López Tafur por hacer posible este informe.

A mis compañeros de clases quienes me acompañaron en esta trayectoria de aprendizaje y conocimientos.

## **RESUMEN**

En este trabajo se revisa las bases del estudio sobre el que se ha realizado la investigación del modelo de sonoridad, sintetiza adecuadamente los conocimientos en el tema de la sonoridad, el cual no es abarcado en su totalidad en este informe.

En este trabajo se presentan las alternativas de solución a la problemática del loudness (sonoridad) que se manifiesta en el audio de la Televisión Digital Terrestre analizando las recomendaciones de los organismos internacionales de normalización. Asimismo se propone los lineamientos de las recomendaciones pertinentes para una posible reglamentación en nuestro país, según las pautas de medición en la sonoridad del audio digital.

Finalmente en este trabajo se muestra el diseño conceptual de un medidor de sonoridad basado en la aplicación del algoritmo propuesto en la solución dada por la norma BS.1770, utilizando programas de software adecuados. Para la simulación del algoritmo y la medición de su desempeño se han utilizado la aplicación diseñada para la edición de audio Adobe Audition, para la medición de los parámetros se ha utilizado los analizadores de espectro SpectraLab y TrueRTA, para las mediciones comparativas de sonoridad se ha utilizado el prototipo del Loudness Meter de Orban y el Loudness Meter vR1.7 y en la programación por software de los filtros de procesamiento digital se ha utilizado el programa MATLAB.

## ÍNDICE

<b>INTRODUCCIÓN .....</b>	<b>1</b>
<b>CAPITULO I</b>	
<b>PLANTEAMIENTO DE INGENIERÍA DEL PROBLEMA.....</b>	<b>4</b>
1.1 Descripción del problema.....	4
1.2 Objetivos del trabajo .....	4
1.3 Evaluación del Problema.....	5
1.4 Síntesis del trabajo .....	5
<b>CAPITULO II</b>	
<b>MARCO TEÓRICO CONCEPTUAL.....</b>	<b>7</b>
2.1 Antecedentes del problema.....	7
2.2 Medición del nivel del audio .....	8
2.2.1 El sonido.....	8
2.2.2 Nivel de Potencia Sonora .....	10
2.2.3 Valores eléctricos .....	11
2.2.5 El vúmetro.....	12
2.2.6 Niveles de referencia .....	13
2.2.7 Medidor de crestas o picos .....	15
2.2.8 El gálibo y la franquía de una señal de audio .....	16
2.3 El Loudness .....	17
2.3.1 Como interpretamos la sonoridad.....	20
2.3.2 Medición del nivel de señal .....	24
2.4 Evaluación del Loudness .....	25
2.4.1 Fisiología auditiva.....	25
2.4.2 Aspectos de la sonoridad.....	28
2.5 Fundamentos de la compresión dinámica .....	33
2.5.1 El sobre rango o saturación digital .....	36
2.5.2 La Guerra de la sonoridad .....	39
<b>CAPITULO III</b>	
<b>METODOLOGÍA PARA LA SOLUCIÓN DEL PROBLEMA.....</b>	<b>41</b>
3.1 Aporte de las distintas normativas .....	41
3.1.1 Recomendación UIT BS.775.....	41
3.1.2 Recomendación UIT BS.1770 – 2. ....	43

3.1.3	Recomendación UIT BS.1771 – 1.....	45
3.1.4	Documento A/85 del ATSC.....	48
3.1.5	Documento CST - RT – 017 / 019.....	49
3.1.6	Documento ARIB TR - B29.....	51
3.1.7	Documento EBU-UER R-128.....	55
3.2	El Loudness en el Audio Digital.....	57
3.2.1	Norma Brasileña ABNT NBR 15602-2.....	57
3.2.2	El caso peruano.....	60
3.3	Criterios de diseño de un medidor de loudness.....	60

#### **CAPITULO IV**

	<b>ANÁLISIS Y PRESENTACIÓN DE RESULTADOS .....</b>	<b>62</b>
4.1	Análisis descriptivo de las variables de estudio.....	62
4.2	Análisis teórico de los datos.....	64
4.3	Costos de Software.....	68
4.4	Resultados obtenidos.....	69
	<b>CONCUSIONES Y RECOMENDACIONES.....</b>	<b>72</b>
	<b>ANEXOS .....</b>	<b>75</b>
	<b>GLOSARIO .....</b>	<b>84</b>
	<b>BIBLIOGRAFÍA.....</b>	<b>89</b>

## INTRODUCCIÓN

Históricamente la posibilidad de lograr un registro sonoro ha sido un anhelo cultural limitado por la tecnología. En este sentido la escritura, la pintura y la arquitectura podrían ser tomadas como productoras de registros culturales.

Los registros sonoros se inician recién a finales del siglo XIX cuando Edison logra los primeros registros por medios mecánicos (1877). El posterior desarrollo de la electrónica permitió mejorar su calidad y posibilitó su uso universal. Esta calidad sonora se definió entonces como Hi Fi o Alta Fidelidad.

En este siglo la masiva utilización de tecnologías digitales, permitió llegar aún más lejos, tanto en calidad como en universalidad. Pero aquí surgió el problema de que las técnicas utilizadas en el entorno analógico no siempre son óptimas en el entorno digital, por lo que se comienza a evaluar y corregir.

Este trabajo busca ser un aporte para la medición de los niveles del audio digital.

### Motivo de elección y desarrollo del tema

Las constantes mejoras en la calidad de audio y video obligan a los realizadores a estar más pendientes de los detalles técnicos, para poder ofrecer productos de calidad a unos usuarios cada vez más exigentes.

La Televisión Digital Terrestre ofrece Televisión de alta definición, imágenes 3D, programas interactivos, son algunos de las tendencias que hoy se toman el mundo de la producción audiovisual; sin embargo, no se debe olvidar otro componente fundamental de un proyecto de calidad: El Sonido.

Aunque los desarrollos en este campo no son tan conocidos, populares o promocionados como las tecnologías mencionadas anteriormente, también hacen parte de la nueva oferta existente en el mercado, por lo que es una necesidad para los actores de la industria conocer a fondo las alternativas con las que pueden llegar a contar.

### Fundamentación

Desde el momento en que la tecnología ha permitido convertir la presión acústica de un **sonido** en una señal eléctrica (**audio**), se ha planteado la necesidad de controlar su nivel eléctrico, equivalente a su volumen sonoro. Tradicionalmente esta magnitud se



ha visualizado y ponderado por medio de un medidor consistente en un instrumento de aguja o galvanómetro d'Arsonval y también por medio de Diodos electroluminiscentes (LED) o gráficamente en una pantalla. [1]

En radiodifusión, las exigencias de controlar y homogenizar los niveles de la señal, en su procesamiento y posterior emisión, ha hecho necesaria la estandarización del equipamiento de estudio, enlaces y planta transmisora, requiriendo de estándares que permitieran la industrialización de sus componente y el intercambio de material. [2,3]

La existencia de grandes variaciones del nivel de audio exigen su limitación en el manejo eléctrico del material sonoro (tanto en la industria discográfica como en radiodifusión), donde una errónea configuración o control lleva a que se perciban grandes diferencias de volumen en los contenidos sonoros (entre programas, estaciones o medios discográficos), con la consecuente disconformidad de los oyentes [4]. Además de la posible saturación y distorsión causadas por la sobrecarga de los equipos.

La necesidad de contar con una correlación objetiva entre este parámetro eléctrico y el característico comportamiento fisiológico del oído, exige la utilización de medidores que evalúen su amplitud ponderada. Hoy en día la radiodifusión y las compañías productoras de música usan una variedad de plataformas de distribución hacia diferentes grupos de oyentes, pero todos tienen una cosa en común, el grado con que perciben los niveles del sonido.

La utilización de tecnología informática en la adquisición, edición y emisión de material sonoro exige que los equipos utilizados tengan una serie de características que deben ser tomadas en cuenta para una correcta operación. Esto permite además la incorporación de metadatos para identificación y control. [6]

En este trabajo se presenta el modelo de sonoridad obtenido de los estudios actuales, examinando los documentos y las recomendaciones disponibles orientadas a que una cadena de audio digital (tanto en radiodifusión como en producción audiovisual), pueda ser optimizada para obtener una placentera experiencia sonora.

## **Limitaciones**

El problema del loudness en la Televisión Digital terrestre no es fácil de resolver, aunque ya muchos países han aprobado estrictas regulaciones para proteger a los consumidores. De igual forma, la industria está realizando un esfuerzo por autorregularse, muestra de ello son las prácticas recomendadas en la norma A/85 del Comité Americano para Sistemas Avanzados de Televisión (ATSC) o en el caso europeo con la norma EBU R-128.

“Los problemas de audio no suelen ser prioridad para la mayoría de las estaciones de televisión. Lo anterior se da, a pesar de que las encuestas a los clientes muestran que el audio representa al menos el 50% de la experiencia del espectador, y un sonido deficiente puede poner en peligro la alegría de ver televisión”, afirma Carlos Watanabe Director Regional en Latinoamérica de *Dolby Laboratories*.

### **Métodos empleados en la investigación del tema**

Para el desarrollo del presente informe, se considera necesario conocer los conceptos ligados a la sonoridad, tales como, las características físicas del sonido, la fisiología auditiva humana, la metodología de medición del nivel sonoro, la métrica aplicada, los requerimientos de producción, los recursos informáticos utilizados y los algoritmos de medición de la sonoridad, que han de ser revisados en este trabajo.

### **Análisis críticos sobre las principales fuentes bibliográficas consultadas**

La información obtenida de las fuentes consultadas ha permitido el abordaje de una visión de referencia histórica, a una solución de aplicación actual. De su análisis se ha obtenido una perspectiva de los criterios, un panorama de las alternativas de solución presentadas hasta ahora y de las recomendaciones y estándares emanados por las organizaciones internacionales y los organismos pertinentes. En esos trabajos se observa que el estado actual de la informática aplicada al sector de la industria del audio profesional está muy vinculado a la capacidad de innovación tecnológica, siendo en estos momentos un nicho de gran expansión que requiere para su aplicación una estandarización de nuevos parámetros.

### **Alcances técnico o científicos del contenido del informe**

Este trabajo plantea la imperiosa actualización de los conocimientos y las técnicas aplicadas a la producción y manejo de señales de audio digital, ya sea detallando las ventajas y los parámetros de los metadatos del audio digital, como también presentando las normativas de medición y determinación de los niveles de sonoridad.

Se incluyen asimismo las definiciones de los términos y los parámetros que están involucrados en la medición de la calidad de un equipo empleado como Estación de Trabajo Digital (Digital Audio Workstation). [6]

# **CAPITULO I**

## **PLANTEAMIENTO DE INGENIERÍA DEL PROBLEMA**

### **1.1 Descripción del problema**

El cambio de paradigma dado a fines del siglo XX involucró la migración de todos los sistemas analógicos anteriores al dominio de la cuantificación digital de las señales y su procesamiento informático, universalizado en el siglo XXI. Hoy en día tenemos la posibilidad de obtener registros sonoros prácticamente perfectos, la utilización de computadoras como estaciones de trabajo de audio digital (Digital Audio Workstation), es hoy la manera común de resolver una producción. [8] Aunque existen fortalezas y debilidades de los sistemas informáticos cuando se intenta usar en la producción de material de audio profesional. [9] Así como esta continua evolución tecnológica lleva a cambios en los estándares. [7]

Ahora la cuestión informática en audio se plantea en la generación, conversión, almacenamiento y difusión, que en definitiva es el manejo de estos altos volúmenes de datos que requieren una gran potencia computacional y controles que en un principio emulen a los analógicos y luego permitan parametrizar valores que eran imposibles de definir en el dominio analógico, como es la utilización de bases de datos, redes, edición no lineal, duplicación sin pérdidas, inclusión de metadatos.

Esto involucra la imperiosa necesidad de migrar desde los medidores analógicos con rudimentarios modelos de evaluación sonora, a los medidores compuestos con etapas de procesamiento, aplicando algoritmos establecidos por profusos estudios, con salida en múltiples e innovadoras presentaciones luminosas.

### **1.2 Objetivos del trabajo**

Los objetivos generales de este trabajo son:

- ❖ Sistematizar los conocimientos en el tema del Loudness.
- ❖ Presentar alternativas de solución a la problemática del Loudness manifestada en el audio de la Televisión Digital Terrestre.

Como objetivos específicos se pretende:

- ❖ Crear un documento de referencia que contenga la problemática del loudness en la televisión Digital terrestre y se proponga alternativas de solución.
- ❖ Diseñar un concepto de medidor de sonoridad (loudness).
- ❖ Proponer los lineamiento para una normativa sobre especificación de niveles de audios producciones audiovisuales

### **1.3 Evaluación del Problema**

Cuando se tienen programas con diferentes niveles de loudness dentro de un canal en particular, por ejemplo intervalos de comerciales, o canales con valores diferentes entre ellos, el consumidor obtiene una mala experiencia de audio, viéndose obligado a ajustar el volumen cada vez que haya una diferencia en el loudness. De hecho, tanto en televisión abierta como de paga, se reporta que este factor, es la principal queja respecto al audio que se recibe de los consumidores.

El audio digital permite una calidad de audio mucho más alta, sin distorsión y con un rango dinámico mucho más alto. Estas mejoras a su vez se convierten en amenazas para los realizadores, pues un audio deficiente será mucho más evidente en un entorno digital.

El paradigma de nivelación por sonoridad afecta a todas las etapas de una señal de radiodifusión, desde la adquisición hasta la transmisión, y el fin último es el de armonizar los niveles de sonoridad dentro de la misma cadena de radiodifusión y, por ende, entre todas ellas a un único nivel de sonoridad universal que beneficie a los oyentes. Por ello, muchos de los profesionales y todos los equipos de medición de audio que forman parte de la cadena de radiodifusión se verán afectados por este cambio.

### **1.4 Síntesis del trabajo**

En este trabajo se trata el tema de la medición de la sonoridad, siendo una recopilación sistemática de las investigaciones que han llevado al cambio de paradigma en la medición de los niveles de audio. También se muestra la descripción de los requerimientos en la producción de material de audio digital y su post-producción aplicando tecnología informática que puede impulsar el desarrollo de hardware y software específicos.

La obtención de un panorama de las necesidades y normativas, ha posibilitado conocer donde se puede desarrollar la innovación tecnológica. Conociendo los parámetros que los estándares definen sobre las características y los mecanismos de

medición que se pueden implementar en las aplicaciones de audio digital, utilizando una PC. [9]

Al analizar las distintas normativas que se están determinando a nivel mundial, se concluye que la creación de una recomendación para las producciones locales permitirá cumplir con estos estándares, dando a la industria local la alternativa de abordar los mercados internacionales, al cumplir con las normas de calidad de sonido. [40, 41, 42]

## **CAPITULO II**

### **MARCO TEÓRICO CONCEPTUAL**

#### **2.1 Antecedentes del problema**

En la búsqueda de registrar y reproducir sonidos de la forma más fiel y perfecta posible, se ha planteado recientemente un cambio de paradigma en la nivelación del audio.

Frente al compromiso de obtener el mejor régimen de nivel sin sobrepasar los márgenes de señal en los equipos electrónicos involucrados, un operador comúnmente se enfrenta al constante ajuste de las señales, que deben ser procesadas para evitar su degradación por causas técnicas.

Transitando desde lo conceptual hacia lo concreto, se reseñan los trabajos que tienden definir objetivamente algo tan subjetivo como es la percepción de la sonoridad (loudness), donde en busca del mejor algoritmo de medición y procesamiento se ha investigado una serie de modelos aplicados en la elaboración de los medidores comerciales de sonoridad (loudness). [10, 11]

De los documentos científicos publicados recientemente, se reconoce que varias organizaciones internacionales como: la Unión Internacional de Telecomunicaciones (UIT), la Comisión Electrotécnica Internacional (IEC), la Organización Internacional de Estándares (ISO), el Instituto de Ingenieros de electricidad y electrónica (IEEE), la Sociedad de Ingenieros en Audio (AES), han encarado estudios y publicado pautas sobre los parámetros recomendados en la operación del audio digital. Asimismo varias organizaciones gubernamentales como la Comisión del Sistema de Televisión Avanzada de los EEUU (ATSC), la Comisión Superior Técnica de la imagen y el sonido de Francia (CST), el Instituto de la Radiodifusión Sonora (IBS) de la Inglaterra, la Unión Europea de Radiodifusión (EBU – EUR), la Asociación Europea para la Coordinación en la Estandarización (ANEC) – Bélgica, han producido reglamentaciones vinculadas a la aplicación de éstas nuevas pautas tecnológicas, efectuando este estudio confinado al análisis de los puntos relacionados con la sonoridad.

## 2.2 Medición del nivel del audio

Con el término *sonido*, se especifica un fenómeno vibratorio transmitido en forma de ondas por un medio elástico (el aire) y que llega a nuestro sistema auditivo, dándonos la sensación de audición.

Mientras que el término *audio* (o *audiofrecuencia*) identifica a la señal electrónica correspondiente al sonido, delimitando la banda que ocupa dentro del espectro electromagnético. Así todo equipo electrónico que maneje una señal de audio debe desempeñarse sin introducir modificaciones no deseadas en sus características. [2, 6, 12, 15]. Parece una tarea fácil pero no lo es, por la innata limitación de los transductores y los dispositivos electrónicos.

### 2.2.1 El sonido

El sonido humanamente audible consiste en ondas sonoras que se producen cuando las oscilaciones de la presión del aire, son convertidas en ondas mecánicas en el oído humano y percibidas por el cerebro. La propagación del sonido es similar en los fluidos, donde el sonido toma la forma de fluctuaciones de presión. [13]. El rango de variación en la presión sonora que el oído humano maneja va desde  $0.0002 \text{ dinas/cm}^2$  ( $20\mu\text{Pa}$ ) hasta  $200 \text{ dinas/cm}^2$  ( $20\text{Pa}$ ), esto corresponde a un rango de un millón de veces, el cual ha sido trasladado al terreno electrónico, por medio de un transductor. Cabe mencionar que con la llegada de la tecnología digital ha sido posible disponer de medios de distribución que manejen esta amplitud de niveles. [17] Para operar convenientemente con este rango de valores se utiliza una magnitud adimensional aplicando la función logaritmo, esto permitió representar este gran rango como una representación de exponentes lineales, tomando además como referencia su relación con el umbral de audición humana.

Su desarrollo fue concebido en los laboratorios de la AT&T, denominado Bel, en honor a Alexander Graham Bell (1847 - 1922). [14] Su uso universal se estableció en la IV reunión Plenaria de la CCITT en 1968 en Mar de Plata, Argentina, aunque no es una unidad ISO nativa, su uso es aceptado. [17, 28]

El Bel acústico es una unidad cuya referencia es el umbral promedio de audición humana ( $20\mu\text{Pa}$ ). Para una mejor representación de valores, se utiliza la unidad 10 veces menor al Bel, este se denomina decibel (dB), que minimiza el uso de números con decimales y se define con la fórmula:

$$\text{Nivel de Presión Sonora [dB]} = 20 * \log\left(\frac{Pres_x}{Pres_{ref}}\right) \quad (2.1)$$

Donde:  $Pres_x$  es la presión medida y  $Pres_{ref}$  es la presión de referencia (umbral de audición).

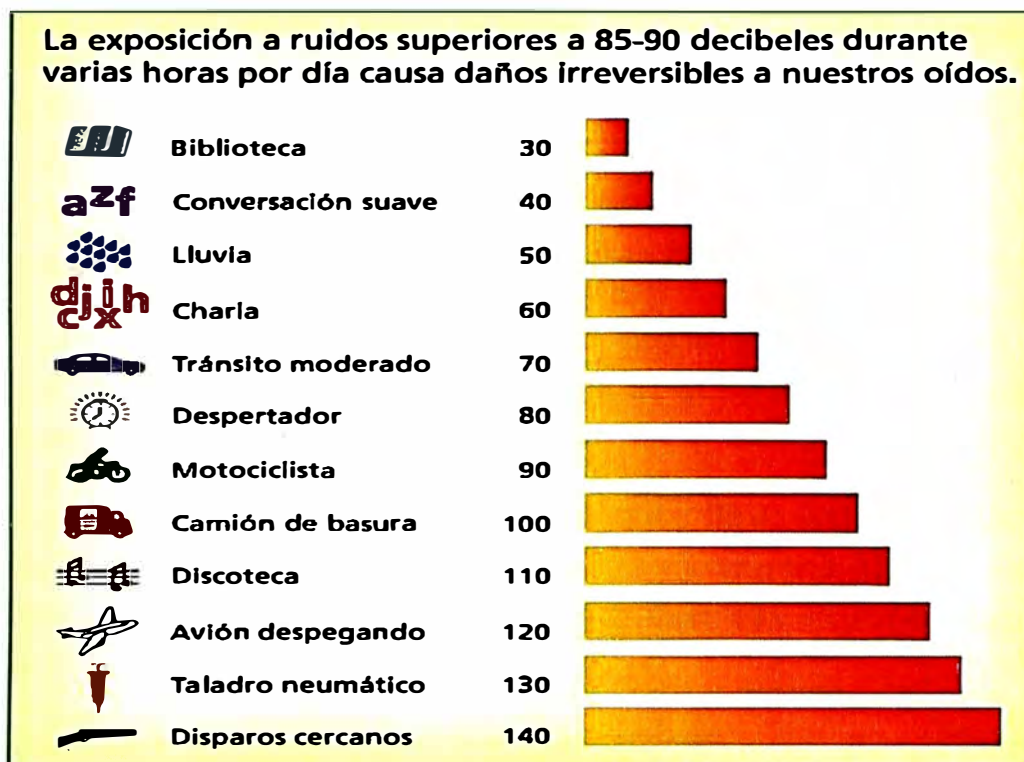
Si la presión medida  $Pres_x$  es menor que la referencia, el valor resultante será negativo, estos valores no se utilizan en acústica ya que los valores negativos indican niveles no audibles.

**TABLA N° 2.1** Rango de audición  
Fuente: Elaboración propia

Presión [Pa]	dB	Sensación	Ejemplos
20	120	Aturdimiento	Turbina jet, concierto de rock, explosión
2	100	Alto volumen	Discoteca, música electrónica, despertador
0.2	80	Sonido fuerte	Sonido de automóvil, musical en primer plano, discurso
0.02	60	Sonido destacado	TV o música suave, charla, oficina, área de recepción
0.002	40	Sonido débil	Conversación en voz baja, ambiente urbano hogareño
0.0002	20	Muy débil	Silencio nocturno
0.00002	0	Umbral Auditivo	Silencio absoluto

Los niveles de presión sonora son identificados como  $dB_{SPL}$  (Sound Pressure Level), con esta unidad se define el rango de audición de 0 a 120 dB como muestra la tabla 2.1, donde se observa que el nivel de referencia de 0 dB, sólo se obtiene en salas acústicamente tratadas, 20 dB es el nivel más bajo que comúnmente encontramos, 40 dB es el nivel normal en una habitación, 60 dB se encuentra en el entorno cotidiano, 80 dB es un sonido fuerte, 100 dB es un sonido muy fuerte, que no se tolera mucho tiempo y 120 dB es el nivel máximo soportable y por muy poco tiempo, luego del mismo se entra en la zona de aturdimiento que puede llegar a la pérdida momentánea o definitiva de la audición, y en muchos casos a la disminución de la capacidad auditiva. [18]





**Figura 2.1 Niveles de intensidad del sonido**

Fuente: <http://contaminacion1auditiva.wikispaces.com>

### 2.2.2 Nivel de Potencia Sonora

El nivel de potencia sonora (NWS, en inglés PWL), no debe confundirse con el nivel de presión sonora, puesto que mientras que el SPL se relacionan presiones en Pascal, en el NWS se relacionan potencias en vatios.

Debido a que el margen de potencias (no presiones), que se encuentran en la vida diaria, están en la proporción 10/1, en la unidad de medida más cómoda es igualmente el decibelio.

La referencia para esta medidas es de  $10^{-12}$  vatios.

La fórmula de cálculo para el nivel de potencia sonora será:

$$\text{Nivel de Potencia Sonora [dB]} = 10 * \log \frac{W}{10^{-12}} \quad (2.2)$$

Donde  $W$  es la potencia acústica en vatios.

Dado que  $10^{-12}$  vatios, corresponde a un nivel de -120 dB, la fórmula anterior se puede expresar:

$$\text{Nivel de Potencia Sonora [dB]} = 10 * \log W + 120 \quad (2.3)$$

### 2.2.3 Valores eléctricos

El desarrollo inicial de un medidor de nivel de audio para radiodifusión involucró el trabajo de Chinn et al, definiendo lo que sería el enlace estudio – transmisor. Se utilizó un sistema de impedancias apareadas (para reducir al mínimo las consecuencias producidas por las grandes distancias), ya experimentado en la telefonía. [3]

Se especificó un sistema de 600 ohms de impedancia, una potencia de señal de 1mW y un instrumento de medición del nivel de señal identificado como *medidor de unidades de volumen*, conocido como vúmetro. [3]

La potencia de 1mW disipada sobre una resistencia de 600 ohm requiere de un voltaje de 0.77459 V eficaces o RMS (redondeado comúnmente a 0.775 V). Por lo que esta potencia se definió como el valor de referencia eléctrico y en una escala graduada en decibeles correspondió a 0 dBm.

Debido a la tecnología de las válvulas electrónicas de esos años, la utilización de impedancias apareadas para los enlaces era lo apropiado. Sin embargo hoy en día con el uso de amplificadores de estado sólido realimentados, que llegan a fracciones de ohm en su salida, la necesidad de elevar su impedancia a 600 ohms apareados representa una pérdida de potencia de -6 dB. [2, 11]

### 2.2.4 Ganancia

En electrónica, la ganancia, en lo referido a señales eléctricas, es una magnitud que expresa la relación entre la amplitud de una señal de salida respecto a la señal de entrada. Por lo tanto, la ganancia es una magnitud adimensional, que se mide en unidades como belio (símbolo: B) o submúltiplos de éste como el decibelio (símbolo: dB).

La potencia eléctrica expresada en decibeles se calcula con:

$$N[dB] = 10 * \log\left(\frac{P}{P_{ref}}\right) \quad (2.4)$$

Dónde:  $N$  = es la cantidad de decibeles expresados en potencia

$P$  = es la potencia media

$P_{ref}$  = es la potencia de referencia (1mW para dBm)

Esta fórmula puede ser expresada en términos de tensiones como:

$$N[dB] = 20 * \log\left(\frac{V}{V_{ref}}\right) \quad (2.5)$$

Dónde:  $N$  = es la cantidad de decibeles en tensión

$V$  = es la tensión media

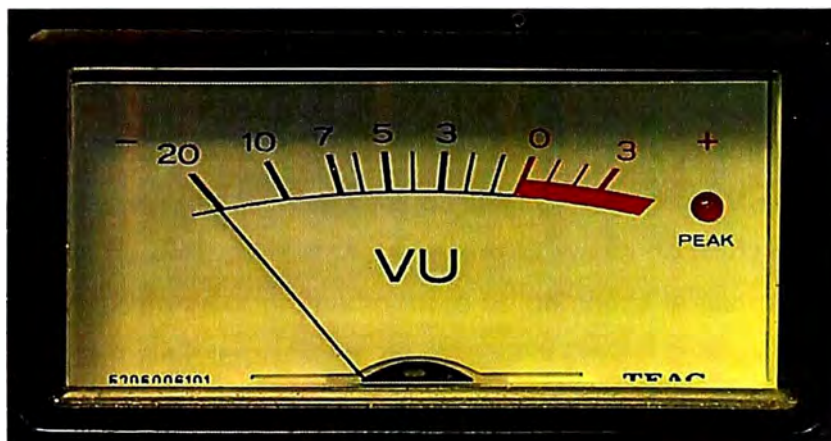
$V_{ref}$  = es la tensión de referencia (0.775V para dBm o dBu)

Asumiendo en este cálculo que la resistencia que interviene en ambos es de 600 ohms. Se definió que el nivel de operación estándar (SOL, Standard Operation Level) sobre las líneas de interconexión entre equipos, sea de +8 dBm, aunque también se ha optado por el uso de +4 dBm, que exige menores potencias. [2] este valor SOL representa el máximo nivel promedio, medido con un medidor de unidades de volumen estandarizado (el vúmetro). [11, 20]

### 2.2.5 El vúmetro

Es un instrumento de medición que se desarrolló inicialmente como un indicador de volumen estándar (*SVI Standard Volume Indicator*), para controlar y monitorear los niveles de audio. Aunque sus especificaciones reflejan la filosofía de los años 30, desarrollado originalmente en 1939 por el esfuerzo combinado de Bell Labs y los organismos de radiodifusión CBS y NBC para la medición y la normalización de los niveles en las líneas telefónicas [16], fue corregido en 1953, redefinido como estándar IEC N° 268-17/1990 e IEEE/ANSI N° 152/1991 y actualizado finalmente en 1999. Es esencialmente un galvanómetro que pretende medir el valor eficaz de la señal de audio, rectificándola y obteniendo el valor medio. [3, 11, 20]

Su escala cubre desde  $-20$  dB hasta  $+3$  dB, que fue el rango de niveles aceptados como útiles, referidos comúnmente como  $-20$  VU y  $+3$  VU. Como se observa en la Figura 2.2



**Figura 2.2** Escala de un vúmetro clásico

Fuente: <http://es.wikipedia.org/wiki/Vumetro>

La unidad de volumen (VU) se define como: "El indicador de volumen marca 0 VU cuando se conecta a la salida una resistencia de 600 ohms, para una señal sinusoidal de 1000 Hz." [16]

La marca de 0 VU corresponde a la modulación del 100% del transmisor. [3]

El vúmetro tiene una impedancia propia de 7.500 ohms para tener la mínima carga sobre una línea de 600 ohms. [3] Se calibra con una señal sinusoidal estable para que marque 0 dB cuando esta aplicado el valor *SOL* (+8 dBm o +4 dBm). [2]

Sus características dinámicas se determinaron en base a la aparente respuesta del oído humano, de manera que si se aplica una señal sinusoidal que exija un salto de +20 dB, el instrumento tardará 0.3 segundo en estabilizar su deflexión, con un amortiguamiento crítico. [22, 26] La señal que produce este efecto está comprendida entre 40 Hz y 10 KHz, con un tiempo de crecimiento de 0.3 segundos, lo que introduce un efecto de enmascaramiento, por lo que el instrumento es incapaz de reflejar correctamente señales de audio complejas y transitorias de rápido crecimiento [21]

### 2.2.6 Niveles de referencia

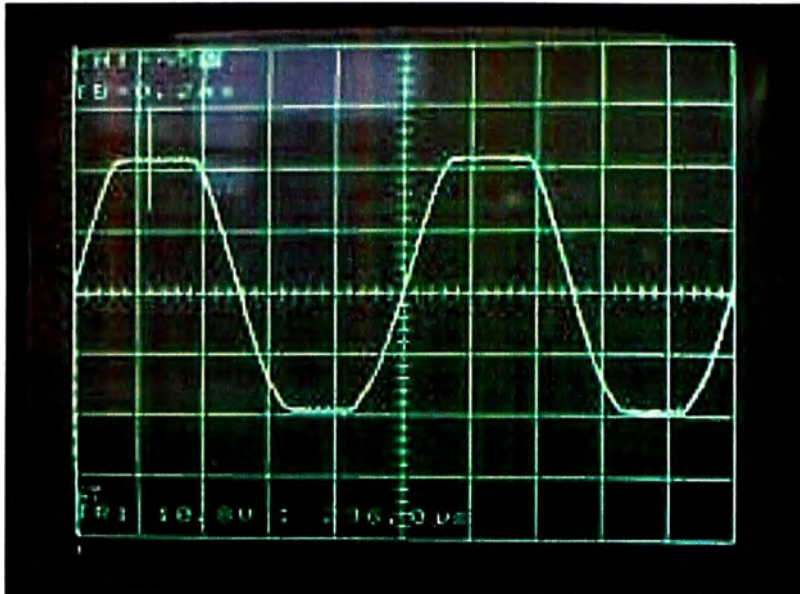
El nivel original fue el mW (0 dBm) y el nivel de salida para 0 VU de +4 dBm o +8 dBm sobre una línea de 600 ohms de impedancia. [2] Como actualmente se utilizan fuentes de señal con impedancias cercanas a cero, se hace innecesario equiparar impedancias para cortas distancias, por lo que se popularizó el dBu, indicando unidades de volumen sonoro, pero referidas a la tensión y no a la potencia

Utilizando impedancias de salida de menos de 100 ohms y de carga mayores a 10 Kohms, la referencia, para la indicación de 0 VU, no siempre es de 0.775 V eficaces, en un equipo doméstico es común encontrar -10 dBm de referencia, por lo tanto un vúmetro indicando un determinado nivel no garantiza conocer la correspondiente salida de tensión. [5, 11]

El uso indiscriminado de indicadores de nivel usados como vúmetros (emulando a los *SVI*), en equipos no estandarizados, hace que la marca de 0 VU solo indique que el equipo está en el máximo valor promedio, pero no garantiza la compatibilidad con otros equipos. Entonces la mala interpretación de este valor lleva a que los usuarios saturen la señal por exceso de amplitud, resultando en un recorte. [23]

Esto acontece en el ámbito analógico, en el ámbito digital el máximo nivel posible corresponderá al máximo valor binario disponible y la relación entre éste y su conversión

a una tensión analógica resultante, por lo tanto dependerá de los conversores A/D y D/A utilizados.



**Figura 2.3** Señal recortada en amplitud

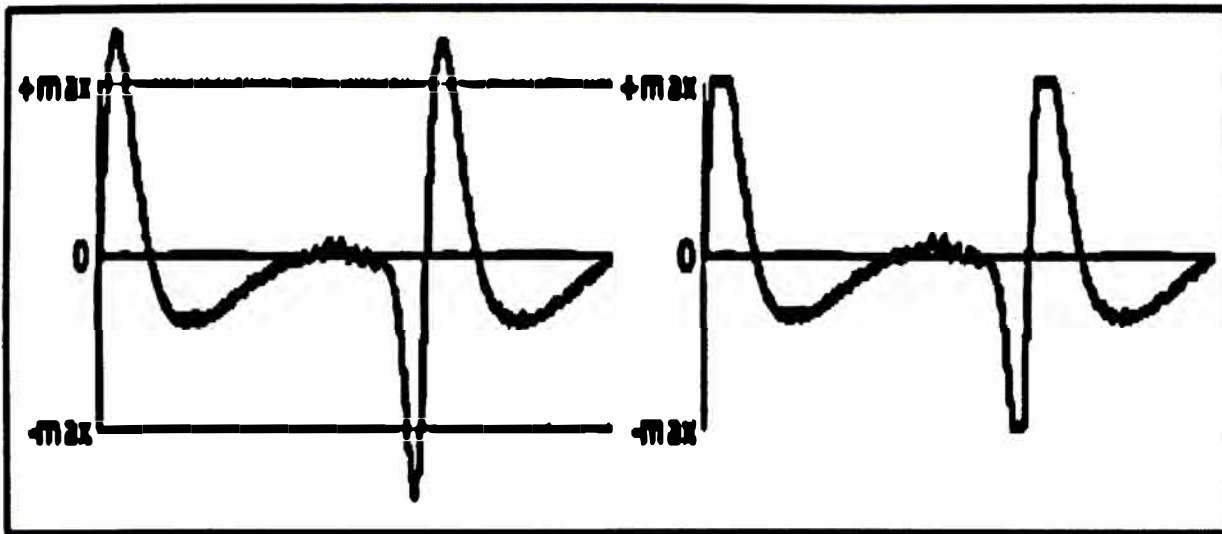
**Fuente:** <http://www.forosdeelectronica.com/>

En el dominio digital este máximo valor debe ser "inalcanzable" por las crestas del material de programa, ya que tiene mayores consecuencias que el dominio analógico, aquí se presenta el dilema de cuánto hace falta disponer por sobre el máximo valor medio de una señal, para que no se produzca un recorte por sobre-rango binario

Tomado de la experiencia analógica, un valor de por lo menos +18 dB es lo mínimo indicado, por ello se estableció que en un equipo digital el valor de referencia debe estar a 18dB del llamado *Fondo de Escala Digital* (Digital Full Scale = DFS o simplemente FS), para indicar el máximo valor posible, que se identifica como 0 dBFS.

El valor de referencia de 0 VU fijado en Europa está en -18 dB del valor FS (EBU R68), la tendencia en EEUU es definir esta valor a -20 dBFS (SMPTE RP-155). [21]

El margen sin recorte, por sobre el valor de referencia, se llama *Headroom*, (que corresponde a los +18 dBu en el mundo analógico), pero un dispositivo debe tener una salida máxima aun mayor, ya que los +18 dB son referidos a una onda senoidal y no al valor de la cresta. La relación entre el valor eficaz de una onda senoidal y su valor de cresta (llamado factor de cresta) es de 3 dB, pero en señales de audio normalmente llega a los 10 dB, este valor por sobre el eficaz se suele llamar *Peackroom*, indicando el margen disponible para las crestas, hasta la saturación de la salida [20]



**Figura 2.4** Recorte de los picos de señal por la falta de margen

Fuente: <http://cycling74.com/docs/max5/tutorials/msp-tut/mspdigitalaudio.html>

En un equipo digital es normal indicar una referencia de  $-20$  dBFS, esto es 20 decibeles por debajo del máximo posible, pero esto involucra la máxima salida posible se señal eficaz y no al nivel de las crestas de señal. [23]

La máxima y absoluta marca debe estar 3 dB por debajo, dando uno  $-23$  dBFS para el valor de referencia respecto al máximo valor binario (EBU – R128). [44]

Para identificar la amplitud de las crestas de señal, existe otro instrumento de medición, denominado medidor de crestas o picos.

### 2.2.7 Medidor de crestas o picos

El medidor de picos de programa o medidor de crestas, es un instrumento que indica el máximo valor absoluto de la señal y su función principal es garantizar que esta no sobrepase la máxima amplitud disponible, además de ajustar la amplitud por el valor medio de la señal, es importante asegurar la no saturación de las etapas electrónicas, sean estas analógicas o digitales. [11, 24, 25]

Los medidores de crestas analógicos indican de manera directa la amplitud y deben estar calibrados para corresponder con la máxima amplitud permitida en un equipo, puede constar de un instrumento de aguja o una barra de LEDs y toman la señal rectificada filtrando el valor máximo, posee un tiempo de crecimiento muy corto para incluir los rápidos transitorios de una señal y un decaimiento más lento para permitir ver

este valor. [11, 20, 24]. A estos instrumentos se les conoce como PPM (Peak Program Meter) y acompañan las lecturas de los vúmetros en las aplicaciones profesionales. Un operador decidirá con la lectura del VU y del PPM como ajustar el valor del nivel de programa, así también un sistema de compresión, en función de este valor, puede variar su ganancia y así evitar la saturación del equipo [11, 25]

Cuando se convierte una señal analógica a una digital, es esencial conocer cómo se relacionan los niveles de la señal analógica con la resolución digital disponible, básicamente hay dos estándares que describen esta conversión, el norteamericano SMPTE RP-155 y el Europeo EBU R-68, la diferencia entre ellos es de 6 dB. [19, 22]

En los primeros conversores A/D si la señal analógica superaba la máxima escala del conversor, aparecían números binarios aleatorios, hoy en día los conversores pueden manejar esto manteniéndose, durante la sobrecarga, al máximo nivel de escala (como si fuera una saturación analógica), pero esto crea una gran componente de tercera armónica posterior al filtro anti-aliasing, comúnmente utilizado en la entrada del conversor para cumplir con el teorema de Nyquist. Estos componentes de alta frecuencia dentro del flujo de datos complican y generan errores en las etapas digitales. [4, 22]

Se desprende de esto que es realmente muy importante no sobrepasar la máxima escala de un conversor A/D, para que la señal digital resultante sea consistente y una vez digitalizada la señal tampoco está exenta de ser saturada por un procesamiento incorrecto de los datos. [22]

### **2.2.8 El gálibo y la franquía de una señal de audio**

En español, podríamos utilizar el término **gálibo** de origen estructural y utilizado también en radio frecuencia, que identifica a las cotas que señalan las dimensiones máximas permitidas para el paso seguro por un túnel o debajo de un puente, se refiere también a las bandas permitidas en un filtro; así aplicado al audio digital, el gálibo (vertical) de una señal, sería el valor que puede alcanzar para pasar por un equipo sin "tocar" su máxima excursión. Sería el máximo rango dinámico permitido de una señal (dado por el LR en la norma EBU R-128). En algún caso podría tomarse equivalente al término '*Headroom*' del inglés.

También podemos tomar de la náutica el término **franquía**, como índice de que se tiene paso franco para tomar determinada vía, correspondiendo a la distancia entre la



parte inferior de la superestructura de un puente y el nivel de la máxima creciente conocida (Dicc. Larousse). Sería entonces el margen con que una señal pasa por una etapa en el momento de su máxima cresta y sería equivalente al término 'Peakroom' del inglés.

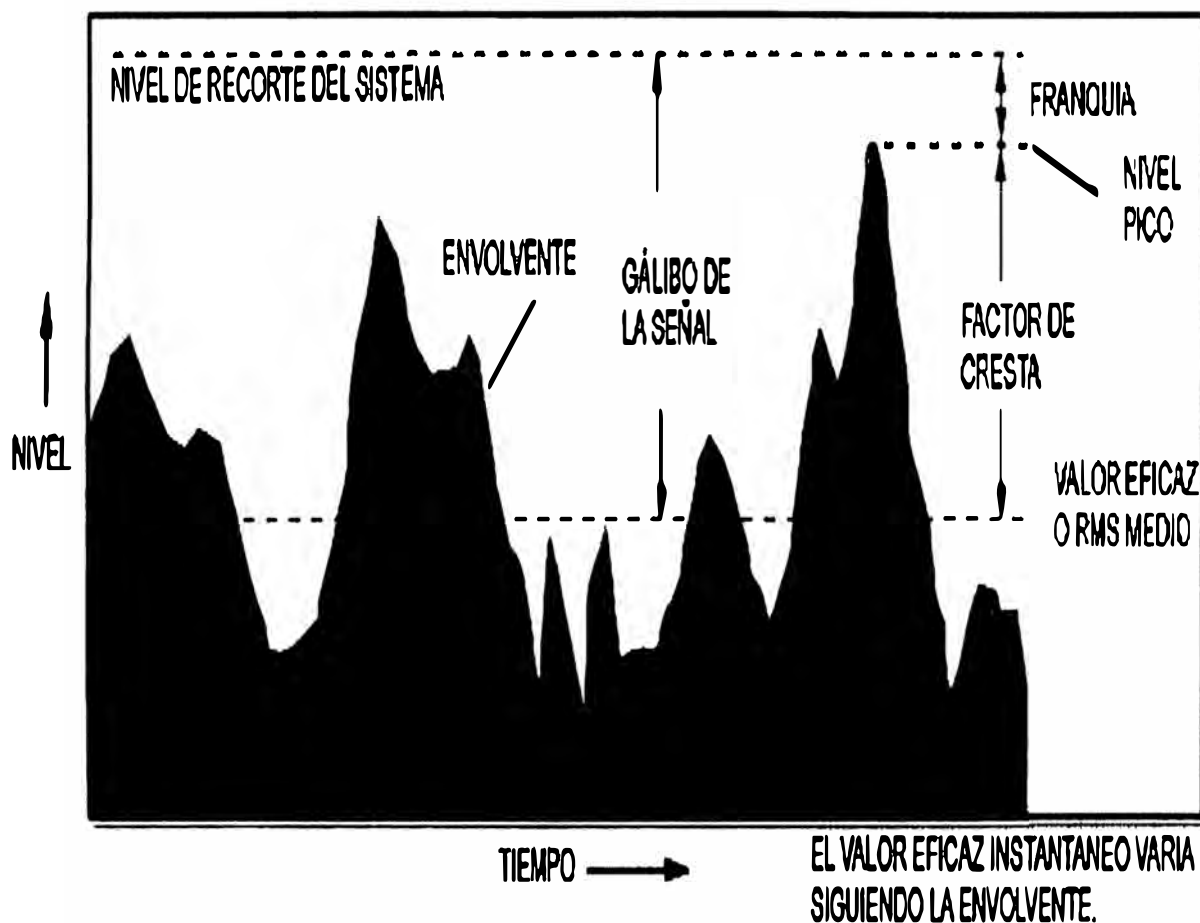


Figura 2.5 Gálibo y Franquía de una señal de audio

Fuente: <http://www.prosoundweb.com/>

### 2.3 El Loudness

Según la definición tomada del American National Standards Institute (American Psychoacoustical terminology, S3.20 de 1973), el Loudness o Sonoridad es el atributo que nos permite ordenar sonidos en una escala que va del más débil al más fuerte. [26]

Se podría decir que el Loudness o Sonoridad es la característica de un sonido que es principalmente un correlato psicológico de la intensidad física (amplitud). Más formalmente, se define como "el atributo de la sensación auditiva en términos de que los



sonidos pueden ser ordenados en una escala que se extiende desde tranquila a carcajadas. [34]

Por ser la sonoridad una medida subjetiva, esta varía para cada persona, ya que la impresión de la intensidad sonora es única, de esta manera, los equipos electrónicos usados en la producción y difusión de material de audio permiten variarla a voluntad, por lo que obtener una medición confiable se hace primordial para poder normalizar los niveles de una señal, la interrogante es, si se puede medir la sonoridad (loudness) y como debe hacerse esta medición.

Las investigaciones sobre la evaluación de la sonoridad están recibiendo un gran impulso como resultado de la transición hacia los sistemas de radiodifusión digitales. Hoy en día gran parte o toda la cadena que va desde el artista o locutor, al oyente, se realiza en formato digital, la globalización en la generación de contenidos también ha traído problemas con los niveles de sonido, estos problemas se producen cuando una producción está ajustada a los estándares de un país y llega a otro, puede que en este último se inserte material que tendrá los niveles locales, medidos con diferentes estándares y medidores. [4, 38]

Es común en el mundo de la radio, la consigna de que “mayor volumen, es lograr más oyentes” y en el de la producción discográfica que un mayor nivel da más “sensación de presencia musical”, por ello se tiene cada vez más volumen y señales más procesadas reduciendo su rango dinámico dentro de una cadena digital, como consecuencia, el oyente tiene cada vez menos control. Esto se denomina “*Loudness War*”. [4, 22, 36, 38,44].

En los trabajos de investigación se utiliza comúnmente el término “sonoridad percibida” o “sonoridad subjetiva”. ¿Será la “sonoridad percibida” una tautología retórica? [27]. Sabemos que en la retórica, una **tautología** (del griego *ταυτολογία*, decir lo mismo) es una afirmación obvia o redundante, aunque la redundancia suele entenderse como una falta de estilo, a veces se utiliza intencionadamente para dar énfasis, en este sentido también se puede llamar pleonasma, que en ocasiones, puede tener la función instructiva de reforzar la memoria de un tema. En este trabajo se usará el término sonoridad, sabiendo que su evaluación siempre será subjetiva por parte del oyente.

En la Figura 2.6, se puede apreciar la tendencia de incrementar el volumen observando esta secuencia de onda de la canción “Something” de The Beatles masterizado en CD cuatro veces desde 1983.

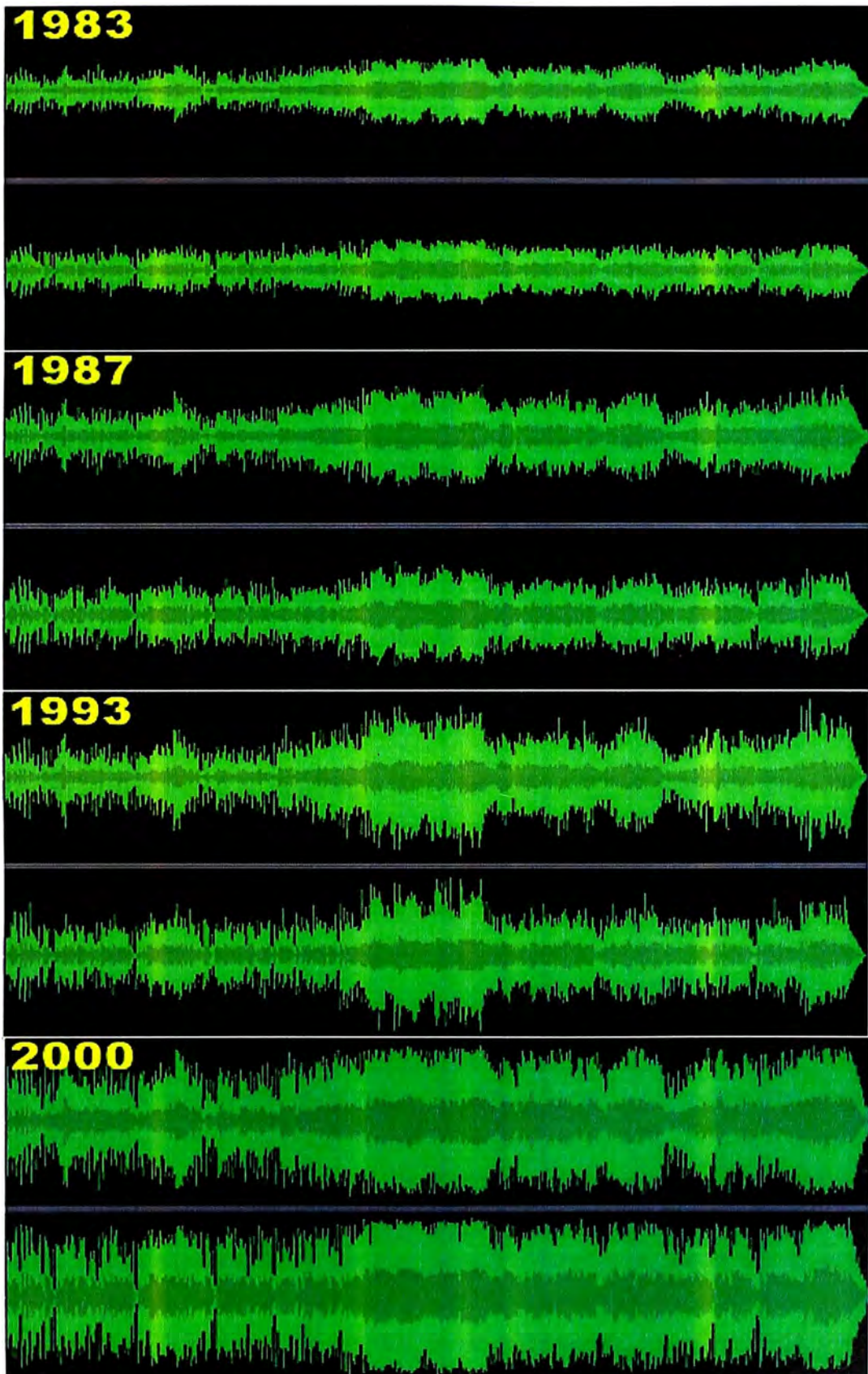


Figura 2.6 Tendencia de incremento del volumen

Fuente: [http://es.wikipedia.org/wiki/Guerra\\_del\\_volumen](http://es.wikipedia.org/wiki/Guerra_del_volumen)

### 2.3.1 Como interpretamos la sonoridad

Una característica primordial de nuestro sentido auditivo es la forma en que nuestro oído percibe la sonoridad y como las células ciliadas y las correspondientes fibras del nervio auditivo son excitadas en la membrana basilar del oído interno. [30]

Esta excitación está distribuida en bandas de frecuencias en la cóclea, formando un tipo de analizador de espectro biológico, cada frecuencia excita la zona de la membrana basilar y cada una agrega valor a la sensación de sonoridad total. Si dos sonidos arriban al oído con contenidos de frecuencias similares, ambos competirán por excitar las mismas células dentro del órgano de Corti, entonces estas células interpretan la información y segregan una sustancia química que será transformada en los impulsos eléctricos que los nervios auditivos llevarán al cerebro. [33] Este sistema tiende al equilibrio, y es por esta razón que una duplicación en la intensidad sonora no es percibida como el doble de nivel de sonoridad.

Por otro lado si dos sonidos conteniendo diferentes frecuencias llegan al oído interno, excitarán diferentes células y fibras nerviosas, por lo que aquí la suma de los dos sonidos dará el doble de sonoridad.

Un aspecto importante es el referido a la capacidad de diferenciar los tonos de un sonido, a bajas frecuencias se requiere cambios hasta del 3% en un tono para detectar la diferencia; en cambio para tonos de frecuencias mayores, variaciones del 0.5% bastan para distinguir la frecuencia. La capacidad de notar diferencias tonales depende también de la duración de los tonos y es en cierto grado independiente de la amplitud del sonido. [33]

La cóclea es un complejo instrumento mecánico del oído y uno de los temas de más estudio en la neurofisiología de la audición. En el análisis del escenario auditivo cotidiano difícilmente se trata con tonos únicos, normalmente el oído tiene que lidiar con conjuntos de diferentes frecuencias y amplitudes, con cadencias variables y superposiciones. Como en el oído interno existen importantes alinealidades en la respuesta de la cóclea a los estímulos sonoros, la no linealidad implica que la respuesta a un sonido complejo no es igual a la suma de las respuestas a cada uno de sus componentes, existen dos clases de efectos de superposición, los efectos de primer orden que se producen a nivel mecánico en la cóclea y los de segundo orden que resulta del procesamiento neuronal. Los tonos de combinación son, generalmente, combinaciones de los tonos originales, que no son importantes en los rangos normales

de sonido porque sólo se producen con estímulos intensos, pero son muy importantes en la audición de música a alto volumen, como sucede en los conciertos de rock o al escuchar música con audífonos a gran volumen. [31, 32]

El rango máximo entre las frecuencias de dos tonos que existen a las mismas fibras nerviosas se llama banda crítica, una manera de medir estas bandas ha sido propuesta por Zwicker en 1957. Estas bandas son angostas a bajas frecuencias (90Hz, para frecuencias por debajo de 200Hz) y más anchas a altas frecuencias (900Hz, a frecuencias de 5KHz), como fue estudiado por John Backus, en *"The Acoustical Foundations of Music"*.

El oído no tiene una respuesta plana en frecuencias, tiene su propia curva de respuesta, Fletcher y Munson en 1933 obtuvieron experimentalmente que la respuesta del oído varía con la intensidad sonora, es decir con la sensación de sonoridad. Estas curvas fueron ajustadas luego por Robinson y Dadson en 1957 y estandarizadas por la norma ISO 223:2003 y el estándar internacional para sistemas de sonido IEC 60268. [33]

En la figura 2.7 se muestran las curvas isofónicas que corresponden al nivel de presión sonora versus la frecuencia para un tono sinusoidal.

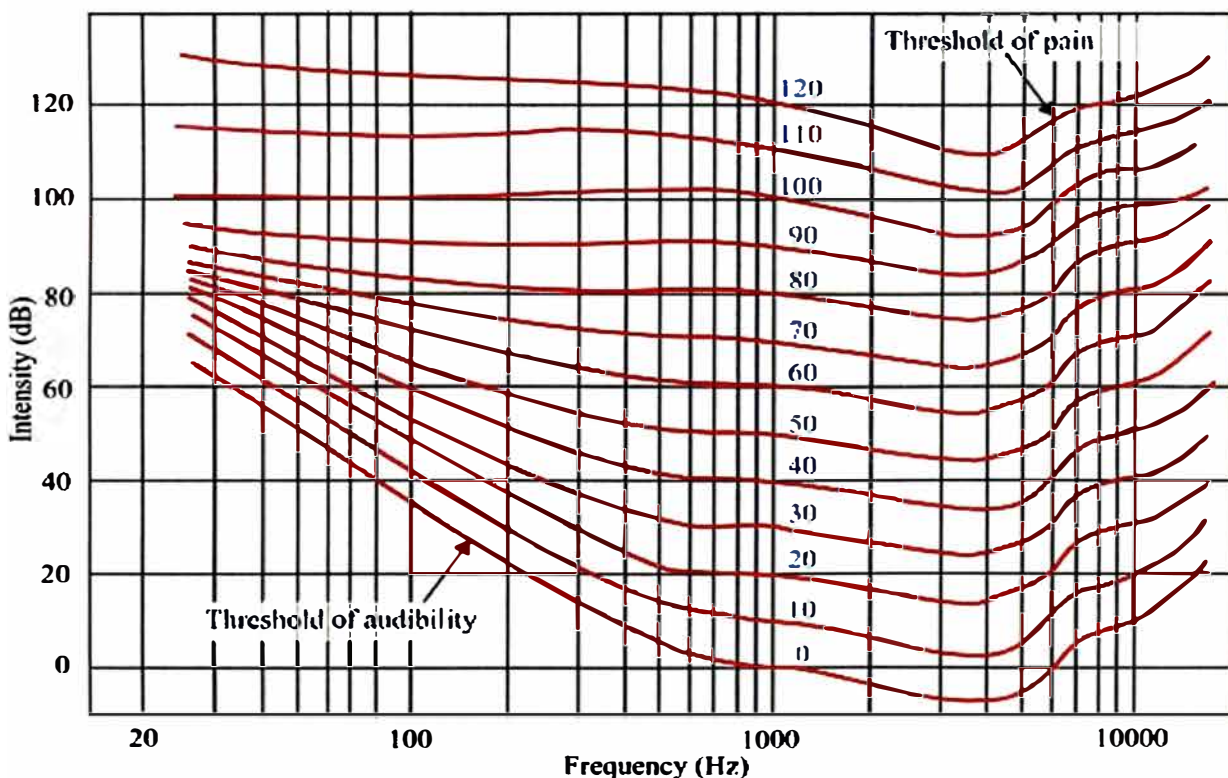


Figura 2.7 Curvas Isofónicas (ISO 226)

Fuente: [http://es.wikipedia.org/wiki/Curva\\_isofonica](http://es.wikipedia.org/wiki/Curva_isofonica)



Un perfil usado para medir la sonoridad está dado por la unidad Fon (Phon en inglés), que está definida como la sonoridad de una onda senoidal de 1 KHz con un nivel de presión sonora (intensidad) de 0 dB<sub>SPL</sub>. Así 0 Fon corresponde a 0 dB y 120 Fones corresponden a 120 dB para sonidos sinusoidales de 1 KHz.

De esta forma, la sonoridad (S) viene expresada por:

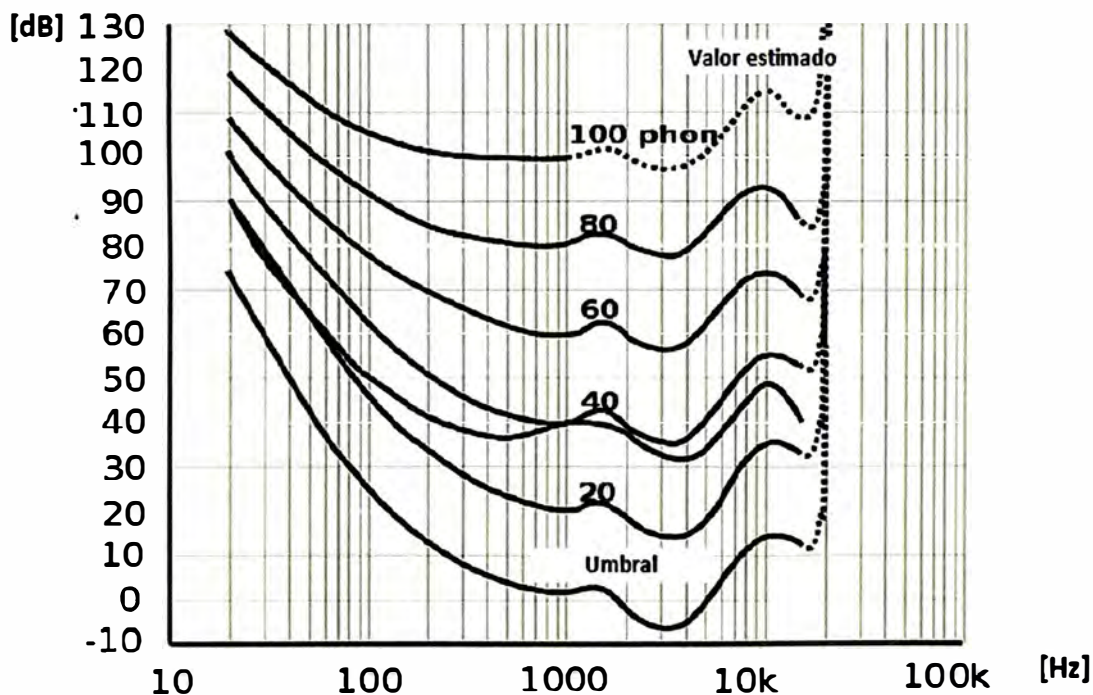
$$S = 10 * \log_{10} \left( \frac{I}{I_0} \right) \text{ Fones} \quad (2.6)$$

Donde:  $S$  = Sonoridad en Fones.

$I$  = Intensidad medida a una frecuencia.

$I_0$  = Intensidad de la curva del umbral de audición a esa frecuencia.

El Fon no se utiliza para comparar la sonoridad de dos sonidos diferentes, sino que da la referencia de sonoridad de un determinado tono, esto se debe a que la escala de Fones está directamente relacionada con la escala isofónica logarítmica, mostrada en la Figura 2.8.



**Figura 2.8** Intensidad expresada en Fones

**Fuente:** "Handbook for sound engineers"

La creación de estas curvas está basada en la respuesta subjetiva a tonos sinusoidales presentados frontalmente al oyente y sólo son válidas para un campo sonoro directo. Estas curvas no tienen en cuenta la diferencia en la percepción de los sonidos

que provienen de diferentes direcciones (campo sonoro difuso). Si se utiliza un ruido de banda angosta y si el sonido proviene de un campo difuso, estas curvas requieren una corrección.

Para modelar este funcionamiento del oído, se utilizan curvas de ponderación. Estas se sintetizan en los tres valores más usados:

- La curva A (curva del nivel de sonoridad de 30 Fones, en decibeles A  $-dB_A$ )
- La curva B (curva del nivel de sonoridad de 70 Fones, en decibeles B  $-dB_B$ )
- La curva C (curva del nivel de sonoridad de 100 Fones, en decibeles C  $-dB_C$ )

**TABLA N° 2.2** Transferencia de las distintas ponderaciones

Fuente: "Redes de Ponderación acústica"

Frecuencia [Hz]	Ponderación [dB]		
	A	B	C
31.5	-39.4	-17.1	-3
40	-34.6	-14.2	-2
50	-30.2	-11.6	-1.3
63	-26.2	-9.3	-0.8
80	-22.5	-7.4	-0.5
100	-19.1	-5.6	-0.3
125	-16.1	-4.2	-0.2
160	-13.4	-3	-0.1
200	-10.9	-2	0
250	-8.9	-1.3	0
315	-6.6	-0.8	0
400	-4.8	-0.5	0
500	-3.2	-0.3	0
630	-1.9	-0.1	0
800	-0.8	0	0
1000	0	0	0
1250	0.6	0	0
1600	1	0	-0.1
2000	1.2	-0.1	-0.2
2500	1.3	-0.2	-0.3
3150	1.2	-0.4	-0.5
4000	1	-0.7	-0.8
5000	0.5	-1.2	-1.3
6300	-0.1	-1.9	-2
8300	-1.1	-2.9	-3
10000	-2.5	-4.4	-4.4

La ponderación A es la más utilizada para la valoración del daño auditivo e inteligibilidad de la palabra, se emplea para analizar sonidos de baja amplitud y es la que universalmente se utiliza como referencia en las leyes y reglamentos contra el ruido producido a cualquier nivel. Los niveles de la presión sonora se expresan en dBa. [1, 10]

La ponderación B es la que corresponde a intensidades medias de sensibilidad del oído, es la que se toma como referencia para las mediciones de música. (dBb). [1, 10]

La ponderación C modela la respuesta del oído ante sonidos de gran intensidad, tiene una respuesta muy plana por lo que se utiliza para medir equipos de sonido y para la evaluación de sonidos de baja frecuencia. (dBc) [1, 10]

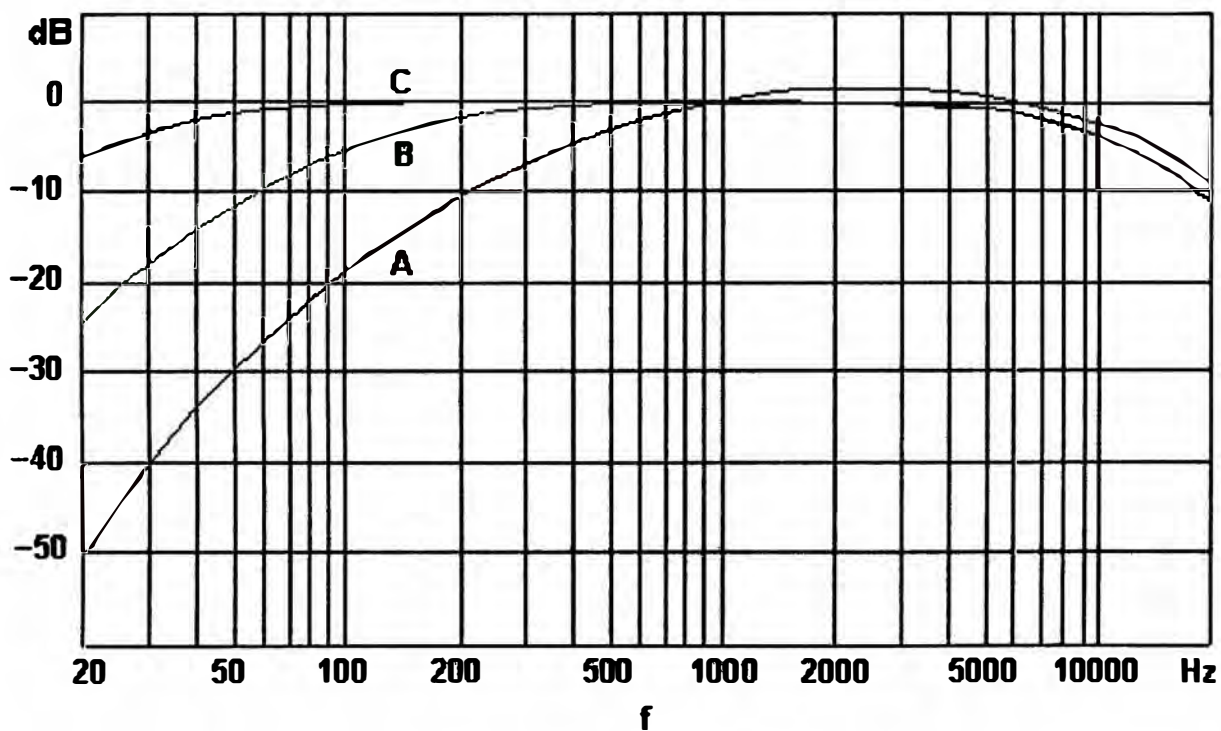


Figura 2.9 Curvas de Ponderación A, B y C

Fuente: <http://www.fceia.unr.edu.ar/acustica/biblio/niveles.htm>

### 2.3.2 Medición del nivel de señal

Tradicionalmente los instrumentos analógicos utilizados para medir el nivel de las señales de audio fueron los medidores de aguja. En el trabajo de H. A. Chinnet *al.* Se ve el primer diseño de un medidor de este tipo, que mide el valor de la señal imitando la respuesta del oído humano [3], pero como no oímos todas las frecuencias con la misma intensidad, su respuesta en frecuencia no debería ser plana; sin embargo, ya los equipos

electrónicos de audio tienen un máximo de señal manejable, independientemente de la frecuencia, todas las crestas deben ser medidas, aún las menos audibles, si se desea evitar la saturación.

En este diseño quedó determinado que la señal óptima era muy cercana al máximo del sistema, evitando en lo posible el ruido, esto trajo como contrapartida que no se dispusiera de mucho rango de variación en los niveles sonoros.

Como se ha visto en la tabla 2.1 y en la figura 2.8, el rango de audición normal del oído humano está entre los 20 dBspl y los 100 dBspl, rango imposible de alcanzar por medios emisores analógicos, cuando se fijaron las normativas de radiodifusión en 1939.

Se logró tener un rango técnicamente optimizado de sólo 20 dB, comprimiendo el rango de la música desde unos 60 dB para conciertos y 40 dB para la música popular, en realidad menos de 40 dB no dan una certera representación del sonido en vivo, pero este rango es demasiado amplio para las transmisiones de radiodifusión. [19, 22]

## **2.4 Evaluación del Loudness**

La sonoridad es la magnitud de un sonido expresada por el “volumen” que es percibido por el oyente y no depende solamente del nivel de intensidad establecido, sino también del formato y fuente sonora del material de audio en particular.

### **2.4.1 Fisiología auditiva**

El proceso de recepción de las ondas sonoras, como se ha mencionado, es un proceso complejo en el que se ven envueltos efectos fisiológicos y psicológicos.

En cuanto a la fisiología auditiva del oído, se puede decir que consta de tres partes:

- Oído externo
- Oído medio
- Oído interno

Como se puede observar en la Figura 2.10, el pabellón auditivo (en la oreja) capta las ondas y las hace converger al conducto auditivo externo, las ondas llegan al tímpano, que recoge las vibraciones de presión y las transmite al sistema de huesecillos (en el oído medio), la vibración pasa a la ventana oval que conecta con el oído interno, donde se encuentra el caracol, allí están las células nerviosas que forman la membrana Basilar



(que resuena en un punto específico para cada frecuencia) y se encargan de traducir la excitación mecánica en un impulso nervioso, que llega al cerebro a través del nervio acústico, dándonos la sensación de escucha. [18]

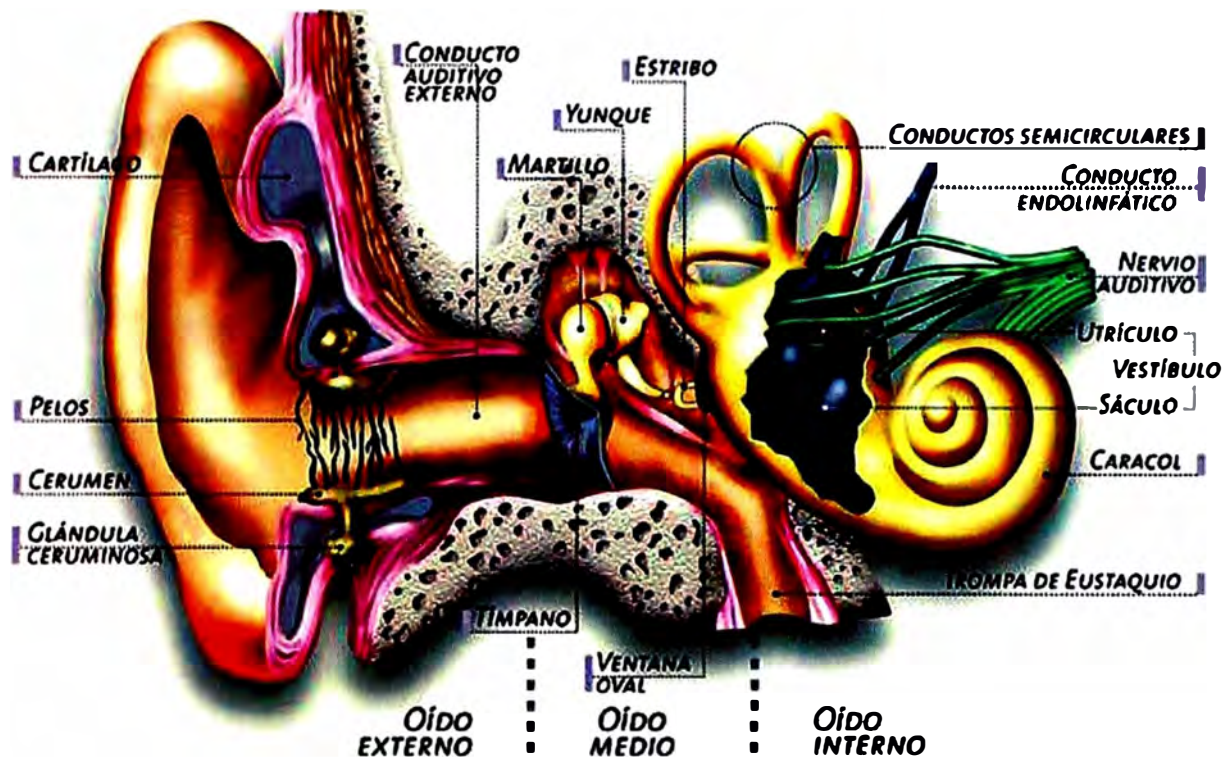


Figura 2.10 Partes del oído

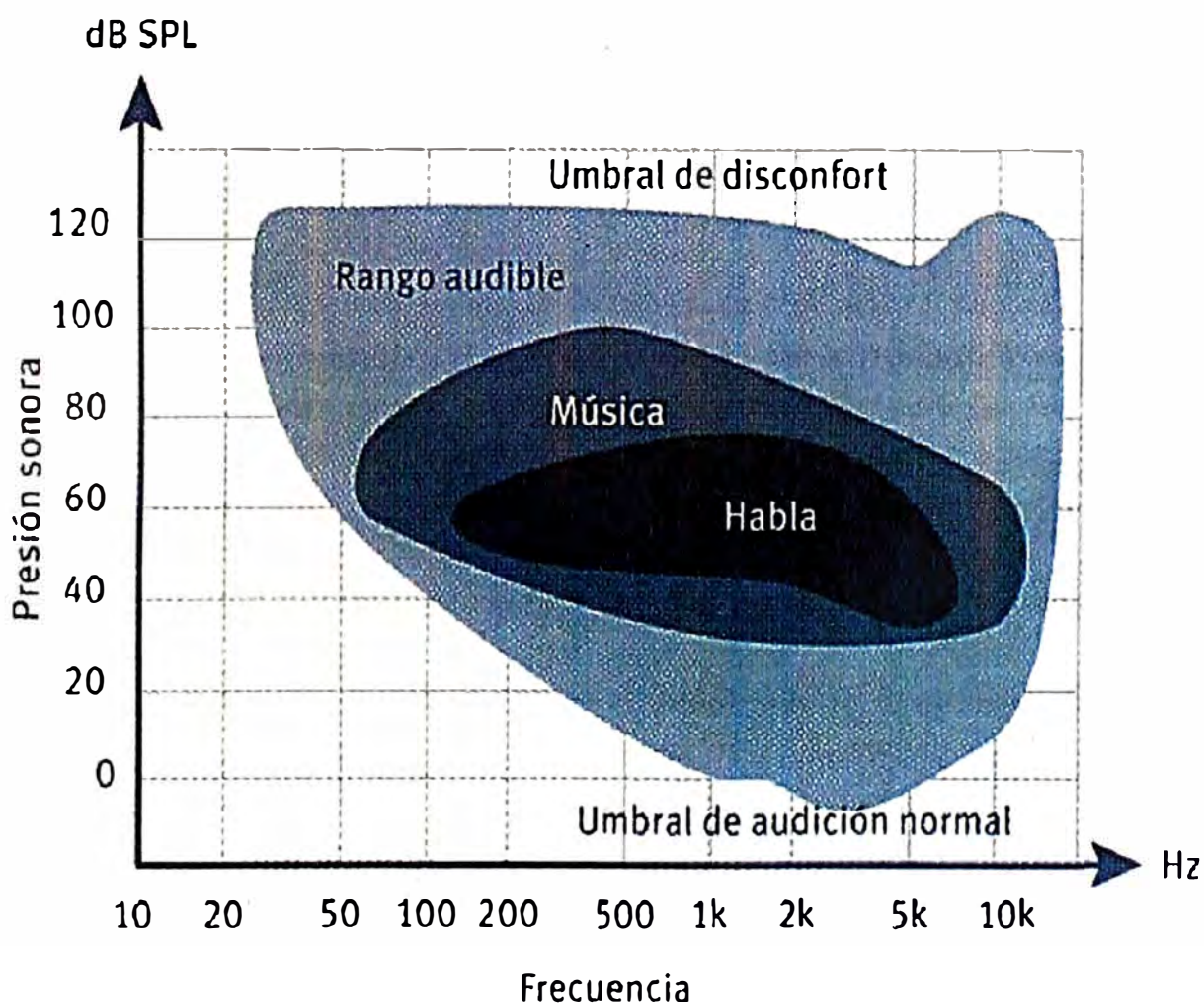
Fuente: [http://usuarios.discapnet.es/ojo\\_oido/el\\_oido.htm](http://usuarios.discapnet.es/ojo_oido/el_oido.htm)

La membrana Basilar es un filtro mecánico que se encarga de descomponer el sonido en sus componentes espectrales; esta descomposición no es tan precisa cuando el sonido es complejo y si tiene componentes muy próximos estos se escuchan como si fueran uno sólo. Al rango de banda mínimo que es capaz de discernir el oído se le llama ancho de banda crítico, pero como los sonidos se interpretan con distintos y complejos umbrales auditivos, en esta respuesta grosera, es donde comienza la respuesta subjetiva del oído o sensación sonora frente a los diferentes sonidos. Las señales que se producen en los vellos de la membrana Basilar, son conducidas al cerebro a través del nervio auditivo, las células pertenecientes a la membrana basilar y que están en contacto con el líquido cloqueal, son las que se ven mayormente afectadas cuando hay exceso de ruido, primero perdiendo los cilios y posteriormente degenerándose si el ruido persiste.

Nuestro sistema de percepción sonora detecta cambios en la presión del aire y sobre estos, hace interpretaciones, se sabe que el oído interno, la cóclea, divide la señal

de audio en bandas de frecuencias y genera señales neuronales basadas en la presencia de componentes dentro de cada banda. Dos de los procesos utilizados para evaluar estas señales neuronales son el análisis gramático y la segregación del flujo, el análisis permite identificar los indicios temporales de que hubo un evento sonoro, la segmentación del flujo es la división de la señal audible en capas, cada capa consiste en un flujo de sonido que proviene de un evento común, la segregación es primordialmente un proceso de agrupación de frecuencias, el proceso final en la identificación categórica de un sonido depende de ambos parámetros y de la pasada experiencia auditiva del oyente. [18, 33]

El área del campo audible limita, a bajos niveles de presión, con el umbral de la audición y a niveles muy altos, con el umbral de malestar y sensación de dolor, mostrado en la figura 2.11.



**Figura 2.11** Rangos de sonoridad

Fuente: <http://audifonos.pro/tipos-y-causas-de-perdidas-auditivas-2>

### 2.4.2 Aspectos de la sonoridad

La sonoridad es subjetiva por naturaleza y por lo tanto más difícil de medir que las propiedades físicas de la materia, por lo que se propone la utilización de un método visual y estadístico para evaluarla tomando en cuenta la incertidumbre y subjetividad de los datos de referencia, así como otros datos importantes, tales como la operación en tiempo real, el nivel de referencia y las características de los medidores a utilizar. [10]

Muchos de los modelos de sonoridad fueron originalmente contruidos para medir señales estáticas, así como el ruido, pero hay modelos que permiten estimar la sonoridad percibida tanto en música como en la palabra; podríamos sintetizarlos como: la integración sobre el eje de la frecuencia perceptual, el enmascaramiento, la adaptación, la compresión sonora y la integración en el eje del tiempo. [26]

Normalmente el material de programa que es difundido en la radio, televisión o distribuido en CD, ha sido dinámicamente y espectralmente procesado, para cumplir con pautas estéticas y técnicas, este procesamiento afecta la sonoridad percibida del material. Los oyentes pueden advertir saltos indeseables en la sonoridad entre diferentes fuentes o canales de televisión y entre diferentes programas o segmentos dentro del mismo canal.

Si la sonoridad puede ser modelada y entonces predicha, es posible medirla y controlarla antes que llegue al oyente, la percepción de sonoridad ha sido extensamente investigada en la psicoacústica bajo condiciones de laboratorio, tradicionalmente por décadas se han usado señales estacionarias o sintéticas, sin embargo las llamadas mediciones objetivas de la sonoridad y sus procedimientos continúan bajo un continuo desarrollo. No hay modelo perfecto que mida exactamente la sonoridad de un material de audio, sea este de música, texto o comerciales emitidos por los canales de radio y televisión. [10] Por lo tanto, en conexión con las producciones de radiodifusión, las aplicaciones de nivel han sido estudiadas tanto para su uso en monitoreo como en nivelación, con la aplicación de compresores y procesadores del rango dinámico, se logró una maximización agresiva de la sonoridad, usada en la masterización de CD de audio, esta práctica mermó la calidad de sonido y llevó a varios problemas técnicos. [4, 22, 36] Un modelo ajustado de la sonoridad puede ser deseable también para otras aplicaciones, como la investigación de audibilidad, donde un conjunto de estímulos sonoros puedan ser evaluados independientemente de su nivel sonoro subjetivo, llevando a cabo experimentos controlados de escucha [10]

La sonoridad de segmentos homogéneos con una duración de 10 a 15 segundos, pueden ser comparados y se pueden definir niveles relativos de referencia, una

característica de la percepción es la sonoridad a largo plazo, aunque en algunas aplicaciones son deseables mediciones a corto plazo, sin embargo se han encontrado divergencias en las mediciones por lo que la mayoría de los estudios se realizan a largo plazo. [10]

### a) Percepción de la Intensidad

Un aspecto básico, es que la función de transferencia desde la magnitud física del sonido al estímulo nervioso que llega a nuestro cerebro no es lineal, una buena (pero no única) aproximación a esta transferencia es la función expresada como *ley de la potencia* presentada por Stevens, en 1957, donde se describe que las características de la función *potencia* se aplican a la magnitud física Intensidad ( $I$ ) para obtener la magnitud Sensación ( $S$ ).

$$S = K * I^n \quad (2.7)$$

En esta fórmula  $n$  es la potencia obtenida empíricamente del tipo de estímulo particular y para la percepción sonora vale alrededor de 0.3;  $K$  es una constante adaptativa a la unidad, utilizada en el juicio subjetivo. [10]

Esto se aplica a sonoridades por encima del umbral de audición, para bajas frecuencias, la potencia requerida es mayor, significando que un cambio en la intensidad sonora se percibirá como mayor al que sucediera con igual intensidad a frecuencias más altas.

La ley de la potencia puede ser reformulada usando logaritmos:

$$\log S = n * \log I + \log K \quad (2.8)$$

Como un incremento de 10 veces en la magnitud física, representa el doble en la magnitud percibida, esto evidencia un tipo de compresión en la función orgánica del oído. [10] Esta correspondencia entre las magnitudes físicas y las percibidas es una de las motivaciones para la utilización universal de la escala en decibeles

### b) Unidades de evaluación sonora

La sonoridad percibida puede expresarse en *Sones* o en *Fones*, la escala en *Sones* es una escala de sensibilidad ajustada, por lo que los valores expresados están directamente proporcionado con la sonoridad percibida, [30] mientras la escala en *Fones* expresa la sonoridad en su magnitud física, de esta manera coincide con la escala en decibeles para 1 KHz, a más de 40 dB de presión sonora y puede ser conveniente, o precisa, para indicar diferencias de niveles de sonoridad.

### c) Propiedades Espectrales

Existe una estrecha correspondencia entre la frecuencia del sonido y su sonoridad, la fuerte dependencia de esta función con la frecuencia, es expresada comúnmente como *contorno de igual sonoridad*. (Estándar ISO: 226, 1997). (Fig. 2.5)

El sistema auditivo humano procesa las señales descomponiendo las frecuencias del sonido en bandas estrechas y luego las ensambla, estas son las bandas críticas, su análisis muestra que hay entre 26 y 32 bandas, con una separación de  $\frac{1}{4}$  de octava por arriba de los 700 Hz y 100 Hz por debajo de los 700 Hz. [29]

Alternativamente las constantes  $n$  y  $K$ , antes presentadas, pueden ser consideradas como funciones de la frecuencia, por su propiedad de sumatoria espectral, lo que significa que la sonoridad percibida aumenta con el ancho de banda del sonido, aún a una potencia total constante; este efecto se aprecia en las diferencias de medición de sonoridad para un tono puro y para un ruido de banda ancha (ruido blanco) cuya diferencia es de 11 a 18 dB dependiendo de nivel de presión sonora. [30]

Zwicker en 1957 mostró que la sumatoria espectral está asociada con el concepto de banda crítica, dentro de cada banda crítica la sonoridad de una señal a un nivel dado es independiente de su ancho de banda, cuando el ancho de banda de una señal crece más allá de la banda crítica, la sonoridad aumenta; aún si el nivel de potencia total se mantiene constante. Aquí se produce una integración en el eje de frecuencia, especificado en unidades dependientes de una particular banda crítica, siendo las más comunes las unidades ERB (ancho de banda rectangular equivalente) presentadas por Hartman en 1998. Las señales de diferentes frecuencias también interaccionan de una manera compleja conocida como propagación, que está muy relacionada al concepto de enmascaramiento, el efecto de este fenómeno es que no sólo las frecuencias de la señal dentro de la banda crítica contribuyen a la sonoridad sino también alguna extensión sobre las frecuencias vecinas. Como consecuencia de esta propagación la sonoridad de los tonos puros y del ruido son diferentes, Como la música y la palabra se parecen más al ruido que a un tono puro, utilizar un tono puro en una medición de sonoridad da resultados errados [31]. Entre las escalas de decibeles y fones con tonos puros las escalas coinciden, pero con ruido se desvían por una causa de esta característica de la propagación. [30]

### d) Propiedades Temporales

Nuestra audición no reacciona instantáneamente al sonido, debido a que la sensación sonora requiere un tiempo, dentro de los límites y hasta 100 mseg, la

sonoridad es proporcional a la energía de un tono o un ruido, más allá de este valor la sonoridad de una ráfaga se percibe como una señal continua, [10] y cuando el sonido cesa, la percepción sonora tampoco cae instantáneamente, ya que existe un efecto de post máscara, que impide especificar una constante de tiempo sencilla para esta integración temporal, entonces la función sonoridad depende tanto del ancho de banda (ruido blanco, ruido de banda estrecha, onda senoidal) de la señal, como de la duración de la misma. [10]

Además existe el efecto temporal de largo plazo, llamado *incremento del umbral de audición*, al oír un sonido fuerte por un determinado tiempo, nuestra sensibilidad de audición se reduce, retornando al estado normal luego de unas horas.

#### e) **Propiedades Espaciales**

La percepción de la sonoridad también es influenciada por el filtrado relacionado a la posición espacial de la fuente del sonido, causado principalmente por el oído externo, ya que la dirección influencia nuestra percepción sonora, normalmente escuchamos con ambos oídos al mismo tiempo, lo que nos da una propiedad no trivial que es la *sumatoria binaural*, [10] al oír un sonido en una habitación, esta capacidad se enmascara por las características reverberantes de la habitación. [30] Este filtrado es complejo y depende de la posición de escucha, como es difícil de modelar se corrige con la utilización de fuentes multicanal. [31]

#### f) **Señales de prueba**

Para obtener una noción de los mecanismos psicoacústicos del sistema auditivo, se han empleado en las experiencias, señales de audio con propiedades analíticas. Como estímulos comúnmente se utilizan ondas senoidales constantes o pulsantes de amplitud variables, ruido de ancho de banda variable y tonos complejos. [31] Estos tonos están lejos de parecerse a un programa de audio de una estación de radiodifusión, tanto en términos de contenido espectral, como dinámico, ya que las señales del mundo real (como la música), tienen una gran fluctuación espectral y dinámica de nivel, entonces, cuanto más se conoce de los mecanismos de audición y su compleja naturaleza, el manejo de las señales del mundo real presentan un desafío a los modelos psicoacústicos tradicionales.

Las investigaciones académicas tienden a enfocarse en modelos sencillos, o en el *Bottom-top approach* para entender la función de transferencia de los mecanismos



individuales de la audición, comúnmente aislándolos. Las investigaciones realizadas por los radiodifusores y los fabricantes de equipos de audio se enfocan en contenidos del mundo real, lo que en ámbitos científicos se denomina *Top-down approach*, lo que hace muy difícil extrapolar sus resultados. [10]

### g) Tipos de Modelos de sonoridad

Los modelos de sonoridad pueden ser categorizados como mono banda o multi banda. Los multi banda separan la señal de entrada en varias bandas de frecuencia, la evalúan y la combinan en una sonoridad estimada, el modelo de sonoridad de Zwicker [30] es el modelo prevalente. Los modelos mono banda sólo tienen una banda pasante, sobre la que se realizarán los cálculos a partir de dos concepciones: los de nivel de sonido equivalente (*Leq*) y los medidores de cresta de programa o *PPM* (Peak Program Meter). [10]

Una cuestión clásica relacionada al modelado, es si el modelo del sistema auditivo debe reflejar la estructura interna del sistema auditivo o simplemente cumplir las funciones de transferencia entre las variables de entrada y salida [10]. Los modelos de la audición más complejos deben cubrir todo el sistema auditivo, desde el oído externo hasta los efectos cognitivos. [32]

Si en la definición de un modelos de sonoridad, sólo se toma un tipo de entrada, el modelo se puede simplificar dramáticamente; si por ejemplo, el análisis se limita sólo a tonos puros, la determinación de la función sonoridad por medio de experimentos de escucha y el subsecuente cómputo de los parámetros da resultados decentes. Pero un modelo basado en las propiedades de un tipo de señal de entrada puede ser apropiado para la palabra, pero no para la música, un modelo así, puede ser más simple que uno de usos generales, pero fallará cuando sea usado fuera de su alcance. [10]

Los primeros modelos fueron diseñados para señales estacionarias, estos son: sonidos que pueden ser descritos completamente por su contenido espectral, tal como ruidos y tonos complejos (modelo de Zwicker). Mientras los modelos contemporáneos tratan de predecir la sonoridad percibida para señales fluctuantes. [10] Entonces, aunque se conoce que el oído interno realiza un análisis de frecuencias sobre el sonido y que las frecuencias de la señal interactúan de una manera compleja, varios modelos se han construido sin un análisis de frecuencias, siendo mono banda. Lo que suele ser común a ambos tipos de modelos, es el dominio en que será evaluada la integración en el tiempo (como ser lineal, cuadrada o logarítmica). [10]

La relación entre el nivel de un sonido y su sonoridad, inicialmente tratada por Stanley S. Stevens, indicaba que una sensación del doble de sonoridad se alcanzaba con un aumento de 10 dB, posteriores investigaciones de Richard M. Warren encontraron que la diferencia de nivel era de sólo 6 dB, el psicólogo Jhon G. Beuhoff apreció que el nivel de crecimiento es más sensitivo que el de caída; entonces para el mismo nivel de diferencia, un cambio de sonoridad creciente es más notables que uno decreciente. Se ha sugerido que la escala de Sones manifiesta una polarización experimental y por lo tanto no representa una relación fundamental entre estímulo y sensación sonora. [33]

## 2.5 Fundamentos de la compresión dinámica

En el mundo analógico, la configuración tradicional de realimentación de las señales de control derivadas de la salida, fue utilizada antes del advenimiento de una conversión lineal – logarítmica precisa, esta configuración introduce una distorsión por modulación y no maneja bien los transitorios de la señal, a pesar de esto, aún se implementa en muchos sistemas de audio profesional. [29]

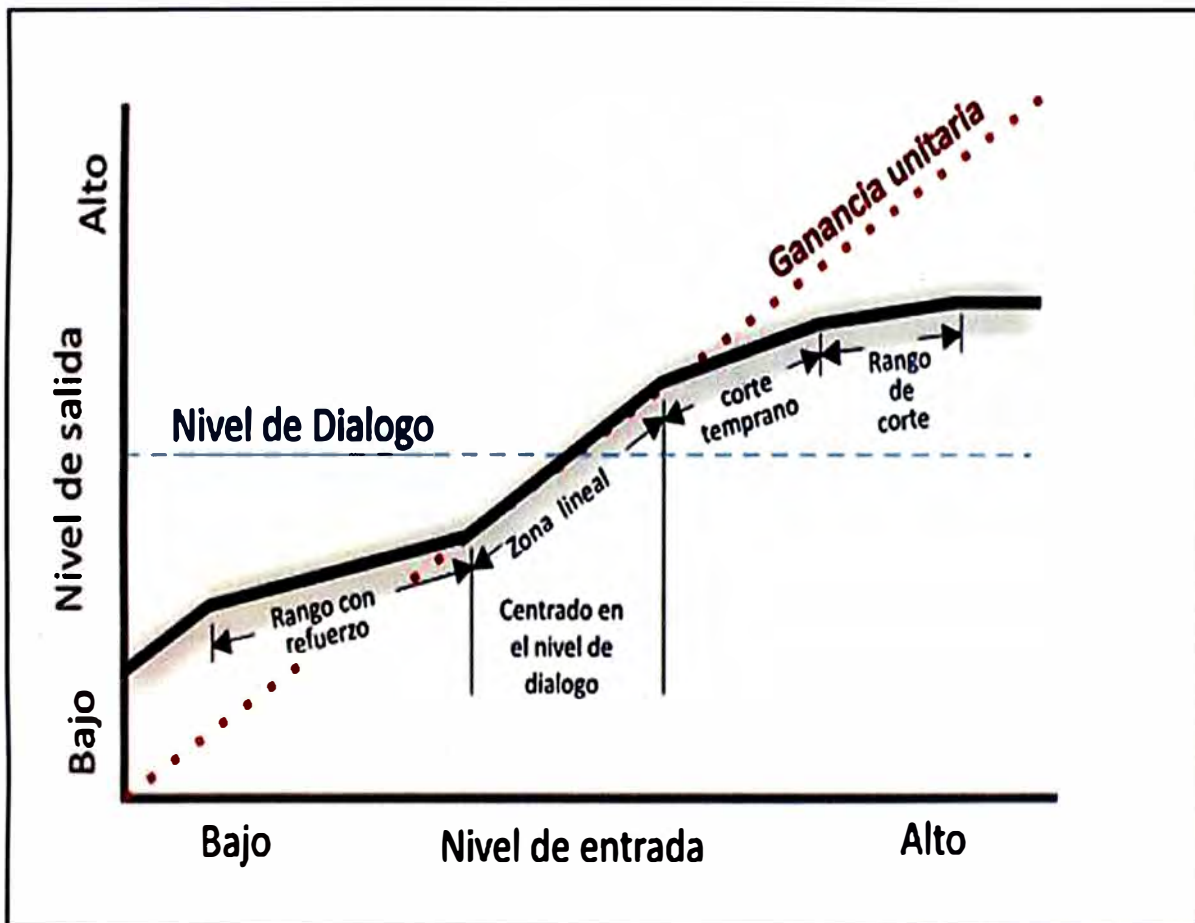
Una manera de minimizar la potencia computacional es “simplificar” las señales a procesar, esto se consigue, entre otras formas, disminuyendo su rango dinámico.

En su trabajo, *Schmith y Rutledge*, [29] desarrollaron un método de compresión de las señales que mantienen su claridad perceptual y por lo tanto un alto grado de transparencia que ha sido la base de posteriores desarrollos, para ello y ante la tarea imposible de alterar la amplitud de la señal mientras ésta está variando, han optado por permitir cierto grado de deformación y la aparición de anomalías o “artifacts” en su acepción de *Percepción de cualquier distorsión o error de datos generado por un dispositivo o técnica involucrada. (Diccionario Merrian / Webster)*, que serán enmascarados por la propia música. El algoritmo incluye una estructura multicanal de banda crítica, rangos de ataque y decaimiento, un control del modo de estimulación del nivel y una normalización estimada del nivel sobre las bandas de frecuencia. [29] Utilizando técnicas de escala multidimensional para realizar el análisis perceptual, estos evaluaron la similitud del estímulo musical entre varios pasajes musicales con distintos parámetros de compresión, lo que se utilizó para realizar un mapeado como indicador para el desarrollo de la métrica del nivel de compresión. [29]

Se pueden obtener una variedad de curvas, que permiten mapear los niveles de entrada sobre los niveles de las salidas (curvas I/O), los aspectos más importantes de



estas curvas son el rango de compresión, el umbral y la suavidad de la curva entre los segmentos lineales. La elección del rango de compresión y el umbral, determinan mucha la conducta del Compresor o Encoder; los diseños actuales están incluyendo una transición suave entre los segmentos, conocido con "codo suave" de manera de lograr curvas suaves en la función de transferencia, haciendo a ésta menos ríspida, como la usada en el sistema de compresión Dolby E, mostrado en la figura 2.12.



**Figura 2.12** Curva de Compresión Dinámica

Fuente: Medición de la Sonoridad del Audio Digital

Los Procesadores de Sonido Digital o DSP (*Digital Sound Processors*) ofrecen una opción para implementar este procesamiento por programación. Los compresores multicanal típicos consisten en un par de bandas y no tienen una buena correlación con las bandas críticas del oído, poniendo aberraciones en la envolvente de la señal. La investigación sobre la síntesis de instrumentos musicales ha encontrado que preservar la forma general de la envolvente es importante para preservar la característica del sonido.

Las envolventes pueden ser simplificadas, pero la forma general debe ser mantenida. Una estimación de nivel y un control temporal son críticos para el éxito de cualquier algoritmo de compresión, por lo que una simple aproximación puede alterar significativamente la distribución espectral de la señal, ya que las frecuencias con más señal, serán más atenuadas, cambiando sustancialmente sus características. [29]

Lo recomendado en trabajos con una estructura de 28 canales, con filtros de fase lineal, con ceros dentro del círculo unitarios en la respuesta, para derivar la señal de control temporal es introducir rangos de ataque y decaimiento definidos en términos de dB/ms implementados como multiplicadores del valor muestreado, lo que da más eficiencia en términos computacionales que la implementación de filtros. Se calculan dos valores, el RMS y el valor pico de la señal que sirven para ponderar la respuesta entre el valor estimado y el designado, un control de nivel estimado que permita variar entre el valor RMS y el de pico da más flexibilidad al control. Para evitar la distorsión espectral por la distinta energía presente en cada banda, se realiza una normalización ponderada del nivel general, dando como resultado una cantidad de compresión similar a cada banda. [5, 29] La dinámica del audio, mostrada en la Figura 2.12, está reflejada por su relación pico/RMS y un Compresor o Encoder, debe producir una señal de salida que, aunque tenga perceptualmente menor rango dinámico, se presente y escuche muy similar al sonido original, dependiendo de los distintos entornos de escucha. *Schmith y Rutledge* determinaron puntos para las posiciones relativas de los estímulos, relacionando uno de los ejes con el nivel de compresión, lo que ayuda a definir una medida analítica del grado de compresión. [1]

Concluyendo que en lo concerniente a la percepción es muy importante (al aplicar compresión a las señales de audio), generar una métrica apropiada para medir el nivel de compresión, con relación a la curva I/O y a la relación pico/RMS de la señal, lo que ayuda a la determinación de los múltiples parámetros de compresión, mientras mantiene el proceso transparente, para un valor deseado de compresión. [29]

En la figura 2.13 se aprecian los distintos rangos dinámicos en función de las distintas condiciones de audición.

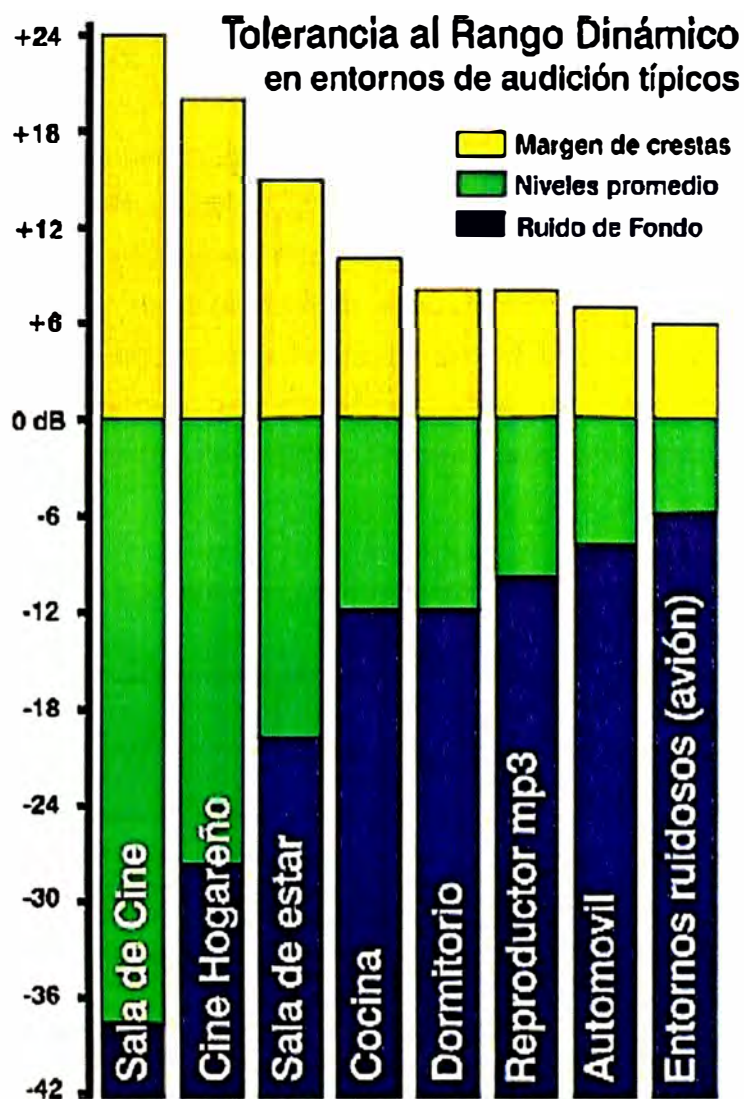


Figura 2.13 Rangos dinámicos típicos

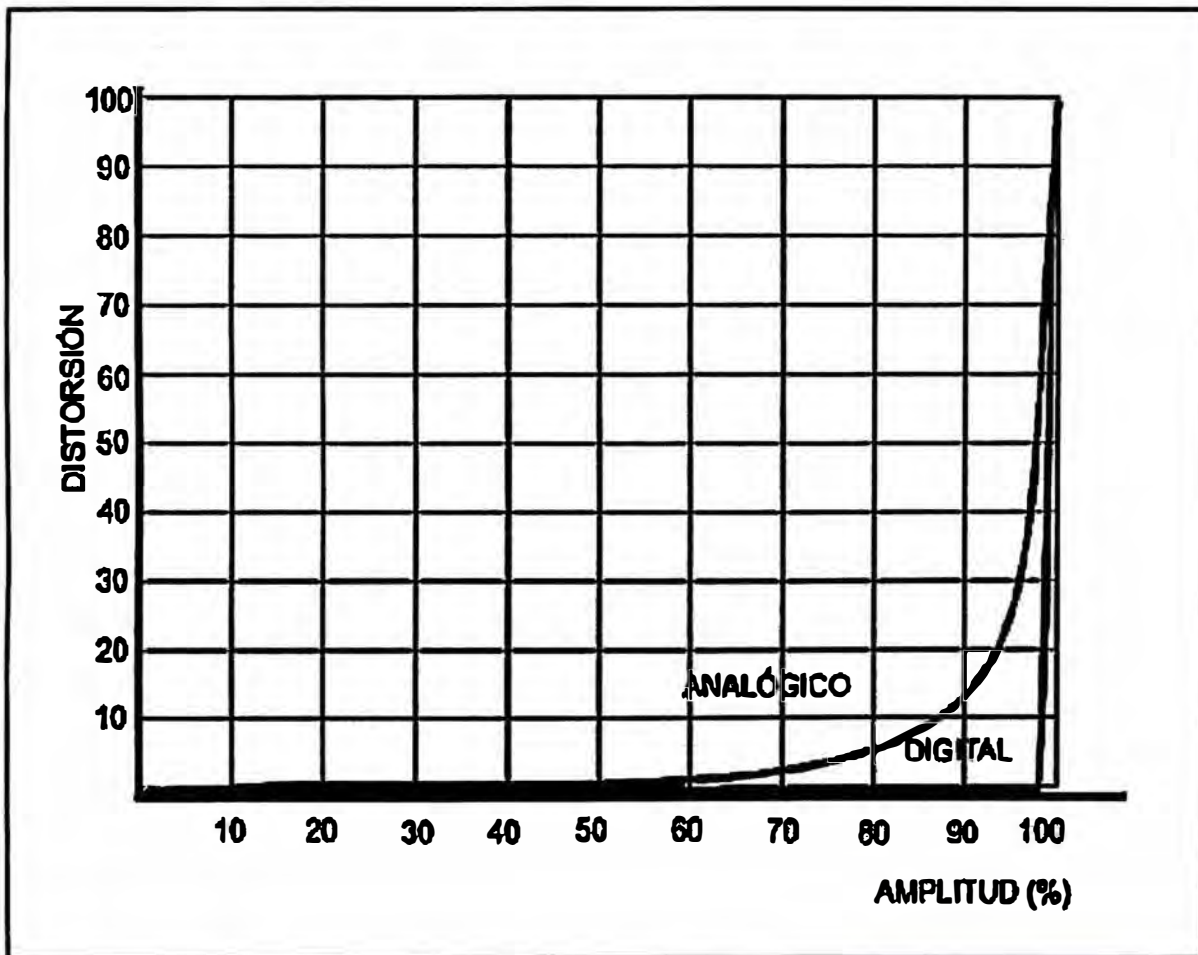
Fuente: <http://www.soundonsound.com/sos/jun08/articles/tclm5.htm>

### 2.5.1 El sobre rango o saturación digital

El concepto de "over" (over-range o sobrerango), se refiere a los valores que quedan por fuera del rango binario codificado y se relaciona al recorte (clipping) de la señal analógica, esto se produce cuando una señal de audio analógica se convierte a digital usando un conversor analógico digital (A/D), hay dos importantes parámetros en esta conversión que son la frecuencia de muestreo y el número de bits utilizados en la muestra digital, la frecuencia de muestreo es importante porque junto al filtro anti-aliasing requerido por el teorema de Nyquist, determina el ancho de banda de la señal muestreada, el número de bits en el muestreo digital determina la resolución que se verá

reflejada en el rango dinámico disponible, como a la relación entre la señal más pequeña representable (ruido de fondo de cuantificación) y el máximo valor binario posible.

En un Disco Compacto, el rango dinámico, en la práctica y dependiendo del equipamiento utilizado está alrededor de 75 dB y la relación señal a ruido (S/N) está en el orden de 96 dB, cuando se produce o graba audio en el mundo analógico, se incluye un margen por sobre el máximo de la señal de audio de manera que el equipamiento pueda manejar esta señal sin la distorsión causada por el recorte al alcanzar su máxima excursión. Hay un tema (relacionado con la psicoacústica), referido a la particularidad de la saturación de las distintas tecnologías, donde cada una tiene su típica forma de distorsión.



**Figura 2.14** Comportamiento aproximado de la distorsión

Fuente: Medición de la Sonoridad del Audio Digital

En un equipo analógico la distorsión a partir de un determinado valor va subiendo paulatinamente, luego del cual aumenta su pendiente exageradamente, comúnmente se

toma el valor donde se comienza a ser notable (1%), como máxima salida útil. [4] En cambio, en una señal digital, la distorsión es despreciable hasta que alcanza el valor máximo, aquí asciende de cero al máximo abruptamente (al alcanzar el máximo valor binario), en la Figura 2.14 se observa una aproximación a estas características.

La sobrecarga distorsionará o recortará (*clipping*) la señal de audio, lo que por supuesto no se desea. Un margen de excursión de la señal es necesario por tres razones importantes:

1. El rango del valor medio al valor pico (factor de cresta) de la señal de audio.
2. La habilidad de un equipo en particular para manejar esta sobrecarga.
3. Margen para una mala calibración de los monitores y medidores de nivel.

Una vez que las características de sobrecarga de un equipo se determinan, sólo se necesita un buen medidor para asegurar que la señal no excederá el rango dinámico del sistema, por lo que muchas soluciones de medición cumplen con la característica de mostrar ambos valores, el valor medio y el valor pico en la misma escala, para que sea fácil de leer, ahora la respuesta de un dispositivo a una sobrecarga depende del diseño y de la tecnología del mismo, algunos lo harán suavemente y otros podrán distorsionar severamente al recortar la señal.

En el audio digital no hay margen más allá de la máxima resolución del audio muestreado, cuando la señal de audio excede la resolución del sistema, se presenta un “*over-range*” y como resultado la señal tiene un tope cuadrado ya que es recortada para valores más allá del máximo valor binario. Cuando una señal binaria utiliza todos los bits disponibles (por ejemplo, 16 en un CD), este nivel de señal se define como de fondo de escala (FULL - SCALE), es cuando la señal viene del mundo analógico y el conversor intenta llevar la conversión más allá del máximo valor binario que resultará esta sobrecarga, además que un valor llegue al fondo de escala no significa necesariamente que la señal ha sobrepasado al sistema; sin embargo, si se presentan una serie de valores de fondo de escala en sucesión, seguramente la señal que ataca al conversor es excesiva ya la conversión está saturada (*clipped*), la señal recortada resultante puede que no sea molesta al oyente, esto dependerá del material de programa y del tipo de música (rock, clásica, etc), por eso, algunos medidores indican los “*over*” digitales en su pantalla cuando detectan 3 valores de plena escala en sucesión. [24]

## 2.5.2 La Guerra de la sonoridad

Conocida mundialmente como "*Loudness War*" se gestó durante los años noventa cuando las compañías discográficas comenzaron a aumentar los niveles de señal en los medios de distribución digitales, básicamente en los CD, ya que estos tienen un rango dinámico mayor a 90 dB, se esperaba que se utilizara el máximo de su capacidad de reproducción, pero el lema comercial hizo que la música de esa década sonara más fuerte; además se encontró que los equipos hogareños no suelen reproducir esta gran rango dinámico. Los grandes bafles de los ochenta se reemplazaron por pequeños bafles exigiendo mayores potencia en los noventa, [36] la aparición de reproductores portátiles utilizando auriculares de mediana calidad, también dio por tierra con el gran rango dinámico.

Internet dio la posibilidad de distribuir archivos de música por la red, por lo que la compresión de datos ayudó que estos fueran reducidos a su mínima expresión, la aparición de un formato de compresión de audio conocido como MP3, permitió reducir el tamaño del archivo hasta dejarlo en un 10% del tamaño original, y por lo tanto reducir el tiempo de transferencia, cuando la mayoría de las conexiones a Internet eran dial-up, esto tuvo su costo en una gran baja de la calidad, pero al público no le importó [4]

En definitiva un archivo comprimido en MP3 a 128 Kbps suena casi como un casete, los reproductores de música digital son este formato, tiene la gran ventaja de que ya no poseen partes móviles, la música está almacenada digitalmente en una memoria flash y se transfiere desde un conector USB, sin copia degenerativa alguna. Estos equipos evolucionaron hasta ser hoy en día incluidos en los teléfonos y demás dispositivos portátiles, pero la calidad de escucha es muy mala, por un lado el ruido del ambiente y por otro, la potencia acústica lograble a partir de auriculares y las baterías atentan contra la Alta Fidelidad que el medio digital puede ofrecer, así las compañías discográficas siguiendo el gusto del público en tener altos niveles constantes "todo el tiempo" redujeron el rango dinámico de sus mezclas hasta llegar a niveles realmente peligrosos. [10]

En la Figura 2.15, se observa diferentes versiones de la canción "Black or White" de Michael Jackson donde se muestran un incremento del Loudness o sonoridad con el tiempo.



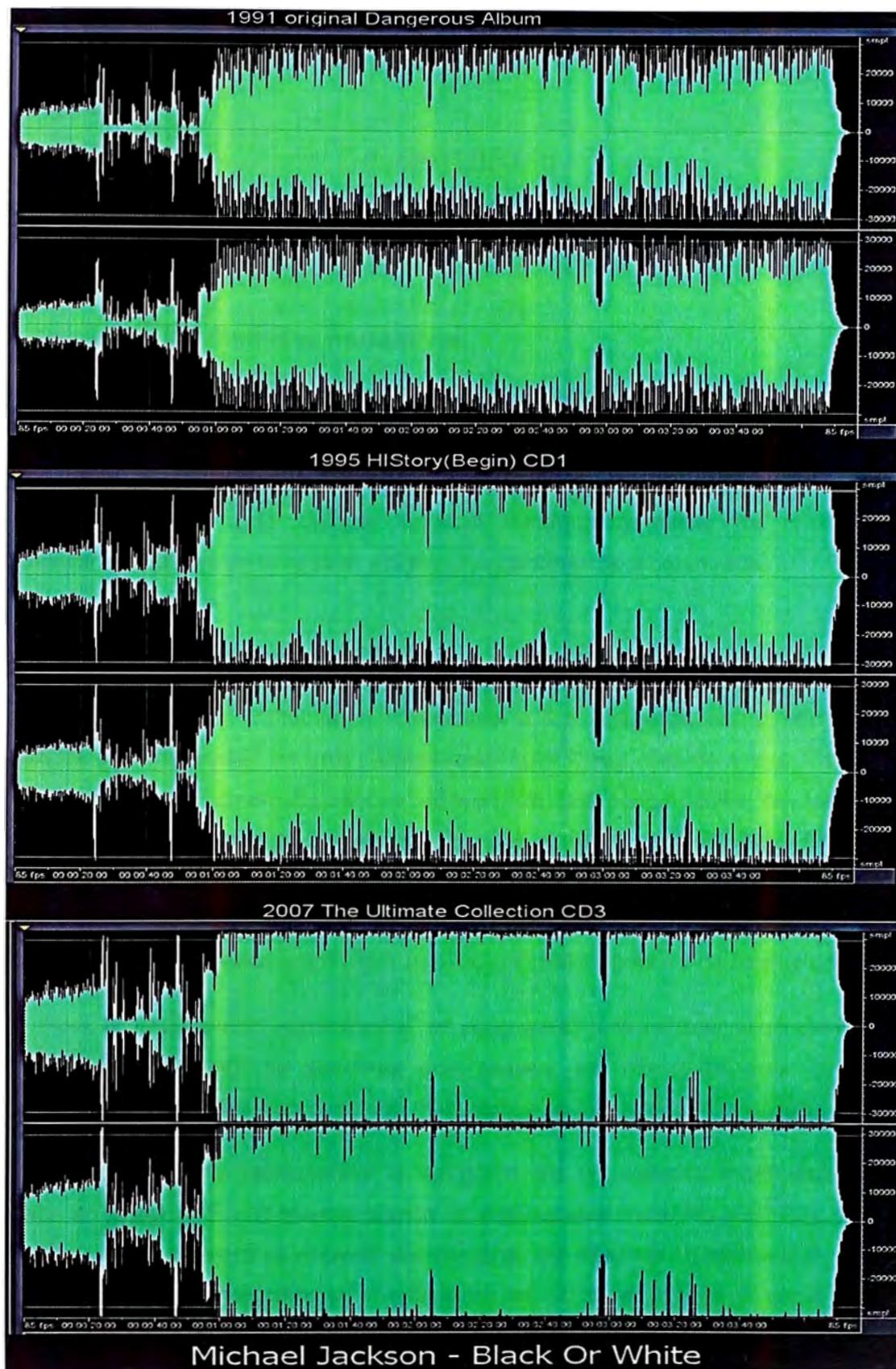


Figura 2.15: Incremento del Loudness con el tiempo

Fuente: [http://en.wikipedia.org/wiki/Loudness\\_war](http://en.wikipedia.org/wiki/Loudness_war)

## **CAPITULO III**

### **METODOLOGÍA PARA LA SOLUCIÓN DEL PROBLEMA**

#### **3.1 Aporte de las distintas normativas**

La solución al problema de la adecuada correlación entre los valores de señal medida y la sensación de sonoridad, se aborda con el estudio de las recomendaciones recientemente presentadas, las cuales intentan presentar una solución al problema planteado, el mismo que permite conocer fundamentalmente los modelos para implementar el sistema de medición y define los parámetros a considerar.

##### **3.1.1 Recomendación UIT BS.775 [37]**

Las importantes limitaciones inherentes a los sistemas de sonido multicanal domésticos y la necesidad de una presentación optimizada, debido a que los requisitos de las proyecciones cinematográficas difieren de las necesidades en los hogares, especialmente en relación al tamaño de la sala y de la pantalla, esto ha llevado a comprender que conviene separar las necesidades de producción, distribución y presentación doméstica, por lo que es conveniente para el oyente implantar un sistema universal de sonido multicanal aplicable a la radiodifusión sonora y de televisión.

Para el intercambio internacional de programas y la mezcla ascendente, es útil establecer una jerarquía de sistemas compatibles, así como disponer de servicios auxiliares destinados a las personas con dificultades en la visión o la audición.

Por lo tanto se recomienda la adopción de un sistema multicanal universal mostrado en la figura 3.1; el mismo consta de dos canales frontales (FR, FL), separados a 60° del arco cuyo centro es el punto de escucha, los altavoces traseros (LS, LR) irán a 120° y la altura debe ser aproximadamente la de los oídos del oyente, el canal central (C) estará en la línea central y el de efectos de baja frecuencia (LFE) en las cercanías de estos; por lo que se necesitan cinco señales, más el canal de bajas frecuencias, denominado comúnmente como 5.1, en la tabla 3.1 se muestran los niveles de matrización, según canales utilizados



TABLA N° 3.1 Ecuaciones de matrización de codificación y decodificación

Fuente: Recomendación ITU-R BS.775

Formato	L	R	C	LS	RS
Estéreo L	1.000	0.000	0.707	0.707	0.000
Estéreo R	0.000	1.000	0.707	0.000	0.707
3/2	1.000	1.000	0.5	0.707	0.707
2/2	$1.0 + 0.7C$	$1.0 + 0.7C$	0.000	0.707	0.707

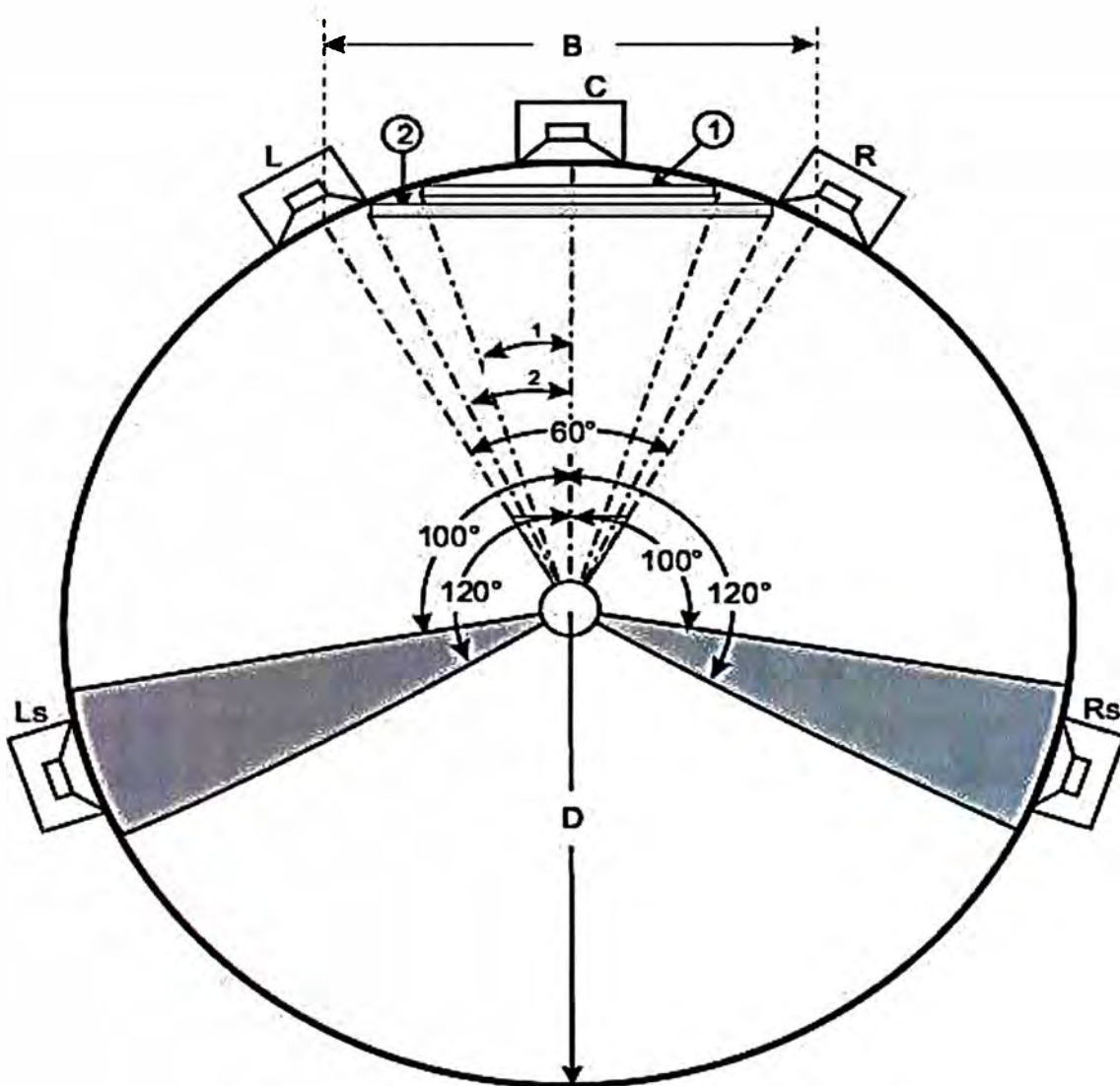


Figura 3.1 Disposición de los altavoces L, R, C y Ls, Rs

Fuente: Recomendación ITU-R BS.775

Donde: (1) es una pantalla SD con distancia de visión  $TVAD = 3H(2\beta_1 = 33^\circ)$

(2) es una pantalla HD con distancia de visión  $TVAD = 2H(2\beta_2 = 48^\circ)$

$H$  = Altura de la pantalla       $B$  = Distancia de los Altoparlantes

En la tabla 3.2 se muestra las posiciones de los parlantes en un sistema multicanal.

**TABLA N° 3.2** Posición de los parlantes multicanal según BS.775

Fuente: Recomendación ITU-R BS.775

Altoparlante	Ángulo Horizontal	Altura (m)	Inclinación (grados)
C	0	1.2	0
L, R	30	1.2	0
Ls, Rs	100 ... 120	$\geq 1.2$	0 ... 15 hacia abajo

### 3.1.2 Recomendación UIT BS.1770 – 2

#### Algoritmo de la sonoridad de los programas de audio y niveles de pico real. [38]

Considerando que los oyentes desean un nivel de sonoridad constante entre programas de distintas fuentes y como el rango dinámico permitido en los sistemas digitales es mucho mayor que en los analógicos, es necesario mantener este valor normalizado, tomando en cuenta el intercambio internacional de programas y como los sistemas digitales saturan abruptamente, deben evitarse los recortes aún de picos momentáneos, como los sistemas de medición de picos de programa no reflejan con exactitud el valor real del pico analógico producido entre muestras y con sistemas de procesamiento digital se puede implementar un algoritmo que estime con exactitud el valor de pico real de la señal, con una indicación precisa del margen entre valor pico y el nivel de recorte. Se recomienda el uso del algoritmo desarrollado para permitir tanto la equalización de niveles en un canal de difusión, como en el intercambio de material sonoro. Este algoritmo, está calculado para material de programa, y por lo tanto no es recomendable para su ajuste, usar tonos puros.

El siguiente diagrama de bloques (Figura 3.2) se indica el algoritmo de medición:



**Figura 3.2** Algoritmo de sonoridad

Fuente: Recomendación ITU-R BS.1770

El pre-filtro tiene una respuesta de corrección de frecuencias, según la curva del modelo acústico de la cabeza (como esfera rígida) con los siguientes coeficientes.

**TABLA N° 3.3** Coeficientes del pre – filtro

Fuente: Recomendación ITU-R BS.1770-2

		$b_0$	1.53512485958697
$a_1$	-1.69065929318241	$b_1$	-2.69169618940638
$a_2$	0.73248077421585	$b_2$	1.19839281085285

El filtro RLB es un filtro pasa altos según la curva de ponderación sonora B modificada con los siguientes coeficientes:

**TABLA N° 3.4** Coeficientes del filtro RLB

Fuente: Recomendación ITU-R BS.1770-2

		$b_0$	1.0
$a_1$	-1.99004745483398	$b_1$	-2.0
$a_2$	0.99007225036621	$b_2$	1.0

Se evalúa el valor medio cuadrático de la señal aplicando la siguiente fórmula:

$$Z_i = \frac{1}{T} \int_0^T Y_i^2 dt \quad (3.1)$$

Donde:  $Y_i$  es la señal ponderada por el filtro RLB, identificada como  $Leq(RLB)$ .

Luego se suman los canales, dando una fórmula:

$$Sonoridad = -0.691 + 10 \log_{10} \sum_i^N G_i * Z_i \text{ LKFS} \quad (3.2)$$

Los coeficientes a aplicar a cada canal se muestran en la tabla 3.3:

**TABLA N° 3.5** Coeficientes de ganancia de los canales

Fuente: Recomendación ITU-R BS.1770-2

Canal	Izquierdo (G <sub>L</sub> )	Derecho (G <sub>R</sub> )	Central (G <sub>C</sub> )	Posterior Derecho (G <sub>RS</sub> )	Posterior Izquierdo (G <sub>LS</sub> )
Ponderación, $G_i$	1.0 (0 dB)	1.0 (0 dB)	1.0 (0 dB)	1.41 (~1.5 dB)	1.41 (~1.5 dB)

### 3.1.3 Recomendación UIT BS.1771 – 1

#### Requisitos de los medidores de sonoridad y de cresta real [39]

Los vúmetros y los medidores convencionales de nivel de cresta de programa no proporciona una indicación precisa de la sonoridad ni una indicación precisa del nivel de cresta real de una señal digital, los oyentes desean que la sonoridad de los programas sea similar entre distintas fuentes y diferentes tipos de programas; dado que el nivel de cresta real de una señal digital puede ser mayor que el máximo valor de muestra. La Recomendación UIT – R BS.1700 especifica un algoritmo de sonoridad y de niveles de cresta reales que cumple con los requisitos de medición de radiodifusión; por lo que los medidores utilizados para medir la sonoridad de programa y determinar el nivel de cresta real, deben satisfacer los requisitos especificados en la Tabla 3.6.

**TABLA N° 3.6** Características de los medidores de Sonoridad

Fuente: Recomendación UIT BS.1771-1

<b>Unidad de Sonoridad (LU)</b>	La unidad de sonoridad es la unidad del medidor de sonoridad, el valor del programa en unidades de sonoridad representa la pérdida o ganancia (dB) necesaria para llevar al programa a 0 LU, un programa con una lectura de -10 LU necesitará 10 dB de ganancia para dar una lectura de 0 LU.
<b>Presentación visual de Tipo I</b>	Presentación visual electrónica con resolución de uno o más segmentos para 3 unidades de sonoridad.
<b>Presentación visual de Tipo II</b>	Presentación visual electrónica con resolución de uno o dos segmentos por cada unidad de sonoridad.

Como ejemplo se puede decir que una señal de referencia es una onda sinusoidal continua con un nivel de presión sonora de 60 dB SPL y  $-24$  dBFS a 1 KHz. El nivel de sonoridad de referencia corresponderá a 0 LU en un medidor de sonoridad. La relación de 0 LU a 0 dBFS (onda sinusoidal a plena escala) es aún objetivo de estudio y el valor de  $-24$  dBFS es sólo un valor ejemplo. Esta señal está destinada a la calibración eléctrica y no es una señal ideal para la medición acústica debido a los efectos de onda estacionaria, una señal que puede utilizarse para la calibración acústica es un ruido continuo en octavos de banda centrado en 1 KHz, con un valor medio de 60 dB<sub>SPL</sub> y  $-24$  dBFS (a título de ejemplo), correspondiendo a un valor medio de 0 LU en un medidor de sonoridad. La medición de ganancia eléctrica utilizando un medidor de sonoridad o calibración cruzada con un vúmetro o un PPM sólo debe realizarse con la señal del nivel de sonoridad de referencia (onda sinusoidal).

Un medidor de sonoridad de radiodifusión tiene al menos dos modos de funcionamiento: rápido (F) e integración (I), se utilizan para distintos objetivos:

- **El modo rápido** se utiliza en la producción, postproducción y presentación. El nivel del programa en un diálogo típico dará una indicación media de 0 LU.
- **El modo integración** se utiliza para control de calidad en la recepción y emisión de programas y en análisis off – line. La salida es un número que indica, sin ambigüedades, la sonoridad y el correspondiente ajuste de ganancia.



Como nuestra percepción de la sonoridad no depende del número de fuentes sonoras involucradas, es lógico especificar una sola presentación visual del nivel de sonoridad para sistemas de sonido multicanal, en vez de una presentación distinta para cada canal; si se incluye el indicador de nivel de cresta para un medidor de programa multicanal, este debe ser activado por el máximo valor que aparece en cualquier canal individual, ello no interfiere con la práctica habitual de medición del nivel por separado para cada canal, debido a que pueden proporcionarse medidores individuales de nivel/cresta para los distintos canales.

Pueden aparecer dificultades ergonómicas si se presentan dos conjuntos de información (sonoridad relativa y nivel de cresta) en una sola presentación visual del medidor, si ambos conjuntos de información tienen la misma categoría y el mismo detalle, el operador podría no saber cuál es más importante y podrá confundirse; por esta razón, en un medidor destinado fundamentalmente a indicar la sonoridad de un programa, la información relativa al nivel de cresta resaltará menos que la información mostrada por un medidor convencional del nivel de cresta del programa, el tiempo de retención mínimo es de 150 ms, tiempo suficiente para que el ojo registre la luz, un tiempo inferior no produciría una iluminación útil.

Al ajustar los niveles de las señales analógicas convertidas en señales digitales, el objetivo fundamental es grabar en un nivel lo suficientemente alto como para evitar el ruido de cuantificación sin que haya riesgo de sobrecarga, para esta aplicación, sería conveniente utilizar un medidor que indique el nivel de cresta verdadero, en lugar de la sonoridad. El objetivo del medidor de sonoridad en radiodifusión es predecir la sonoridad subjetiva en condiciones de reproducción controlada con un nivel de referencia mínimo de 60 dBA<sub>SPL</sub>, el modelo perceptivo para la sonoridad es una función no lineal de la amplitud, la frecuencia y el ancho de banda; en general, la modificación del nivel de audio en  $x$  dB no modifica la percepción de la sonoridad en esa misma cantidad, debido a la respuesta no lineal del sistema auditivo humano.

Por razones prácticas, se utilizan unidades en dB, lo cual es lógico ya que es una tradición en la mediciones de las señales de audio, sin embargo el dB no es una unidad perceptiva y no debe utilizarse para medir la sonoridad; no obstante es conveniente escoger una unidad vinculada al dB, de manera que le medidor pueda indicar cuantos decibeles (en ganancia o pérdida) deben aplicarse para ajustar el programa a la sonoridad de referencia. Como unidad de medición se ha propuesto la *unidad de sonoridad*, esta se define como la unidad que representa la ganancia o pérdida en

decibeles que debería aplicarse a una señal para llevarla a la sonoridad de referencia, las unidades de sonoridad tiene la ventaja de que se distinguen claramente de los dB, de forma que el medidor no se confundirá fácilmente con un PPM o con un Vúmetro.

### 3.1.4 Documento A/85 del ATSC

#### **Técnicas para Establecer y Mantener la Sonoridad del Audio en la Televisión Digital. [40]**

Este documento provee una guía a los radiodifusores y creadores de contenidos de audio para Televisión Digital, recomendando prácticas en la producción, distribución y transmisión, para proveer mejor calidad en la bandas de sonido a la audiencia de Televisión Digital; el documento se enfoca en la medición del audio, en las técnicas de monitoreo, en producción y postproducción y en los métodos que permitan un control efectivo de la sonoridad en la distribución en intercambio de contenidos, adicionalmente recomienda métodos para in control de sonoridad durante la emisión de material intersticial; discute el uso de metadatos y describe un moderno control del rango dinámico, incluyendo información específica del manejo de la sonoridad entre programas.

Principalmente se utiliza un metadato llamado "*Dialnorm*" (definido en la norma A/53), que indica el nivel del elemento ancla del contenido, este refiere al nivel del diálogo o en su defecto al nivel medio del programa, que debe configurarse correctamente, para proveer variaciones de sonoridad potencialmente severas entre transiciones de contenidos y entre canales, sin un adecuado control se pueden esperar saltos de hasta 30 dB. Este parámetro es numéricamente igual al valor absoluto del nivel de diálogo, consta de 5 bits sin signo, indicando cuan debajo está de 0 LKFS, los valores válidos van desde 1 a 31, siendo "0" un valor reservado.

En el codificador AC3, este parámetro se puede definir de tres maneras:

- **Fijo**, se tiene definido un valor fijo y los contenidos estarán conformes a este valor.
- **Predefinido**, se tienen varios valores prefijados, seleccionados por medio de la *Interface de Propósitos Generales* (GPI), según sea necesario para corresponder con el contenido
- **Ágil**, esta configuración permite recibir el metadato del exterior y el flujo de datos (*upstream*) se usa para indicar los cambios dinámicamente, con los valores correspondientes a los cambios de sonoridad tomados desde los contenidos.

Con el objetivo de hacer los programas intercambiables, se fijan una serie de procedimientos de calibración de las salas de control de operación, definiéndolas en cinco categorías, para lo que se proveen tonos senoidales y ruido estocástico (aleatorio) de ajuste, la presión recomendada de escucha está entre 76 y 78 dB*spl*, según el tamaño del recinto de control, para una señal de  $-24$  LKFS. Define además que el valor de 0 dBFS es para una onda senoidal, según lo establecido en AES17-1998/R2004.

### 3.1.5 Documento CST - RT – 017 / 019 [41, 42]

La CST (Commission Supérieur Technique de l'image et du son), define las especificaciones de calidad para la distribución de contenido de audio y video digital en Francia, en lo correspondiente al audio, está basado en las normas IEC 60268 – 5 con distribución de altoparlantes según la norma ITU BR.775.

La presión sonora recomendada para la sala de escucha es de 79 dB*spl* para una señal de 18 dBFS y para el canal de bajas frecuencias LFE se estiman dB*spl*, el nivel de referencia en la sala de control es de 74 dB*spl* para un ruido rosa de  $-18$  dBFS, excepto el canal LFE de 78 dB*spl*. Esto medido con un medidor de picos según la norma DIN 45406 con 10 ms.

La relación entre el medidor PPM y uno en dBFS es de 0 dB en PPM corresponden a +12 dBu y a  $-9$  dBFS, así como  $-9$  dB en PPM corresponden a +4 dBu y a  $-18$  dBFS. La alineación se realizará a  $-18$  dBFS con un tono de 1 KHz, dando  $-9$  dB en PPM DIN 45406 y 0 VU en un vúmetro. El máximo valor pico no debe exceder 0 dB PPM, DIN 45406 y  $-3$  dBFS.

La dinámica según la norma RT-19 debe estar centrada en el nivel de diálogo, se definen las mediciones en unidades LKFS, que es el valor de sonoridad (el nivel) expresado como una medición ponderada de K ( $Leq(R2LB)$ ) a plena escala, la escala LKFS esta graduada en pasos de 1 dB, el nivel de diálogo puede tener un pico de  $-31$  LKFS permitiendo una mayor dinámica y un techo de  $-18$  LKFS con una dinámica más reducida; el valor elegido para el nivel de diálogo normal es de  $-25$  LKFS  $\pm 2$  LKFS, para definir el perfil de la dinámica, se recomienda usar el procesamiento Dolby para "Film Light" ya que es el mejor para radiodifusión televisiva.

Las excursiones permitidas son:



+/- 7 dB para diálogo, valor instantáneo.

+/- 12 dB para todo sonido, valor instantáneo.

Para medir la dinámica, se debe utilizar una herramienta que permita ver el nivel de energía durante 10 segundos en una ventana deslizante en LKFS (ITU BS.1770), para cumplir la sección A2.2.3 del RT17 de CST.

El formato para distribución del audio digital es de 48 KHz de frecuencia de muestreo y 16 bits mínimo para cuantificación PCM (20 bits para el Dolby E) y libre de errores de CRC (Código de Redundancia Cíclica).

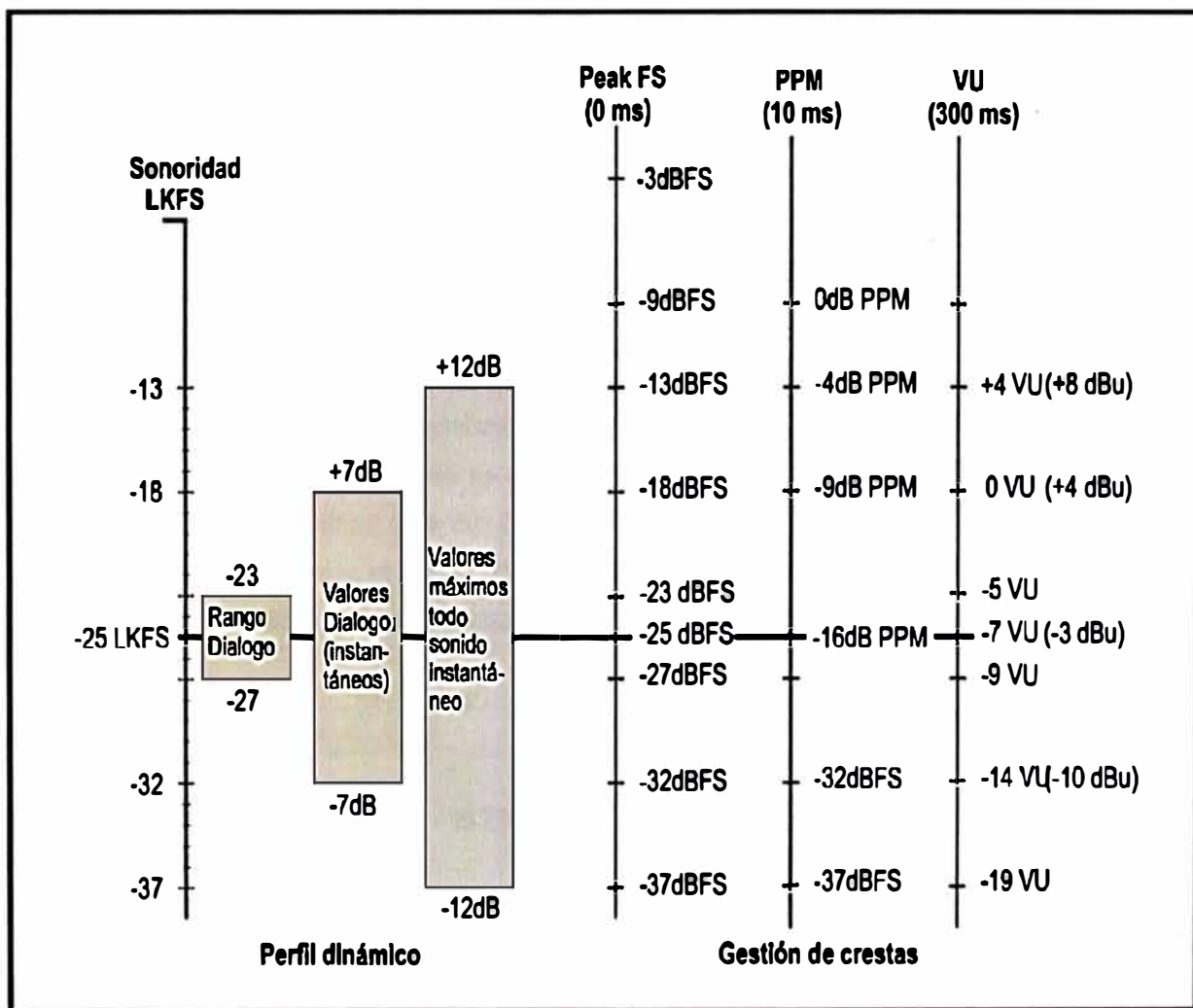


Figura 3.3 Escalas de los distintos niveles según la norma RT – 17

Fuente: Documento CST - RT – 017

Con respecto a la calidad se expresa la importancia de una banda sonora libre de zumbidos, clics y ruidos, con un balance natural y una respuesta en frecuencia plana; los

convertidores 24/25 fps requieren una armonización para mantener el sonido original, en 5.1 la armonización está al límite de la tecnología y en algunos casos puede ser terriblemente destructiva, por lo que es una operación demandante que exige el máximo de esmero y la autorización técnica correspondiente.

Para la transmisión de material monofónico se utilizarán los dos canales con la misma información y fase, cuando se transmita en estéreo, ambos canales tendrán fase positiva; para la reducción de multicanal a estéreo y mantener la coherencia espacial se debe usar el matrizado según la norma EBU R91. Se debe usar la metadata para identificar el programa y la cantidad de canales asignados.

Para la reducción de multicanal a estéreo se recomienda mantener lo siguiente:

- La coherencia espacial
- El nivel e inteligibilidad de las voces
- El balance entre los elementos mezclados
- El balance de frecuencias

La señal obtenida debe satisfacer las exigencias que una señal estéreo tradicional, hay dos tipos de categorías de metadatos para programas multicanal, estos son los musicales (que pueden traer problemas en la mezcla a estéreo) y los no musicales. El nivel de diálogo (*Dialnorm*) se define usando *Leq* (RLB) según la ITU BS.1770, para ello se deben medir todos los canales y el valor obtenido se ingresa en los metadatos del sistema de compresión de audio Dolby AC3, los programas que midan menos de  $-31$  dB deben rechazarse

### 3.1.6 Documento ARIB TR - B29

El estándar de Televisión Digital Terrestre adoptado por nuestro país es el ISDB-Tb (Integrated System Digital Broadcasting – Terrestrial, Brazilian version) desarrollado en Japón y elegido por Brasil en su versión mejorada, la que incluye la compresión MPEG – 4 parte 10 (H.264) en video y la HE – AAC v.2 en audio. Este documento ARIB, define los metadatos en la codificación y los parámetros del audio en las transmisiones digitales terrestres.

En el apartado 2.2.4 *Audio Parameters* se detalla la definición sintáctica de los parámetros:

TABLA N° 3.7 Parámetros de Audio

Fuente: Documento ARIB TR-B29

Sintaxis	Número de bits	Mnemónico
Audio parameter () {		
audio_input_error	1	bslbf
audio_processing	3	bslbf
audio_aes_channels_minus1	2	uimsbf
reserved	2	
For (i=0;i<4:i++){		
audio_ii	10	uimsbf
audio_oi	10	uimsbf
audio_rms_1	10	uimsbf
audio_rms_2	10	uimsbf
}		
}		

Donde: **bslbf** (bit string left bit first) = flujo de bits, bit izquierdo primero

**Uimsbf** (unsigned integer, most significant bit first) = entero sin signo, bits más significativo primero

**Audio\_input\_error** indica si se ha detectado un error en la capa física de la interface de audio (p.ej. errores detectados por el CRCC de la interface serial). Cuando al diagnóstico no está disponible en el punto de monitoreo se define como 0.

**Audio\_processing** indica si se realiza algún procesamiento del audio en el punto de monitoreo, cuando esta información no está disponible se define como 0.

**Audio\_aes\_channel\_minus1** da el número de canales AES menos uno, hay un máximo de cuatro flujos AES y cada flujo contiene dos canales (I y D), dando un total de 8 canales.

**Audio\_ii** indica la información del audio en fase entre dos canales del flujo AES calculado según:

$$A_{ii} = INT \left[ \frac{1}{8} \left( \frac{1}{2N} \sum_{i=0}^{N-1} abs(X(i) + Y(i)) \right) \right] \quad (3.3)$$

Donde:

$X(i)$  y  $Y(i)$  es el valor de la muestra  $i$  en los canales  $X$  e  $Y$ , y  $N$  es el número de muestras de audio durante la duración de un campo de video, la función  $abs(x)$  retorna el valor absoluto de  $x$ ; los canales  $X$  e  $Y$  corresponden al par de canales en un flujo de audio digital AES, el factor de escala  $1/8$  se ha adoptado para indicar el valor del atributo. En este cálculo, se utilizan 16 bits más significativos de la señal de audio, se representan con una notación de 10 bits enteros sin signo, si el valor está por encima de  $2^{10} - 1$  se recorta a  $2^{10} - 1$ .

**Audio\_oi** indica la información del audio fuera de fase entre dos canales AES, según lo calculado por:

$$A_{oi} = INT \left[ \frac{1}{8} \left( \frac{1}{2N} \sum_{i=0}^{N-1} abs(X(i) - Y(i)) \right) \right] \quad (3.4)$$

Descripción de las variables igual que en el caso anterior.

**Audio\_rms\_1** indica la información de la magnitud del audio (AMI) del canal de audio 1 en el flujo de datos AES, calculado por la fórmula indicada en el siguiente ítem.

**Audio\_rms\_2** indica la información de la magnitud del audio (AMI) del canal de audio 2 en el flujo de datos AES, calculado por la fórmula indicada en el siguiente ítem.

$$AMI = INT \left[ \frac{1}{8} \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} X(i)^2} \right] \quad (3.5)$$

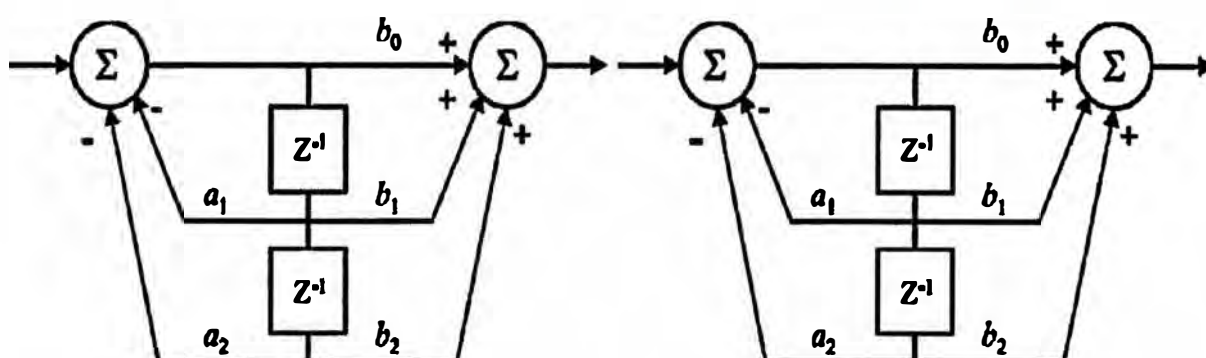
Descripción de las variables igual que en el caso anterior.

Se aplica un pre-procesado a la señal de audio con un filtro de 20 Hz, mostrado en las figuras 3.4 y 3.5, estructurado como dos filtros de 2° orden en cascada, con los siguientes coeficientes:

**TABLA N° 3.8** Coeficientes del pre-filtro de entrada

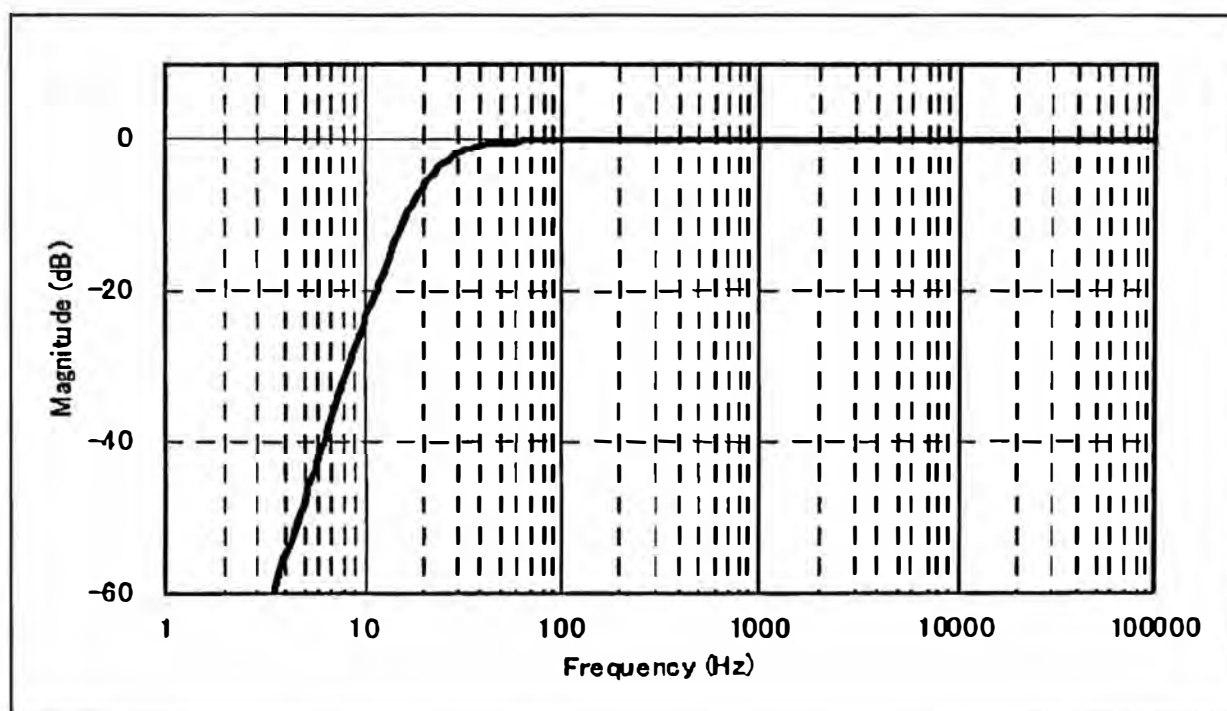
Fuente: Documento ARIB TR-B29

		$b_0$	0.9981318
$a_1$	-1.9962602	$b_1$	-1.9962636
$a_2$	0.996267	$b_2$	0.9981318



**Figura 3.4:** Diagrama de flujo de la señal como un filtro de 4° orden

Fuente: Documento ARIB TR-B29



**Figura 3.5:** Respuesta en frecuencia del pre-filtro

Fuente: Documento ARIB TR-B29

**TABLA N° 3.9** Indicador de las asignaciones de calidad del audio

Fuente: Documento ARIB TR-B29

Indicador de calidad	Nombre de Modo	Criterio de calidad	Detalles
00	Reservado		
01	Modo 1	Modo estándar de televisión terrestre	Bit-rate 192 a 256 Kbps / estéreo o más (2)
10	Modo 2	Distinto a Modo 1, no calificado como Modo 3	Otro modo no especificado
11	Modo 3	Audio con calidad limitada < 192 Kbps	Los Teledifusores elegirán este modo a su criterio

La tabla 3.9 muestra que el Modo 1 es el mejor y es el estándar, ya que cuanto mayor es el bit-rate del flujo, mayor es el ancho de banda y por ende la calidad del audio, con una tasa de más de 192 Kbps el audio es indistinguible del sonido original.

El documento brasilero ABNT NBR 15602-2:2007 especifica formatos, interfaces, niveles de la señal y metadatos para el formato de entrada de audio, mientras el documento japonés ARIB STD-B32:V2.1:2007, parte 2, no especifica estos ítems.

### 3.1.7 Documento EBU-UER R-128 [44]

La EBU (*European Broadcasting Union*) es de la opinión que se necesita un nuevo paradigma en la nivelación del audio basado en su sonoridad, propone que se usen, además del descriptor del nivel de la sonoridad de un programa (*'Programme Loudness'* o *'Sonoridad de Programa'*), los descriptores *'Loudness Range'* (*'Rango de Sonoridad'*) y el *'Maximum True Peak Level'* (*'Nivel Máximo de Cresta Real'*) para la normalización de las señales de audio, cumpliendo los límites técnicos en la totalidad de la cadena de señal, así como las necesidades estéticas de cada programa y/o estación, dependiendo de su género(s) y de la audiencia destino.

Considera que la normalización por los valores pico en la señal de audio, ha llevado a diferencias considerables en la sonoridad entre programas y entre canales emisores de radiodifusión, ya que las inconsistencias de sonoridad resultantes entre

programas y entre canales son la causa de la mayoría de las quejas de los televidentes/oyentes, cuando se usa el Medidor de Cuasi-Crestas de Programa o QPPM (*Quasi-Peak Programme Meter*) especificado en la EBU TechDoc 3205-E, este no refleja la sonoridad de una señal de audio, ya que no está diseñado para indicar el valor promedio a largo plazo. Con la proliferación de los sistemas de producción, distribución y transmisión digital, el nivel máximo permitido de la señal de audio, especificado en la ITU-R BS.645, ya no es apropiado, por lo que se ha definido un nuevo estándar internacional para la medición de la sonoridad de los programas de audio, el ITU-R BS.1770; se introducen además, las unidades LU (*Loudness Unit*) y LUFS (*Loudness Unit*, referenciados al fondo de escala), 'LUFS' es equivalente a 'LKFS' (usada en la ITU-R BS.1770-1).

La medición de la sonoridad de un programa de audio con una compuerta de nivel – gate – podrá mejorar la concordancia de la sonoridad entre programas con un rango sonoro amplio, y el descriptor '*Loudness Range*' o Rango de Sonoridad, puede ser utilizado para evaluar la necesidad de una reducción del mismo, a fin de adecuar los programas a la ventana de tolerancia (gálibo acústico) de la audiencia destino. La UER recomienda que los descriptores '*Programme Loudness*', '*Loudness Range*' y '*Maximum True Peak Level*' sea normalizado a un '*Target Level*' (nivel índice) de  $-23$  LUFS, con una desviación máxima permitida de  $\pm 1$  LU. La medición se hará con un medidor de sonoridad en conformidad tanto al documento ITU-R BS.1770, como al EBU Tech 3341 [4], incluyendo un método de compuerta de nivel – gating – con un umbral relativo a 8 LU por debajo del nivel de sonoridad integrado LUFS especificado en la EBU Tech 3341, el '*Loudness Range*' se medirá con un medidor conforme al documento EBU TechDoc 3342 [5], el '*Maximum Permitted True Peak Level*', durante la producción de un programa, será de  $-1$  dBTP (dB True Peak), medido con un medidor conforme al documento ITU-R BS.1770 y al EBU TechDoc 3341.

Los metadatos de sonoridad serán ajustados para que indiquen  $-23$  LUFS en los programas que hayan sido normalizados a ese nivel y deben indicar siempre y correctamente la sonoridad real del programa, incluso si por alguna razón un programa no puede ser normalizado a  $-23$  LUFS.

Los procesamientos, los sistemas de audio y las operaciones respecto a la producción e implementación, serán realizados de acuerdo al documento EBU TechDoc 3343 [6].



Los procesamientos, los sistemas de audio y las operaciones respecto a la distribución se realizarán de acuerdo al documento EBU TechDoc 3344 [7].

Este documento define que:

Un **Programa** es un elemento audiovisual o sólo de audio autocontenido e individual, para ser presentado en radio, televisión u otro medio electrónico. Un anuncio comercial, un tráiler, un elemento promocional, un intersticial u otros elementos similares deben ser considerados como un programa en este contexto.

La **Sonoridad de programa** (*Programme Loudness*), es la sonoridad integrada en toda la duración de un programa. El nivel de Sonoridad de Programa (*Programme Loudness Level*) es el valor, en LUFS, de la sonoridad.

El **Rango Sonoro** (*Loudness Range, LRA*), describe la distribución de la sonoridad dentro de un programa.

El **Máximo Nivel de Cresta Real** (*Maximum True Peak Level*), es el valor máximo (de cresta) de la forma de onda de la señal de audio de un programa, en el dominio continuo del tiempo.

En el momento de la publicación de esta recomendación (Sep. 2010), los instrumentos de medición, conformes a la ITU-R BS.1770 [3] y a la EBU TechDoc 3341 [4], recién habían comenzado a estar disponibles.

## **3.2 El Loudness en el Audio Digital**

### **3.2.1 Norma Brasileña ABNT NBR 15602-2**

Esta norma define el sistema de audio HE-AAC del estándar de televisión Digital Terrestre ISDB-T, utilizado en nuestro país, analizando el ítem audio, en este documento se encuentra la referencia específica a los parámetros de nivel de las señales de audio y el sistema de codificación y decodificación del sonido.

En las referencias normativas menciona varios documentos indispensables para la aplicación de la misma, incluyendo la recomendación UIT BS 775-1, esta como se ha visto, se refiere a la estandarización de la distribución para señales multicanal (ya sea 2/0, 2/2 y 3/2 o 5.1) y su necesaria homogenización para poder implementar una compatibilidad descendente (remezclado a menor cantidad de canales); esto es



importante porque define el formato de transmisión para que los oyentes tengan la misma distribución espacial con la que fue creado el material de programa, siendo mundialmente aceptado y por lo tanto esencial en la generación de contenidos.

En el apartado 5.2.3 Niveles de señal de audio, de esta norma ABNT, se definen las pautas del volumen sonoro.

Este contiene 4 párrafos:

- 1) *El nivel de referencia para la intensidad o presión sonora debe ser obligatoriamente igual a 0 dB*

La intensidad sonora (WA) y el nivel de presión sonora (SPL) son unidades distintas, en cuanto a que la primera se refiere a la potencia sonora de la fuente y la segunda a la fuerza recibida sobre un área (por ejemplo la membrana del oído). La intensidad sonora está relacionada a la presión sonora por una proporción cuadrática ( $I = P^2/R$ ) el nivel de referencia cuando nos referimos a ondas sonoras es el umbral de audición (20  $\mu\text{Pa}$ ) y referido a intensidad sonora serían unos  $10^{-16} \text{ W/cm}^2$ .

- 2) *La banda dinámica admisible de excursión debe estar limitada obligatoriamente a + 20 dB (headroom) y – 70 dB con respecto a la referencia, correspondiendo a una banda dinámica típica de 90 dB.*

Este párrafo se refiere a que el valor de referencia eléctrico debe tener una capacidad de picos de 20 dB, no se aclara pero se presupone que se refiere a 20 dB del valor pico de la tensión, ya que se menciona el término *headroom*, con respecto a los -70 dB es el valor de ruido del sistema y no se aclara si es el valor pico o medio, por lo que los 90 dB de rango no están perfectamente determinados.

- 3) *Es conveniente que los niveles de audio medio se encuentren a –20 dBFS (0 dB), para permitir homogeneidad en el volumen entre canales distintos.*

Aquí se explica lo mencionado en el ítem anterior, entendiendo que los –20 dBFS corresponden al valor pico de señal, sin indicar si es senoidal, pero como no lo obliga, sólo recomienda, no garantiza una homogeneidad.

- 4) *La señal debe acomodar picos de por lo menos 4 veces su potencia media RMS.*

Se determina que la señal de audio debe tener como mínimo un factor de cresta de 6dB en potencia, por lo que es el doble en tensión (+6 dB). Este criterio, se

entiende, limitaría la posibilidad de sobre comprimir el material de programa, lo que es difícil de cumplir.

Se puede observar que las definiciones para determinar el valor de señal con que acometer al codificador son ambiguas, no se toman en cuenta las distintas ponderaciones para material de programa de Definición Estándar (SD) o de Alta Definición (HD); no se garantiza la universalidad sobre los criterios de sonoridad especificado por la recomendación UIT BS.1770.

En el apartado **5.2.5 ABTN** referido a los metadatos, se determinan los valores obligatorios a cumplir en la transmisión del audio, estos tendrán dos niveles de jerarquía. El primer nivel debe ser obligatoriamente normativo, este nivel afectará la operación del receptor (decodificación de los flujos de bits) incluyendo, por ejemplo, información de cantidad y modo de los canales y el perfil y nivel de codificación, extraídos directamente de las tablas *Program Service Information* (PSI). Como los datos en esta categoría son esenciales para la decodificación y reproducción correcta del servicio de audio en el receptor, son obligatorios. El segundo nivel debe ser obligatoriamente informativo, ese nivel no afectará la decodificación, pero contendrá información sobre los programas de audio asociados a cada *Packet Identification* (PID). Los datos en esta categoría se deben usar obligatoriamente para el procesamiento de la información de los programas en el receptor.

Metadatos utilizados en el stream ISDB-T:

***Matrix\_downmix\_idx:***

Identificador del coeficiente a ser utilizado en el downmix de multicanal a estéreo, se debe transmitir obligatoriamente en el flujo de bits como metadato, conforme a la norma ISO/IEC 14496-3. Es de uso obligatorio al transmitir un programa en modo multicanal.

***Program\_ref\_level:***

Valor representativo de la intensidad media del volumen del programa de audio a largo plazo para todos los canales combinados con relación a la referencia 0 dBFS, es representado en 128 niveles (7 bits), cuantificado en pasos de 0.25 dB, totalizando una excursión de 32 dB con relación al fondo de escala (0 dBFS). Este parámetro debe contener el volumen de referencia adoptado por la emisora con relación al fondo de escala (0 dBFS), para la normalización del diálogo y tomar más confortable el cambio de canales al usuario. Este metadato es de uso obligatorio. Se recomienda utilizar

*prog\_ref\_level* = 80 (0x50), que corresponde a un valor indicativo de -20 dBFS como referencia a 0 dB, el parámetro deberá ser transmitido junto a la estructura DRC, conforme a la norma ISO/IEC 14496-3:2005, sección 4.5.2.7

### ***DynamicRange Control (DRC)***

El control de rango dinámico es especialmente indicado para las transmisiones en modo multicanal y debe señalar los metadatos de conformidad con la norma ISO/IEC 14496-3:2005, sección 4.5.2.7. Su uso en la codificación es opcional, aunque el decodificador debe soportar obligatoriamente esa herramienta, en el caso que la información de DRC no sea enviada el decodificador del receptor no debe aplicar la herramienta DRC.

### **3.2.2 El caso peruano**

En Perú al elegir el estándar ISDB-Tb para la transmisión de la televisión Digital Terrestre, aún no se han definido las normas técnicas que la regirán, ya que el sistema ha sido desarrollado por Japón y por Brasil. En nuestro país se han tomado estas normativas como válidas en las implementaciones realizadas hasta la fecha, pero en éstas no se han incluido aún las pautas de sonoridad del audio digital que se están definiendo, por lo que su análisis y recomendación es un área de trabajo.

Será entonces recomendable definir los parámetros de sonoridad correctos para la producción, contribución, distribución y transmisión para proveer una banda sonora de calidad internacional a la audiencia de televisión digital local, cumpliendo las técnicas contemporáneas al producir contenidos de audio que garanticen la homogénea sonoridad del material.

### **3.3 Criterios de diseño de un medidor de loudness**

Inicialmente se estudiaron de los antecedentes para comprender las necesidades a cumplir en la medición de los niveles de señal, por lo que se decidió desarrollar la aplicación del algoritmo propuesto en la solución dada por la norma BS.1770 utilizando un software específico de edición de audio, se ha establecido que se debe estudiar los tres modos de medición, el instantáneo, el de corto plazo y el de integración.

Para cumplir con los objetivos, el estudio se orientó inicialmente en la comprobación de los modos instantáneos y de corto plazo del medidor de sonoridad, ya que su utilidad sería el reemplazo del vúmetro, para lo que se diseñó un medidor autónomo conceptual. En primer lugar se realizaron las mediciones simulando el algoritmo con aplicativos de

audio y luego se implementó y calibró el prototipo autónomo, conectado a la salida de la PC.

Para la simulación del algoritmo y la medición de su desempeño se han utilizado las siguientes aplicaciones diseñadas para la edición de audio, el *Adobe Audition CS6*, para la medición de los parámetros, se ha utilizado el *SpectraLab v4.32.17* de *Sound Technology Inc.* Y el *TrueRTA Level 4 v3.11* de *True Audio Inc.*

Para las mediciones comparativas de sonoridad se ha utilizado el prototipo del *Loudness Meter V2.08* de *Orban* y el *Loudness Meter vR1.7* de la *BBC R&D*, posteriormente para la realización del algoritmo y la programación por software de los filtros de procesamiento digital se ha utilizado el *MATLAB Student (R2013a)*.

Un criterio importante en el desarrollo del prototipo, fue obtener un dispositivo de bajo costo y realizable con tecnología disponible en nuestro mercado, para lo que se comprobaron diversas configuraciones circuitales, se evaluó, el comportamiento de la medición de sonoridad integrada, para lo que se desarrolló el software correspondiente, encontrando que en archivos de audio con grandes silencios los valores predichos por el algoritmo BS.1770 eran menores que los correspondientes al sonido real, dando en consecuencia un resultado que, aplicado como normalización a la señal de audio, elevaba sus niveles más de lo necesario.

Se comprobaron distintos tipos de compuertas de nivel, inicialmente se probó con compuertas fijas, encontrando que no cumplían con los requerimientos, luego se realizaron las modificaciones al software para trabajar con compuertas dinámicas, donde el valor del umbral está referido al valor de la sonoridad media del archivo; este tipo de compuerta dio un desempeño satisfactorio para la mayoría del material estando por debajo del umbral definido como no aceptable, encontrando que valores de  $-10$  LU dan los resultados más satisfactorios.

## **CAPITULO IV**

### **ANÁLISIS Y PRESENTACIÓN DE RESULTADOS**

#### **4.1 Análisis descriptivo de las variables de estudio**

El diseño pretende mostrar que un indicador de sonoridad autónomo es útil como indicador del “volumen” de un programa, reemplazando a los vúmetros como indicador de nivel de audio, esto se logró haciendo un dispositivo dependiente del modo de integración utilizado. Definidos 3 modos de medición: el momentáneo, el de corto plazo y el de integración. El momentáneo tiene un tiempo de lectura de 400 ms y no utiliza compuerta de nivel, es indicado para la medición de programas en vivo, en tiempo real, sirve para evaluar el nivel de un material de audio, como un medidor de nivel estándar, reemplazando a los vúmetros. El modo de corto plazo tiene un tiempo de lectura de 3 segundos y no utiliza compuerta de nivel, indica el valor promedio instantáneo, dando una idea de cómo varía la sonoridad a lo largo de un programa, archivo o stream de audio. El modo integración lee todo el programa, en este modo puede haber una significativa diferencia, dependiendo de la cantidad de silencios que contenga el material de programa, el valor instantáneo de la sonoridad puede variar mucho y es el que por lo tanto requiere utilizar la compuerta de nivel propuesta en este informe.

Analizando los valores definidos como objetivo de la normalización, se encontró que el valor de referencia o ancla, elegido por la norma BS.1864 es de  $-24$  LKFS, el valor propuesto por la NBR 15602, es de  $-20$  dBFS, el valor propuesto por la EB es de  $-23$  LUFS  $\pm 2$  LU y el propuesto por la CST es de  $-25$  LKFS  $\pm 2$ .

Se define entonces el 0 LU en el valor  $-23$  UFS en reemplazo del LKFS, el valor mínimo de señal indicado en la NBR 15602 es de  $-70$  dBFS, que corresponderán a  $-70$  LUFS y se puede llevar a un equivalente de  $-47$  LU; este valor sería el mínimo indicado o el umbral de lectura, la norma CST define que valores por debajo de  $-31$  LU no son útiles, por su baja amplitud y seguro enmascaramiento por el ruido ambiente.

Se propuso que le dispositivo autónomo pueda indicar el valor de sonoridad y también el valor pico. Para lo que se realizaron dos prototipos analógicos autónomos, el

primero para probar el concepto del visualizador matricado en decibeles, logrando un dispositivo con una escala de 30 dB, constituido por 30 Leds. Incorporando un medidor de valor medio y el medidor de valor pico con una especial constante de descarga (Modelo de utilidad, registrado en el INPI con el N° 20100101198). Ambos valores seleccionables mediante un interruptor.

Este medidor se desempeña como un vmetro, midiendo el valor medio. Esto permiti comprobar el funcionamiento de la matriz logarmica de 30 diodos led y el concepto bsico de disponer de un medidor con un paso de 1 dB por led.

El segundo prototipo integr la incorporacin del filtro Head, el filtro RLB y el conversor de RMS. Con una presentacin simultnea de ambos valores, mediante un sistema de multiplexado, permitiendo seleccionar y visualizar el valor pico, el de sonoridad o ambos simultneamente adems de tener dos constantes de tiempo para la integracin del valor de sonoridad.

Este prototipo es un dispositivo de 30 leds, que concuerda con la especificacin BS.1771, para medidores de tipo II.

Se utiliz adems para comprobar la lectura de ambos parmetros, a fin de apreciar las velocidades de integracin del instrumento, la relacin de los niveles de sonoridad a los picos de la seal y los valores de seal reales medidos a la salida de la PC de trabajo (y no solo los medidos dentro de la PC).

El algoritmo para el valor de sonoridad resulta:

1. Se aplica el filtro de ajuste del modelo de la cabeza humana (FHead).
2. Se aplica el filtro de ponderacin RLB modificado,  $l_{eq}$  (RLB).
3. Se calcula el valor RMS con distintos valores de integracin (RMS).
4. Se convierte la salida en una escala logarmica en decibeles (dB).

El algoritmo para el valor pico resulta:

1. Rectificacin completa de la seal.
2. Filtro pasa-bajo asimtrico. El tiempo de ataque es menor a 10  $\mu$ seg y la descarga tiene una corriente constante de 0.27  $\mu$ A dando una pendiente exponencial.
3. Integra el valor de la envolvente.
4. Convierte la salida en una escala logarmica en decibeles.

Se calibr el prototipo usando la referencia dada por las mediciones de software.

La tabla 4.1 muestra los valores experimentales.

**TABLA N° 4.1** Valores experimentales del prototipo de medidor

**Fuente:** Medición de la Sonoridad del Audio Digital

N° Led	Escala LUFS	Encendido LEDs
30	2	1.97
29	1	1.07
28	0	0.05
27	-1	-0.92
26	-2	-2.09
25	-3	-3.15
24	-4	-3.93
23	-5	-5.09
22	-6	-6.1
21	-7	-6.95
20	-8	-7.95
19	-9	-9.17
18	-10	-10.12
17	-11	-11.15
16	-12	-12.05

N° Led	Escala LUFS	Encendido LEDs
15	-13	-13.01
14	-14	-13.98
13	-15	-15.01
12	-16	-15.95
11	-17	-16.95
10	-18	-17.97
09	-19	-19.2
08	-20	-20.05
07	-21	-21.03
06	-22	-22.17
05	-23	-22.8
04	-24	-23.7
03	-25	-25.21
02	-26	-25.89
01	-27	-26.85

## 4.2 Análisis teórico de los datos

Durante este proyecto se comprobaron los valores recomendados así como las modificaciones al algoritmo propuesto por la BS.1770, inicialmente se desarrolló una serie de mediciones para comprobar los valores de los parámetros.

Como se observó que este algoritmo no logra correlacionarse con todos los rangos de sonoridad presentes en los contenidos de audio cuando hay grandes silencios

presentes, propuso incluir una corrección de lectura en el modo integración, constituida por una compuerta de nivel, con un umbral relativo.

Se encontró que el valor de  $-70$  LU como mínimo valor útil, permite manejar los archivos de audio con un margen sobre el ruido de fondo, dando lecturas consistentes, en las pruebas sólo se tomaron los algoritmos modificados que dieron menor error con respecto a los originales y por lo tanto son los más representativos.

El experimento de medición sobre la modificación del algoritmo se basó en el modo de integración ya que los otros modos se utilizan en tiempo real, verificados con el prototipo, para lo que se utilizó una computadora PC en la evaluación del algoritmo por software. El procedimiento utilizado en la experiencia, consistió en la medición comparativa de archivos de audio, primero se midió el valor RMS del archivo y luego se aplicó el filtrado tomando el valor corregido, esto se realizó con un total de 16 archivos.

Todos los archivos fueron normalizados antes de la prueba para tener un nivel similar durante el experimento, para lo cual se realizó el siguiente procedimiento, se emuló la ecualización BS.1770 en el programa *Adobe Audition*, de la siguiente manera:

1. Se hizo una copia del archivo.
2. A esta copia se la pasó por el filtro FFT del modelo de cabeza humana.
3. Luego se pasó por el filtro RLB modificado, dando  $Leq(RLB)$
4. Se midió su valor RMS por software y se sumaron los canales.
5. Se ajustó el valor obtenido, según la fórmula de sonoridad.
6. Se aplicó esa cantidad al archivo original, quedando normalizado.

Esto es, todos los archivos deberían dar el mismo nivel según el algoritmo.

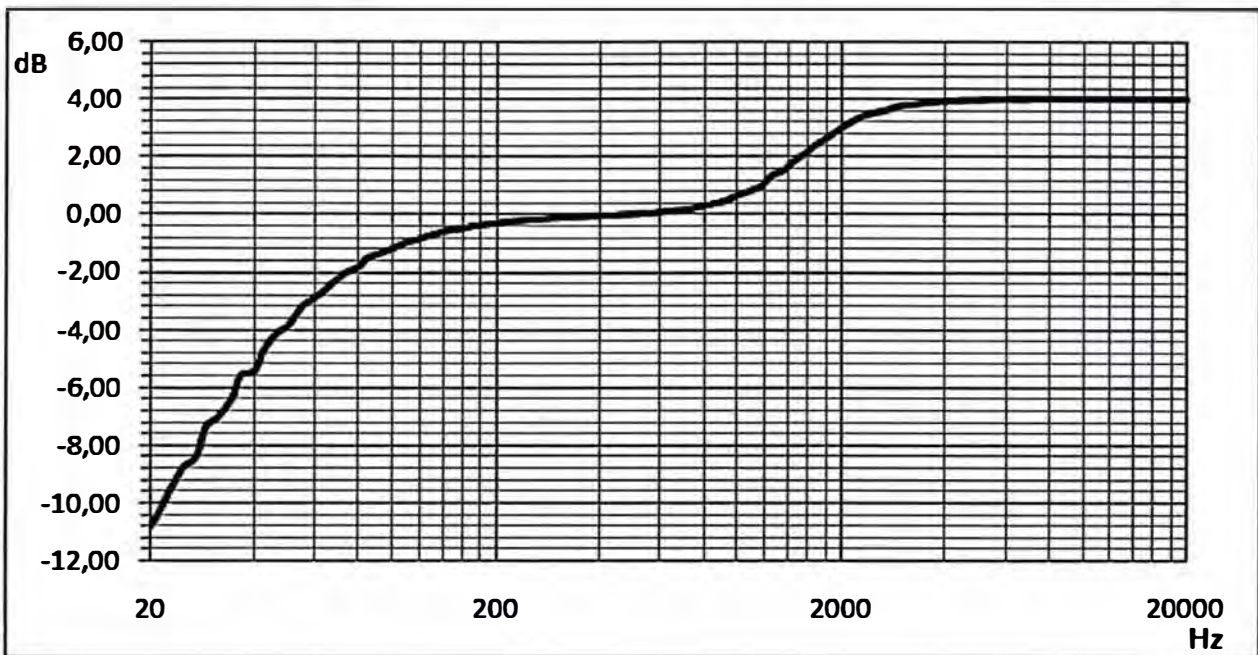
La curva de ecualización implementada emula la respuesta del oído a niveles medios de intensidad sonora [47]

Esta curva está generada a partir del filtro RLB utilizado en la ponderación de frecuencias, cuya fórmula es:

$$h(z) = \frac{1 - 2z^{-1} + z^{-2}}{1 - 1.99004745483398z^{-1} + 0.99007225036621z^{-2}} \quad (4.1)$$

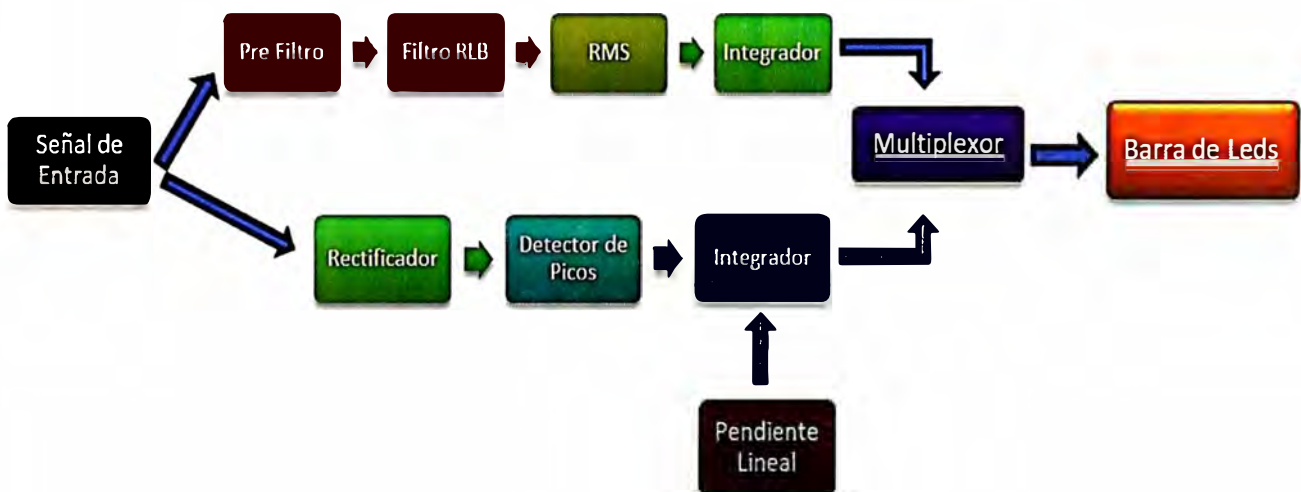
Además se incluyó el filtro de pre-énfasis a 1 KHz de  $+4$  dB, dado por el efecto de la cabeza, la curva de respuesta en frecuencia resultante se observa en la Figura 4.1.





**Figura 4.1** Curva de ecualización aplicada  
**Fuente:** Medición de la Sonoridad del Audio Digital

En la figura 4.2 se muestra el diagrama en bloques del algoritmo utilizado en el prototipo.



**Figura 4.2** Diagrama del algoritmo desarrollado en el prototipo  
**Fuente:** Medición de la Sonoridad del Audio Digital

Luego se midieron con la simulación del algoritmo en el MatLab, tomando sus valores.

La referencia se tomó en función de que cada archivo (luego de pasar por los filtros) se equiparó al valor RMS de un ruido rosa. Se usó un nivel acústico de referencia

ajustado a 70 dBspl, como valor confortable para la experiencia de comparación auditiva, utilizando ruido rosa. Los archivos seleccionados consistieron en tres categorías, locución hablada, música moderna con altos niveles de compresión y bajo rango dinámico y música con altos valores dinámicos. [48] El sonido de las voces fue monofónico y con una duración no mayor a 30 segundos. Los archivos de música se tomaron de CDs y se ajustaron a 48 KHz.

Luego se desarrollaron diferentes algoritmos para acercarse al valor final, con práctica cualitativa, a medida que avanzaba la experiencia se fueron corrigiendo los cálculos. Se probaron los algoritmos correspondientes a la ponderación RLB y a la ITU BS.1770, primero sin compuerta, luego con compuerta fija y por últimos con compuerta dinámica, relacionada al valor de sonoridad medido.

El valor de la compuerta de nivel será el umbral por sobre el que se medirá la sonoridad, descartando los valores por debajo de este, que se calcula según:

$$Z_i = \frac{1}{T_c} \int_{T_{c_j}}^{T_c(j+1)} y_i^2 dt \quad (4.2)$$

Donde  $Z_i$  = Valor de sonoridad del bloque

$T_c$  = Tiempo asignado al bloque

$j$  = Valor del bloque analizado ( $j = \{0,1,2, \dots (T/T_c) - 1\}$ )

$y_i$  = Valor de nivel medido

Para un umbral de compuerta asignado por un valor  $\Gamma$ , se obtendrán los índices que cumplan con la condición de ser inferiores al umbral, la compuerta dinámica se obtuvo de restar el valor de la sonoridad media y el valor del umbral, donde  $j_c = \{j: l_j < \Gamma\}$ .

La medición de sonoridad controlada por compuerta en el intervalo T se define como:

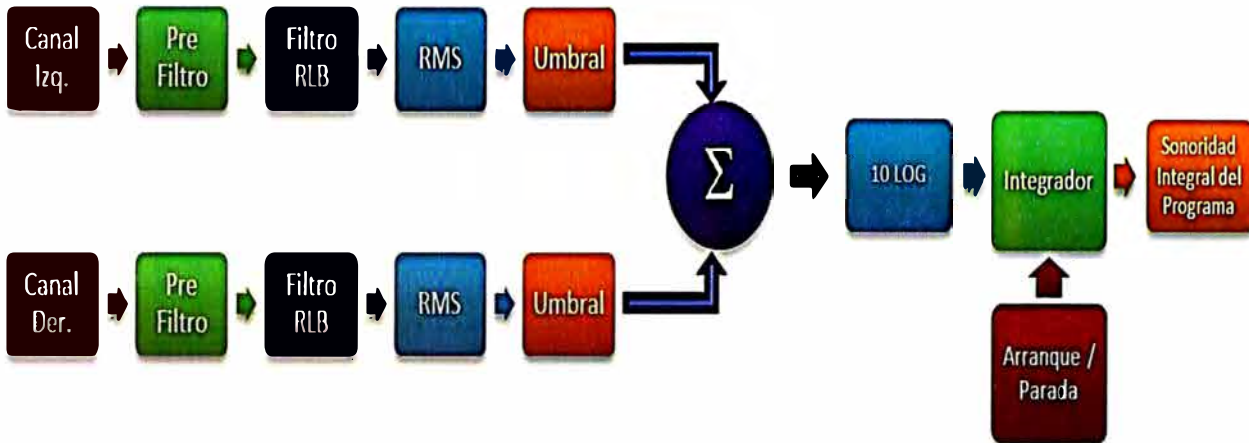
$$\text{Sonoridad controlada} = -0.691 + 10 \log \sum G_i \left( \frac{1}{|j_g|} \cdot \sum_{j_g} Z_{ij} \right) \quad (4.3)$$

En las mediciones, los valores resultantes permitieron comprobar el comportamiento de esta fórmula con diferentes valores de umbral y compararla con los obtenidos en el cálculo de los fragmentos seleccionados.

El umbral de compuerta se eligió desde 6 dB hasta 12 dB por debajo del valor medido, para filtrar los fragmentos por sobre el índice de sonoridad y así calcular el valor

ajustado, se realizó el análisis por medio del *Adobe Audition* seleccionando los fragmentos que estaban por encima del valor calculado y obteniendo la sonoridad de ellos, descartando los demás.

En la figura 4.3 se observa el diagrama del algoritmo desarrollado en el MATLAB.



**Figura 4.3** Diagrama del algoritmo desarrollado en software

Fuente: Medición de la Sonoridad del Audio Digital

### 4.3 Costos de Software

La tabla 4.2 muestra los costos de los softwares utilizados

**TABLA N° 4.2** Costos de Software

Fuente: Elaboración propia

SOFTWARE	Costo
<b>AUDITION CS6 ONLY SUBSCRIPTION</b>	
Month to Month Audition CS6	\$29.95/mo USD
Yearly Audition CS6	\$19.95/mo USD
<b>Fuente:</b> <a href="https://blogs.adobe.com/audition/">https://blogs.adobe.com/audition/</a>	
<b>TRUERTA - LEVEL 4</b>	\$99.95
<b>Fuente:</b> <a href="http://www.trueaudio.com/catalog1.htm">http://www.trueaudio.com/catalog1.htm</a>	

<b>MATLAB STUDENT</b>	
Includes MATLAB only. Add-ons must be purchased separately.	USD \$45.00
Fuente: <a href="http://www.mathworks.com/academia/student_version/">http://www.mathworks.com/academia/student_version/</a>	
<b>LOUDNESS METER</b>	FREE
Fuente: <a href="http://www.orban.com/meter/">http://www.orban.com/meter/</a>	
<b>SPECTRALAB</b>	FREE
Fuente: <a href="http://keygen.la/S/SpectraLAB-v4-32-17.html">http://keygen.la/S/SpectraLAB-v4-32-17.html</a>	

#### 4.4 Resultados obtenidos

El método de evaluación aplicado sobre los datos resultantes de la experiencia consistió en el análisis de la diferencia en LUFS de los archivos comparados (DN). Se aplicó un análisis de regresión para encontrar la relación entre la sonoridad percibida y la predicha por el modelo con las variables de referencia.

El resultado del análisis de regresión es el valor estimado del nivel de cada fragmento musical (NF) basado en la solución del método de mínimos cuadrados.

También se obtuvieron los histogramas de los archivos para su análisis comparativo.

La tabla 4.3 muestra los niveles de corrección de los 16 archivos utilizados en la experiencia.

**Tabla 4.3** Resultados de las mediciones

Fuente: Medición de la Sonoridad del Audio Digital

Archivo	Nombre	Sonoridad medida		
		Audition	Algoritmo	Audición
1	Fantasia.wav	-23.47	-21.6	+
2	Bataille1.wav	-28.88	-28.7	+



3	Música Country mujer.wav	-26.23	-26	+
4	Multitud_1.wav	-26.14	-26.5	-
5	Documental.wav	-25.88	-26	-
6	Drama_1.wav	-24.27	-25	-
7	Drama3_mujer.wav	-25.23	-25.5	-
8	Entretenimiento.wav	-26.85	-26.8	+
9	Interviú_mujer.wav	-28.75	-28.7	+
10	Interviú_hombre.wav	-23.20	-23.1	+
11	Interviú2_hombre.wav	-30.03	-29.1	+
12	voces_hombre1.wav	-21.60	-22.2	-
13	Conferencia.wav	22.50	-22.3	+
14	Música_Pop.wav	-23.31	-23.2	+
15	Vocal1_hombre.wav	-20.30	-19.9	+
16	Vocal1_mujer.wav	-22.01	-22.4	-

En la tabla 4.4 se muestran los valores de sonoridad para los distintos umbrales.

**Tabla 4.4:** Valores de sonoridad para distintos umbrales

**Fuente:** Medición de la Sonoridad del Audio Digital

Pista	Sin umbral	Umbral -3LU	Dif	Umbral -8LU	Dif	Umbral -10 LU	Dif	Umbral -12LU	Dif
Fantasia.wav	-21.6	-21.4	0.2	-21.6	0.0	-21.6	0	-21.6	0
Bataille1.wav	-28.7	-28.6	0.1	-28.7	0.0	-28.7	0	-28.7	0
Música Country mujer.wav	-26	-25.8	0.2	-25.9	0.1	-26	0	-26	0
Multitud_1.wav	-26.5	-26.3	0.2	-26.4	0.1	-26.5	0	-26.5	0

Documental.wav	-26	-25.4	0.6	-25.8	0.2	-26	0	-26	0
Drama_1.wav	-25	-23.5	1.5	-24.6	0.4	-25	0	-25	0
Drama3_mujer.wav	-25.5	-24.6	0.9	-25.1	0.4	-25.3	0.2	-25.5	0
Entretenimiento.wav	-26.8	-26.7	0.1	-26.7	0.1	-26.7	0.1	-26.7	0.1
Interviú_mujer.wav	-28.7	-27.9	0.8	-28.6	0.1	-28.7	0	-28.7	0
Interviú_hombre.wav	-23.1	-22.7	0.4	-23.1	0.0	-23.1	0	-23.1	0
Interviú2_hombre.wav	-29.1	-28.8	0.3	-29.0	0.1	-29	0.1	-29	0.1
voces_hombre1.wav	-22.2	-21.6	0.6	-22.0	0.2	-22.1	0.1	-22.1	0.1
Conferencia.wav	-22.3	-21.9	0.4	-22.2	0.1	-22.3	0	-22.3	0
Música_Pop.wav	-23.2	-22.3	0.9	-23.2	0.0	-23.2	0	-23.2	0
Vocal1_hombre.wav	-19.9	-19.5	0.4	-19.6	0.3	-19.7	0.2	-19.8	0.1
Vocal1_mujer.wav	-22.4	-21.3	1.1	-22.1	0.3	-22.3	0.1	-22.4	0
<b>Diferencia Máxima</b>			<b>1.5</b>		<b>0.4</b>		<b>0.2</b>		<b>0.1</b>

## CONCLUSIONES Y RECOMENDACIONES

### Conclusiones

1. La necesidad de ajustar el nivel de las señales de audio de acuerdo a un modelo existente consiste en la normalización de los niveles por su sonoridad y no por sus valores picos. Esto constituye un cambio de paradigma, una nueva visión del entorno de producción y emisión de material de audio. Este concepto es posible gracias a que en la modulación digital no importa el nivel de la señal en función del alcance de las emisoras, hace falta, para una exitosa implementación de este concepto, una revisión de las técnicas de operación y control del audio en la producción y post producción de contenidos. Esto involucra un cambio en los parámetros de calibración y ajuste de los equipos utilizados, que exige una actualización del personal interviniente.
2. Cuando se mencionan valores de amplitud de una señal, es necesario indicar si se trata de valores pico o valores eficaces, porque de no estar indicado claramente el tipo de medición especificada se tiene una incertidumbre y el error puede ser grave, ya que la relación entre el valor eficaz (RMS) y el valor pico, depende de la forma de la señal, es decir, de su factor de cresta. En las mediciones y las especificaciones normativas es necesario definir explícitamente estos valores.
3. No es conveniente, en audio, referirse a valores de sonoridad con ondas senoidales ya que la música tiene mucha más factor de cresta, es más lógico utilizar ruido rosa o blanco para las mediciones críticas. En el caso de usar los dBFS hay que aclarar que se refieren a la máxima señal senoidal, por lo tanto su indicación máxima está 3 dB por debajo del máximo binario. El máximo absoluto o fondo de escala digital, está definido como dBDFS (dB Digital Full Scale) o dBTP (dB True Peak).
4. En la evaluación de la sonoridad del audio digital, la utilización de los LKFS o LUFS, significa que se está ponderando la medición en función de la respuesta del oído y en el tiempo (integración temporal). En cambio, los dBFS no tienen ninguna ecualización en frecuencia y podría decirse que son instantáneos.



5. La utilización de tecnología informática permite fácilmente realizar estos cálculos y es posible construir indicadores que cumplan con estos nuevos parámetros, es recomendable el uso de la magnitud LUFS de la EBU R128, ecualizada para el oído humano con la sumatoria de todas las fuentes de sonido. El grupo de trabajo P/LOUD de la EBU ha recomendado el uso de una unidad referida a  $-23$  LKFS denominada LU (Loudness Units), similar a la de  $-24$  LKFS de la ITU. El incremento de 1 LKFS entre ambas, se debe al hecho de poder aplicar la compuerta de nivel necesaria para armonizar las fuentes de audio, aun conteniendo pasajes silenciosos.
6. El material de audio puede definirse entre material con Gran Rango de Sonoridad (GRS) y Poco Rango de Sonoridad (PRS), que se diferencian por su Rango de Sonoridad (RAS). El material con GRS se encuentra en películas, música clásica, efectos sonoros y deportes, el material con PRS se encuentra en comerciales, promociones y avisos.
7. El nivel máximo de sonoridad momentánea (SM), el valor del rango de sonoridad (RAS) y el valor medio de la sonoridad integrada (SI), son descriptores que permiten conocer detalladamente el contenido y sus características de emisión. Estos pueden ser obtenidos por tratamiento de la información dentro de un dispositivo autónomo o por una aplicación de software específico en una PC.
8. Cuando se utiliza directamente el algoritmo de la recomendación ITU BS.1770, se presenta insuficiente para la evaluación sonora de un tipo de material con gran rango de sonoridad (GRS), tal como películas con considerable proporción de fondos silenciosos. La introducción de una compuerta en la medición puede resolver este problema, la elección de un correcto valor de umbral relativo es crítica para su efectividad, ya que no debe arruinar la compatibilidad con el material de poco rango de sonoridad (PRS)

## **Recomendaciones normativas**

Se necesita determinar en Perú un patrón de referencia para la normalización del audio digital en televisión, como se ha realizado en otros países [49] Esta norma podría ser emitida por el IRAM, e base a las propuestas desarrolladas en este trabajo. Tomando como partida la norma NBR15606-2, por lo que se propone:

- Utilizar la nomenclatura LU (Loudness Units) para referirse al nivel del audio.
- Estandarizar el nivel de referencia o ancla de sonoridad en  $-23$  LUFS (Unidades de sonoridad referidas al fondo de escala), que se denominará nivel de 0 LU.
- Que los niveles de audio sean ajustados a  $-23$  LUFS, con una compuerta de nivel de  $-10$  LU, en su defecto se aceptarán  $-24$  LUFS sin compuerta.
- Que le valor máximo absoluto del material de audio sea de  $-1$  LUFS.
- Que se procure mantener el rango de sonoridad dentro de los 10 LUFS, para no producir elevados picos de nivel.
- Que el material sea identificado como de poco rango o PRS y de mucho rango o GRS.
- Que las salas de control de audio dispongan de una configuración de parlantes según la ITU BS.775 y se adopte este formato como estándar en la Televisión Digital Nacional.
- Que se incluya en el stream de audio del Transport Stream de la señal ISDB-T los metadatos definidos en la NBR15602-2 sección 5.2.5.
- Que los STB hogareños posean control del audio referido al parámetro de metadatos de segundo nivel (*matrix-downmix* y *prog\_ref\_level*).

**ANEXOS**  
**IMPLEMENTACIONES EN CÓDIGO MATLAB**

## Implementación en MATLAB del código para el filtro RLB

```
% Cálculo de la sonoridad leq(RLB)
% Este script calcula la sonoridad leq(RLB) , para un
% archivo mono o estéreo tipo .wav
% la entrada del script requiere el nombre del archivo .wav
% con la forma 'nombre.wav', la salida es el valor de la
% sonoridad calculada para todo el archivo
```

```
function RLBloudness = RLB(a1)
```

```
% lee el archivo de audio
```

```
audio=wavread(a1);
```

```
% Coeficientes del filtro Leq(RLB) para 48 kHz
```

```
ARLB = [1 -1.99004745483398 0.99007225036621];
```

```
BRLB = [1 -2 1];
```

```
% Cálculo para un archivo estéreo
```

```
if size(audio,2)==2
```

```
    disp('archivo stereo')
```

```
    can1filt = filter(BRLB, ARLB, audio(:,1));
```

```
    can2filt = filter(BRLB, ARLB, audio(:,2));
```

```
    can1filt2=can1filt.^2;
```

```
    can2filt2=can2filt.^2;
```

```
    z11 = mean(can1filt2);
```

```
    z12 = mean(can2filt2);
```

```
    sono = -0.691+10*log10(z11+z12);
```

```
% Cálculos para un archivo mono
```

```
else
```

```
    disp('archivo mono')
```

```
    can1 = filter(Bhead, Ahead, audio(:,1));
```

```
    can1filt = filter(BRLB, ARLB, can1(:));
```

```
%    can1filt2 = af1ch1filt.^2;
```

```
    can1filt2 = can1filt.^2;
```

```
    z11 = mean(can1filt2);
```

```
    sono = -0.691+10*log10(z11);
```

```
end
```

```
RLBloudness = sono;
```

```
end
```

## Implementación en código MATLAB para el algoritmo ITU-R BS. 1770

```

% Cálculo de la sonoridad de acuerdo a ITU-R BS. 1770-1
% Este script calcula el valor de la sonoridad para archivos .wav
% mono o estéreo de acuerdo a ITU-R BS. 1770-1
% la entrada del script requiere el nombre del archivo .wav
% de la forma: 'nombre.wav' la salida calcula el valor de la sonoridad
% integrada para todo el archivo

function ITUloudness = ITUOriginal(a1)
% lee archivo de audio
audio1=wavread(a1);

% Coeficientes del filtro para el modelado de los efectos
% acústicos de la cabeza (48 kHz)
Ahead = [1 -1.69065929318241 0.73248077421585];
Bhead = [1.53512485958697 -2.69169618940638 1.19839281085285];

% Coeficiente del filtro RLB (48 kHz)
ARLB = [1 -1.99004745483398 0.99007225036621];
BRLB = [1 -2 1];

% Cálculos del filtrado para archivo estéreo (2 canales)
if size(audio1,2)==2
    can1 = filter(Bhead, Ahead, audio1(:,1));
    can2 = filter(Bhead, Ahead, audio1(:,2));

    can1filt = filter(BRLB, ARLB, can1(:));
    can2filt = filter(BRLB, ARLB, can2(:));

    can1filt2 = can1filt.^2;
    can2filt2 = can2filt.^2;

    z11 = mean(can1filt2);
    z12 = mean(can2filt2);

    sono = -0.691+10*log10(z11+z12);

% Cálculo del filtrado para archivo mono (1 canal)
else
    can1 = filter(Bhead, Ahead, audio1(:,1));
    can1filt = filter(BRLB, ARLB, can1(:));
    can1filt2 = can1filt.^2;
    z11 = mean(can1filt2);

    sono = -0.691+10*log10(z11);

end
ITUloudness = sono;
end

```

## Implementación del código MATLAB para RLB con compuerta fija

```

% Cálculo de la sonoridad RLB con la función compuerta
% Cálculo de acuerdo a RLB
% pero los valores debajo del umbral se ignoran

function RLBgate = RLBgate(a1)
% lee el archivo de audio
audio = wavread(a1);

% Coeficientes del filtro RLB para 48 kHz
ARLB = [1 -1.99004745483398 0.99007225036621];
BRLB = [1 -2 1];

% Valor de la compuerta de discriminación
% (el valor corresponde a -50 dBFS)
% Cambiar para distintos valores de umbral
% 0.00001 = -50 dBFS
% 0.001 = -30 dBFS
% 0.01 = -20 dBFS

umbral = 0.00001;

% Cálculo de acuerdo a RLB para archivo estéreo
if size(audio,2)==2
    can1filt = filter(BRLB, ARLB, audio(:,1));
    can2filt = filter(BRLB, ARLB, audio(:,2));

    % Se elevan los valores al cuadrado
    can1filt2 = can1filt.^2;
    can2filt2 = can2filt.^2;

    % Se ordenan los valores elevados al cuadrado
    ordcan1 = sort(can1filt2);
    ordcan2 = sort(can2filt2);

    % Busca en los vectores ordenado los índices menores que el umbral
    a = find(ordcan1 < umbral);
    b = find(ordcan2 < umbral);

    % Calcula el valor medio incluyendo solo los valores sobre
    % el umbral en los vectores ordcan1 y ordcan2
    z11 = mean(ordcan1((length(a)+1):length(ordcan1)));
    z12 = mean(ordcan2((length(b)+1):length(ordcan2)));

    sono = -0.691+10*log10(z11+z12);

% Cálculo de acuerdo a RLB para archivo mono
else
    can1filt = filter(BRLB, ARLB, audio(:,1));
    can1filt2 = can1filt.^2;

    % Se ordenan los valores al cuadrado
    ordcan1 = sort(can1filt2);

    % Busca el índice en los vectores donde el umbral es mayor
    a = find(ordcan1 < umbral);

    % Calcula el valor medio cuadrático solo para los valores

```

```
% sobre el umbral en el vector ordcan1  
z11 = mean(ordcan1((length(a)+1):length(ordcan1)));
```

```
sono = -0.691+10*log10(z11);
```

```
end
```

```
    RLBgate = sono;
```

```
end
```



## Implementación del código MATLAB para ITU BS. 1770 con compuerta fija

```

% Cálculo según ITU-R BS. 1770 con función de compuerta fija
% Cálculo según el estándar ITU, pero los valores debajo
% del umbral son ignorados

function ITUG = ITUGate(a1)
% lee el archivo de audio
audio = wavread(a1);

% Coeficientes del filtro de efecto
% acústico de la cabeza (48 kHz)
Ahead = [1 -1.69065929318241 0.73248077421585];
Bhead = [1.53512485958697 -2.69169618940638 1.19839281085285];

% Coeficiente del filtro RLB (48 kHz)
ARLB = [1 -1.99004745483398 0.99007225036621];
BRLB = [1 -2 1];

% Valor del umbral de discriminación este valor corresponde a -10 dBFS
% Cambiar para distintos valores de umbral
% 0.00001 = -50 dBFS
% 0.001 = -30 dBFS
% 0.01 = -20 dBFS

umbral = 0.001;

% Cálculos de acuerdo a ITU-R BS. 1770-1 (estéreo)
if size(audio,2)==2
    can1 = filter(Bhead, Ahead, audio(:,1));
    can2 = filter(Bhead, Ahead, audio(:,2));

    can1filt = filter(BRLB, ARLB, can1(:));
    can2filt = filter(BRLB, ARLB, can2(:));

    % Se elevan los valores al cuadrado
    can1filt2 = can1filt.^2;
    can2filt2 = can2filt.^2;

    % Se ordenan los valores elevados al cuadrado
    ordcan1 = sort(can1filt2);
    ordcan2 = sort(can2filt2);

    % Busca en los vectores ordenado los índices donde el umbral
    % es mayor
    a = find(ordcan1 < umbral);
    b = find(ordcan2 < umbral);

    % Calcula el valor medio incluyendo solo los valores sobre
    % el umbral en los vectores ordcan1 y ordcan2
    z11 = mean(ordcan1((length(a)+1):length(ordcan1)));
    z12 = mean(ordcan2((length(b)+1):length(ordcan2)));

    sono = -0.691+10*log10(z11+z12);

% Cálculo de acuerdo a ITU-R BS. 1770-1 (archivo mono)
else
    can1 = filter(Bhead, Ahead, audio(:,1));
    can1filt = filter(BRLB, ARLB, can1(:));

```

```
can1filt2 = can1filt.^2;

% Se ordenan los valores al cuadrado
ordcan1 = sort(can1filt2);

% Busca el índice en los vectores donde el umbral es mayor
a = find(ordcan1 < umbral);

% Calcula el valor medio cuadrático solo para los valores
% sobre el umbral en el vector ordcan1
z11 = mean(ordcan1((length(a)+1):length(ordcan1)));

sono = -0.691+10*log10(z11);

end

ITUG = sono;

end
```

## Implementación en código MATLAB para el algoritmo ITU-R BS. 1770

```

% Con compuerta dinámica
% Cálculo de la sonoridad de acuerdo a ITU-R BS. 1770-1 con
% compuerta dinámica
% Este script calcula el valor de la sonoridad para archivos wav
% de acuerdo a ITU-R BS. 1770-1 con compuerta dinámica
% La entrada del script requiere el nombre del archivo .wav
% y la cantidad de LU de compuerta. La salida calcula
% el valor de la sonoridad integrada aplicando una compuerta
% de discriminación.

function ITUdb = ITUdb(a1,db)
% Lee el archivo de audio
audio1=wavread(a1);

% Coeficientes del filtro para el modelado de los efectos
% acústicos de la cabeza (48 kHz)
Ahead = [1 -1.69065929318241 0.73248077421585];
Bhead = [1.53512485958697 -2.69169618940638 1.19839281085285];

% Coeficiente del filtro RLB (48 kHz)
ARLB = [1 -1.99004745483398 0.99007225036621];
BRLB = [1 -2 1];

% Cálculos del filtrado para archivo estéreo (2 canales)
if size(audio1,2)==2
    can1 = filter(Bhead, Ahead, audio1(:,1));
    can2 = filter(Bhead, Ahead, audio1(:,2));

    can1filt = filter(BRLB, ARLB, can1(:));
    can2filt = filter(BRLB, ARLB, can2(:));

    can1filt2 = can1filt.^2;
    can2filt2 = can2filt.^2;

    z11 = mean(can1filt2);
    z12 = mean(can2filt2);

    sono = -0.691+10*log10(z11+z12);

% Cálculo del filtrado para archivo mono (1 canal)
else
    can1 = filter(Bhead, Ahead, audio1(:,1));
    can1filt = filter(BRLB, ARLB, can1(:));
    can1filt2 = can1filt.^2;
    z11 = mean(can1filt2);

    sono = -0.691+10*log10(z11);

end

% Cálculo de dB de umbral
x = str2num(db);
umbral = sono-x;

% Calcula el valor diferencia de umbral
umdin=10.^(umbral/10);

```

```

% Cálculo de valor de sonoridad con compuerta dinámica estéreo
if size(audio1,2)==2
    % Se ordenan los valores elevados al cuadrado
    ordcan1 = sort(can1filt2);
    ordcan2 = sort(can2filt2);

    % Encuentra el índice del vector, donde la sonoridad es menor
    % al umbral
    a = find(ordcan1 < umdin);
    b = find(ordcan2 < umdin);

    % Calcula el valor medio incluyendo solo los valores sobre
    % el umbral en los vectores ordcan1 y ordcan2
    z11 = mean(ordcan1((length(a)+1):length(ordcan1)));
    z12 = mean(ordcan2((length(b)+1):length(ordcan2)));

    sonodin = -0.691+10*log10(z11+z12);

% Cálculo de acuerdo a ITU-R BS. 1770-1 (mono)
else
    can1 = filter(Bhead, Ahead, audio1(:,1));
    can1filt = filter(BRLB, ARLB, can1(:));
    can1filt2 = can1filt.^2;

    % Se ordenan los valores al cuadrado
    ordcan1 = sort(can1filt2);

    % Busca el índice en los vectores donde el umbral es mayor
    a = find(ordcan1 < umdin);

    % Calcula el valor medio cuadrático solo para los valores
    % sobre el umbral dinámico en el vector ordcan1
    z11 = mean(ordcan1((length(a)+1):length(ordcan1)));

    sonodin = -0.691+10*log10(z11);

end

ITUdb = sonodin;

end

```

## **GLOSARIO**

### **Reproducción o Playback Digital**

Un dispositivo de audio de PC tiene varios caminos para la señal que pueden estar caracterizados independientemente y en variadas combinaciones.

Esta conexión va desde el bus interno de la PC ruteado a través de un conversor digital – analógico, (C-D/A) el mezclador de Play y el amplificador de salida.

Las señales digitales estarán presentes en el bus de la PC cuando se reproduzcan o se transmiten archivos de ondas (wave) desde el disco rígido, desde la memoria RAM o desde una entrada digital al bus de la PC.

### **Lazo Analógico**

Es la conexión entre la entrada analógica ruteada por el mezclador de reproducción y enviada de vuelta a la salida analógica.

Pueden existir varias entradas analógicas tales como: entrada de línea, micrófono, auxiliar, CD, audio del MODEM telefónico (TDA) y otras.

### **Grabación o Record Digital**

Es la conexión entre la entrada analógica ruteada por el mezclador de grabación y a través del conversor analógico – digital al bus de la PC. La señal digital en el bus puede residir en memoria o ser enviada al disco rígido. Alternativamente la señal puede ser enviada hacia la salida digital.

### **Lazo Grabación – Reproducción (Record-Play Loop)**

Esta es una combinación de los alzos de grabación y reproducción a través del bus de la PC. La señal desde el mezclador analógico de entrada se envía al mezclador analógico de grabación al conversor analógico – digital al bus de la PC y a través del conversor digital analógico y del mezclador de reproducción a la salida analógica.

### **Lazo Digital**

Es la conexión que corre desde la entrada digital a través del bus de la PC a la salida digital. Estas entradas están definidas por la norma IEC 60958-3 también llamada S/PDIF.

### **Mezclador de Grabación**

Es la parte del driver asociado al dispositivo de audio de la PC que forma la interface de entrada o de grabación, está más asociada a las conexiones analógicas y ocasionalmente incluye controles para las digitales. Este mezclador tiene varios atenuadores uno para cada entrada analógica y no están calibrados por lo que no es posible especificar condiciones de ganancia específicas, particularmente la unitaria a demás asociado con cada entrada comúnmente hay una celda para marcar, seleccionando la entrada, lo que lo hace mutuamente exclusivo siendo más un selector que un mezclador.

### **Mezclador de Reproducción**

Es la parte del driver del dispositivo de audio de la PC asociado a la reproducción. Como el de grabación está más asociado a las entradas analógicas y ocasionalmente contiene controles para las digitales. Tiene varios atenuadores que controlan el nivel de salida de cada entrada analógica. Tampoco están calibrados, por lo que es imposible especificar una ganancia determinada. Cada uno contiene una celda que permite silenciar su salida. Es posible tener más de una salida habilitada por lo que se comporta bien como un mezclador de señales.

### **Archivos de Onda (Wav)**

Un archivo de ondas con la extensión "WAV", contiene la información de audio en PCM, este formato se usa para almacenar y transferir audio digital entre dispositivos.

Comúnmente reside en la memoria o en el disco rígido. La señal de audio puede ser mono estéreo o multicanal, indicada en la cabecera identificadora que especifica las características de la señal digitalizada. Este tipo de archivo en televisión se muestra a 48 Khz con 20 o 24 bits. En audio doméstico es de 44.1 Khz, a 16 bits.

### **Decibeles de Fondo de Escala (Full-Scale decibels) (dBFS)**

Es la amplitud eficaz de la señal digital expresada en decibeles, relativa a la señal de fondo de escala. Se utiliza 20 veces el logaritmo de la amplitud de la señal dividido por la amplitud máxima. Este dBFS expresa niveles de señal digital y no debe usarse para indicar niveles de señales analógicas.

### **Máximo Nivel de Señal (analógico)**

El máximo nivel está referido a 0.5 dB por debajo de la máxima señal sin recorte para una señal sin distorsión, que en este contexto significa que no debe exceder los  $-40$  dB de THD + N.

**Nivel de Señal de Referencia**

Es la amplitud de señal usada para computaciones radiométricas como el rango dinámico y la relación señal ruido. La referencia se estipula como 0 y puede ser el máximo nivel o la amplitud a fondo de escala.

**Respuesta en Frecuencia**

Rango de frecuencia donde la salida casio varía en relación a la salida a 997 Hz. Los límites de variación deben especificarse y comúnmente son de +/- 1 dB o +/- 3 dB. Para medir en sonoridad se utiliza también 440 Hz, por ser la parte unitaria de la curva de respuesta en frecuencia.

**Distorsión Armónica Total (THD+N)**

Esta expresada como 20 veces el logaritmo de la relación entre la amplitud eficaz de todas las señales más el ruido y la amplitud de la señal de prueba, dentro del rango de frecuencia de prueba. Las mediciones son dependientes de la amplitud por lo que debe especificarse el nivel de referencia usado.

**Distorsión por Intermodulación**

Es la medición de la distorsión producida por la interacción de dos o más señales. La técnica más usada es una señal compuesta por dos senoidales y se mide su producto. Las frecuencias son 18 KHz para señales analógicas y 17.987 Hz y 19.997 Hz para las digitales. Esta medición permite obtener valores de distorsión válidos para la banda alta de frecuencia.

**Rango Dinámico**

Es 20 veces el logaritmo de la relación entre el nivel de señal de referencia y el ruido de fondo en valores eficaces, expresado en dB.

Esta especificación es a veces mal llamada relación señal ruido (SNR). Pero no debe confundirse porque los niveles en ausencia de señal son engañosos ya que actúan circuitos enmudecedores cuando no hay señal presente.

**Diafonía (Cross-talk)**

La pérdida de información de un canal hacia otro expresada como 20 veces el logaritmo de la relación entre la señal de prueba medida a la salida del canal y la amplitud de la señal medida sobre otro canal sin señal de entrada, pero con su entrada terminada.

**Diferencia de Fase**

Es la repuesta en fase entre los canales referida a una referencia. La diferencia de fase entre los canales y el de referencia debe reportarse como grados en función de la frecuencia.



### **Exactitud en la Frecuencia de Muestreo**

Es la comparación entre la frecuencia utilizada para el muestreo de las señales de entrada y una referencia patrón. Se define en porcentaje de frecuencia expresada en Hz.

Por ejemplo: si la frecuencia teórica es de 48.000 Hz se puede medir 48.001,45 Hz, un error de 1,45 Hz o sea 0,003%.

### **LKFS**

Es la ponderación de nivel absoluta, basada en una escala en decibeles filtrada por la influencia de la cabeza y la respuesta del oído a niveles medios (70 Fones), referenciada al fondo de escala digital. Corresponde a L: nivel, K: ponderación de frecuencia usada, FS: fondo de escala. En definitiva son decibeles pero ponderados con una curva especial. No son valores pico, son valores eficaces.

### **LU**

Definido por al EBU como niveles de sonoridad, en reemplazo de los LKFS. Si se utilizan en modo absoluto (referidos al máximo de señal) son LUFS. Respetando el formato ISO.

### **Lista de abreviaturas utilizadas**

AES	Sociedad de Ingeniería de Audio – Audio Engineering Society
CENELEC	Comité Européen de Normalisation Electrotechnique (European Committee for Electrotechnical Standardization)
CODEC	Codificador y decodificador Encoder and decoder system (coder/decoder)
dB	decibel
dBFS	decibel referido a fondo de escala.
dBTP	Decibel de pico real – decibel True Peak.
DD	Dolby Digital (AC3)
DRC	Control de Rango Dinámico – Dynamic Range Control.
EBU	Unión Europea de Radiodifusión – European Broadcasting Unión.
ETSI	European Telecommunications Standards Institute.
FM	Modulación de Frecuencia – Frequency Modulation
HE-AAC	Codificación de Audio Avanzada de Alta Eficiencia – High Efficiency Advanced Audio Coding.

IEC	Comisión Internacional Electrotécnica – International Electro - Technical Commission.
ISO	Organización Internacional de Estándares – International Standards Organization.
ITU-R	Unión Internacional de Telecomunicaciones Sector de radio – International Telecommunications Union / Radio Communications Sector
LU	Unidad de sonoridad – Loudness Unit
LUFs	Unidad de sonoridad a fondo de escala – Loudness Unit to Full Scale
LKFS	Sonoridad ponderada K fondo de escala – Loudness K relative to Full Scale
MPEG	Grupo de expertos en imágenes – Moving Pictures Experts Group
SPDIF	Interface digital Sony Philips – Sony Philips Digital Interface
PCM	Modulación por códigos de pulso – Pulse Code Modulation
PID	Identificador de paquetes – Packet Identifier
PPM	Medidor de picos de programa – Peak Programme Meter
PRL	Nivel de programa – programme Reference Level (prog_ref_level ISO/IEC 144496-3)
RMS	Raíz media cuadrática – Root Mean Square
SDI	Interface serial digital – Serial Digital Interface
THX	Tomlinson Holman's Experiment
RF	Radio Frecuencia
TL	Nivel Destino Target Level (for HE-AAC systems: target_level specified in ISO/IEC 144496-3)

## BIBLIOGRAFÍA

- [1] Glen M. Ballou, (1992) "*Handbook for sound engineers*", © SAMS, ISBN 0-672-22752-5.
- [2] Michael Robin, (2002) "*Audio Signal Distribution & Level Measurements*", Broadcasting Engineering Digital Handbook, Intertec Publishing Corp., pp: 20 – 26.
- [3] H.A. Chinn, D.K. Morris, R.M. Morris (1940) "*A new Standard Volume Indicator and Reference Level*", IEEE, Proceedings Vol 28, Pages 1 - 17. USA.
- [4] Lund, Thomas, (2004) "*Distortion to the People*" AES Conference Proc. Of the 23<sup>rd</sup> Tonmeistertagung. Leipzig 2004. Paper A05, TC Electronics, Risskov, Denmark.
- [5] Bonello, Oscar, (2008) "*Procesado Psicoacústico del Audio en FM*", VI Congreso Iberoamericano de Acústica, FIA 2008 – A017. Argentina.
- [6] Jones, Waive, et al. (2003), "*Testing Challenges in PC Audio Devices*", AES 114<sup>th</sup> Convention, Paper 5814, Amsterdam.
- [7] Intel Corp. (2010), *Especificación de Audio de Alta Definición*, "AZALIA", Junio, <http://www.intel.la/content/www/xl/es/standards/standards-high-def-audio-specs-general-technology.html> (último acceso Dic. 2013)
- [8] AEG Sutcliffe y M. Yonge, (1994 - 1998) "*Computers and professional audio*", "*Workstations Moving into the Studio*", IEE Colloquium, pp: 1 – 4 London, INSPEC:4865060.
- [9] AES-6id-cf061110, (2006) – *Information Document for Digital Audio* – "Personal Computers – Audio Quality measurements" (Draft), AES Standards
- [10] Skovenborg, E. & Nielsen, S.H. (2005) "*Evaluation of Different Loudness Models with Music and Speech Material*", AES 117<sup>th</sup> Convention, San Francisco, CA.
- [11] Brixen, E.B. (2010), "*Audio Metering*", Broadcasting Publishing, Focal Press, DK Audio A/S. ISBN: 978-0-240-81467-4.
- [12] Ir. Maykel Post. (2003), "*Measuring Audio Specifications*", Axon Digital Design B. V, Ziff Davis White Papers Library, The Netherlands.
- [13] Wikipedia (2013), "*Sonido*", (último acceso Dic. 2013) <http://es.wikipedia.org/wiki/Sonido>
- [14] Chapuis R. J. (2003) "100 Years of Telephone 1878 - 1960", Part 1m IOS Press, Netherlands, ISBN: 1-58603-349-2.

- [15] Dennis Bohn, (2003) "*Audio Specificactions*", RANE Note N° 145, Paper N° 11628, RANE Corporation Inc. Mukilteo, WA. USA.
- [16] Wikipedia (2013), "*Vúmetro*", (último acceso Dic. 2013), <http://es.wikipedia.org/w/index.php?title=Vúmetro&oldid=66051144>.
- [17] Millard, Andre (1995), "*America on Record: A History of Recorded Sound*". Cambridge University Press, © 1995. USA.
- [18] <http://www.eumus.edu.uy/eme/ensenanza/acustica/apuntes/SistemaAuditivo.pdf>
- [19] Bauer, B.B. et al (1997), "*A loudness-Level Monitor for Broadcasting*", IEEE Transactions on Audio, vol. 15-4, pp 177 – 182. USA.
- [20] Dennis Bohn (2008), "*No Such Thing as Peak Volts dBu*", RANE Note 169, Paper N° 1-08 – Mención a Pat Brown, *Synergetic Audio Concepts*, RANE Corporation INC. Mukilteo, WA. USA.
- [21] Jhon G. McKnight (2006), "*Some Questions and Answers on the Standard Volume Indicator (vumeter)*", Mag. Ref. Laboratory, Mountain View, CA. USA.
- [22] Lund, Thomas (2007), "*Level and Distortion in Digital Broadcasting*" EBU Technical review – April Notes. Switzerland
- [23] Whitaker, Jerry C. (2006), "*Master handbook of audio production: a guide to standards, equipment and system design*", McGraw-Hill Prof, ISBN 0071408762.
- [24] Dorrough (2003) "*Stereo Meters with over indicator – DE-280-D*", Dorrough Electronics Inc. User Manual. USA.
- [25] EBU (1979), "*Standard Peak Programme Meter*", Tech 3205-E, 2° Ed. Nov 1979
- [26] ISO (1975), "*International Standard ISO 532-1 Acoustics. Method for calculating loudness level of steady sounds*", International Organization for Standardization.
- [27] MattiZemack, (2007), "*Implementing Methods for Equal Loudness in Radio Broadcasting*", Master of Science Thesis, Royal Institute of Technology, KTH, Stockholm, Suecia.
- [28] NIST (2008) "*The International System of Units (SI)*", National Institute of Standards and Technology NIST, Special Publication #330
- [29] Jon Schmith y Janet Rutledge (1996), "*Multichannel Dynamic Range Compression for Music Signals*", Acoustics, Speech and Signal Processing, IEEE pp: 1013 – 1016, Atlanta, GA INSPEC:5465063.
- [30] Zwicker, E. et al. (2000), "*Psychoacoustics; Facts and Models*" 2° Ed. Information series 22, Springer – Verlag, Berlin.
- [31] Zwicker E. (2002), "*Audio Engineering and Psychoacoustics...*" parte del libro de Jeffay, Kewvin "*Reading in Multimedia Computing and Networking*", Chapter 1, Digital Audio, Academic Perss, USA, ISBN; 1-55860-651-3.

- [32] Greenberg, S. (2001), "*Computational Models of Auditory Function*", IOS Press.
- [33] Richard M. Warren (2008), "*Auditory Perception: Analysis and Synthesis*", Wisconsin University, Milwaukee, Cambridge Pres; 3° Ed. ISBN: 9780521688895.
- [34] Wikipedia (2013), "*Loudness*", (último acceso Dic. 2013),  
<http://en.wikipedia.org/wiki/Loudness>.
- [36] Bob Katz "*The Lost Interview*" (2013), Written by Brent Randall, © ProRec.com  
<http://www.prorec.com/index.php/home/interviews/item/63-bob-katz-the-lost-interview-part-1.html> último acceso: Nov 2013
- [37] UIT-R BS.775-2 (2006), "*Sistema de sonido estereofónico multicanal con y sin acompañamiento de imagen*", recomendación ITU-R, grupo R-6
- [38] UIT-R BS. 1770-1 (2007), "*Algorithms to measure audio programme loudness and true-peak audio level*", ITU-R recommendation 2/6.
- [39] UIT-R BS.1771 (2007), "*Requisitos de los medidores de sonoridad y de valores de cresta reales*", recomendación UIT-R grupo 2/6.
- [40] ATSC (2009), "*Recommended Practice: Techniques for Establishing and Maintain Audio Loudness for Digital Television*", Advanced Television Systems Committee, Inc.Doc A/85:2009, Nov 2009, USA.
- [41] CST\_RT – 017 – TV – 2009 V.2 (2009), "*Recommendation Technique PAD Diffuseurs*" (CST/FICAM/HDFORUM), Commission Supérieure Technique de l'image e du son, France.
- [42] CST\_RT – 019 – TV – 2009 (2009), "*Methodology for evaluating sound dynamics for television broadcast*", (RT019 – TV Dynamic measurement methodology for RFB – v1f.doc), France.
- [43] ARIB STD-B32, (2007), "*Audio transmission in ISDB-T*", Part 2, Version 2.1 – E1, ISDB-T Technical Specification, Japan.
- [44] EBU R-128, (2010), "*Loudness normalization and permitted maximum level of audio signals*", EBU Recommendation – P/LOUD Task Group. Switzerland.
- [45] ABNT NBR 15602-2 (2007), "*Codificación de video, audio y multiplicación Parte 2: Codificación de audio*", Documento normativo del sistema SBTVD, Brasil.
- [46] ISO/IEC 13818-7 (2004), Part 7, "*Advanced Audio Coding (AAC)*", International Standard, 3° Edition, Switzerland.
- [47] Enrique Soto (2002), Elementos No. 44, Vol. 8, Dic. 2002, Página 11, Instituto de Fisiología de la BUAP, [www.elementos.buap.mx](http://www.elementos.buap.mx), México (accedido Nov. 2010)
- [48] Iglesias Simón, Pablo (2004), "*El diseñador de sonido: función y esquema de trabajo*", *ADE-Teatro* N° 101. Julio – Agosto Págs. 199 – 215 España.

- [49] Public Broadcasting System – PBS (2010), “*Technical Operating Specifications (TOS)*”, Program Submission, PBS TOS-1, 2010 Edition. USA.
- [50] Norcross; Scott G; Lavoie; Michel C. (2009), “*Investigations on the Inclusion of the LFE Channel in the ITU-R BS.1770-1 Loudness Algorithm*” AES Convention Paper 7829, 127<sup>th</sup> AES Convention, New York.
- [51] Earl Vickers, (2001), “*Automatic Long-term Loudness and Dynamics Matching*”, presentada en la convención AES 111 New York.
- [52] Gilbert Soulodre, Scott Norcross (2003), “*Objective Measures of Loudness*”, presentada en la convención AES 115 New York.
- [53] Earl Vickers, (2010), “*Metrics for Quantizing Loudness and Dynamics*”, presentada en la convención AES 119, San Francisco, CA. USA.
- [54] <http://www.replaygain.org/>, último acceso 20 febrero 2011.
- [55] C. Zhang & F. G. Zeng (1966), “*Loudness of Dynamics Stimuli in Acoustic Hearing*”, Journal of Acoustic Society of America, vol. 40, pp. 71 – 78, USA.