



UNIVERSIDAD ABIERTA INTERAMERICANA
Facultad de Tecnología Informática

Tesis de la Maestría en Tecnología Informática

Medición de la Sonoridad
del Audio Digital

Víctor M. S. Acuña

Tutores: Ing. José Simonetta – Dr. Carlos Neil

Director de la Maestría: Dr. Gustavo Rossi

Secretario Académico: Dr. Carlos Neil

Buenos Aires, Mayo de 2012

Resumen

La transmisión en los medios de comunicación audiovisual de contenidos digitales ha perfeccionado y mejorado la calidad técnica de los mismos. Específicamente, la posibilidad de contar con grandes rangos dinámicos ha permitido que los contenidos sonoros se difundan sin merma o modificación alguna desde las fuentes originales, manteniendo la identidad y el sentido artístico dados por el autor en su creación.

Disponer de este gran rango dinámico trae un nuevo problema, ya que los radiodifusores y los oyentes necesitan un recurso para evitar los abruptos cambios en el nivel de sonido transmitido y por lo tanto un ajuste constante del nivel del volumen.

La solución tradicional de normalizar por los valores de pico, ha traído una serie de graves consecuencias. La búsqueda se ha centrado, entonces, en la obtención de un nivel correcto de sonoridad para los distintos contenidos emitidos. A raíz de los distintos estudios realizados se ha elaborado un modelo de las características auditivas, y a partir de este se están estableciendo normativas para obtener niveles correctos de emisión.

Este trabajo proveerá las bases de estudio sobre las que se han realizado estas investigaciones del modelo de sonoridad, se analizarán las recomendaciones emanadas de los organismos de estandarización, presentando un algoritmo de aplicación que propone una mejora en las mediciones de la sonoridad en el modo integral, el diseño de un medidor de sonoridad momentánea y finalmente se propondrán los lineamientos de las recomendaciones pertinentes para una reglamentación en nuestro país según las pautas de medición en la sonoridad del audio digital.

Abstract

The digital transmission of audiovisual contents has allowed the technical quality of these get better. In the area of sound, the possibility of having, among other improvements, a greater dynamic range, allows that audio content reach listeners without loss or modification from the original sources, maintaining the artistic sense of the author. But this has brought a new problem; it is that broadcasters and listeners need a resource to avoid the abrupt changes in the transmitted sound level that forces listeners to adjust the volume level of their receivers constantly. The traditional solution to normalize for the peaks has brought a series of serious consequences. The search has been centered, then on obtaining the right level of loudness for different emitted contents. From the realized studies, a model of the auditory characteristics has been developed an accepted, and from this, are dictating ground rules to obtain correct levels of emission.

This work provides the bases of studies, on which these investigations have been conducted, about the loudness model presented. The recommendations from the organisms of standardization, and will be presented an application algorithm that proposes an improvement on the measures in the integral mode, presenting the designing of a meter for momentary loudness, and the guidelines of the pertinent recommendations for its regulation at our country, keeping the measurement guidelines of the loudness in the digital audio.

Agradecimientos:

A mi familia y a mis tutores.

ÍNDICE

<i>Portada</i>	
<i>Resumen</i>	
<i>Abstract</i>	
<i>Agradecimientos</i>	
<i>Visión Global de la Tesis</i>	4
1: INTRODUCCIÓN	
1.1 Fundamentación	5
1.2 Objetivos	7
1.3 Descripción del problema	8
1.4 Metodología	8
1.5 Contribuciones de la Tesis	9
2: DESARROLLO TEÓRICO	
2.1 Antecedentes del tema	11
2.2 Medición del nivel de la señal de audio	12
2.2.1 El sonido	12
2.2.2 Los valores eléctricos	14
2.2.3 El vúmetro	15
2.2.4 Niveles de referencia	17
2.2.5 El medidor de crestas	19
2.2.6 El gálibo y la franquía	21
2.3 La sonoridad	22
2.3.1 Como interpretamos la sonoridad	23
2.3.2 Medición del nivel de señal	28
2.4 Evaluación de la sonoridad	29
2.4.1 Fisiología auditiva	29
2.4.2 Aspectos de la sonoridad	31
2.4.2.1 Percepción de la intensidad	32
2.4.2.2 Unidades de evaluación sonora	32
2.4.2.3 Propiedades espectrales	33
2.4.2.4 Propiedades temporales	34
2.4.2.5 Propiedades espaciales	34
2.4.2.6 Señales de prueba	34
2.4.2.7 Tipos de modelos de sonoridad	35
2.5 Sistema de audio en una PC	36
2.5.1 Situación Actual del audio en PC	39
2.6 Fundamentos de la compresión dinámica	41
2.6.1 El sobre-rango o saturación digital	44
2.6.2 La guerra de la sonoridad	46
3: HIPÓTESIS DE TRABAJO	
Mejorar la medición de la sonoridad	47

4: PROPUESTA DE SOLUCIÓN

4.1 Aportes de las distintas normativas

4.1.1 Recomendación UIT BR.775	48
4.1.2 Recomendación UIT BS.1770-1	49
4.1.3 Recomendación UIT BR.1771-2	51
4.1.4 Documento A/58 ATSC	53
4.1.5 Documento CST RT – 017 v2.0	55
4.1.6 Documento ARIB TR- B29	57
4.1.7 Documento EBU R-128	60

4.2 La sonoridad en el audio digital

4.2.1 Norma Brasileira ABNT NBR 15602-2	62
4.2.2 El caso argentino	64
4.2.3 Análisis de otras Propuestas	65
4.2.4 Análisis de la arquitectura del Audio PC.....	69

4.3 Implementación de un medidor de sonoridad

4.3.1 Criterios	72
4.3.2 Diseño conceptual	73
4.3.3 Experiencias	76
4.3.4 Mediciones objetivas	80

5: CONCLUSIONES

5.1 Conclusiones	82
5.2 Recomendaciones normativas	84
5.3 Futuras líneas de investigación	85
5.4 Actividades relacionadas	87

6: BIBLIOGRAFÍA

Numerada y ordenada por orden de aparición en el texto.....	88
---	----

7: ANEXOS

7.1 Anexo 1 Algoritmos experimentales	94
7.2 Anexo II Definición de términos	103

ÍNDICE DE FIGURAS

Figura 1: Escala de un instrumento vúmetro	16
Figura 2: Señal recortada en amplitud	17
Figura 3: Recorte de los picos por falta de margen	18
Figura 4: Galibo y Franquía de la señal	21

Figura 5: Curvas isofónicas	25
Figura 6: Intensidad expresada en Fones	26
Figura 7: Curva de Ponderación B	28
Figura 8: Partes del oído	29
Figura 9: Rangos de sonoridad	30
Figura 10: Curva de Compresión Dinámica	42
Figura 11: Rangos dinámicos típicos	43
Figura 12: Comportamiento de la distorsión	44
Figura 13: Disposición de los altavoces L, R, C, LS y LS	49
Figura 14: Algoritmo de sonoridad	50
Figura 15: Escalas de los distintos niveles según la norma RT-17	56
Figura 16: Diagrama de flujo de la señal como un filtro de 4° orden	59
Figura 17: Respuesta en frecuencias del pre-filtro	59
Figura 18: Diagrama de la sección de audio en una PC	71
Figura 19: Regresión lineal	76
Figura 20: Curva de ecualización aplicada	78
Figura 21: Diagrama del algoritmo del prototipo	78
Figura 22: Diagrama del algoritmo en software	80
Figura 23: Gráficos de correlación lineal	82

ÍNDICE DE TABLAS

Tabla 1: Rango de Audición	15
Tabla 2: Transferencia de las distintas ponderaciones	27
Tabla 3: Ecuaciones de matrización de codificación y decodificación	48
Tabla 4: Posición de los parlantes multicanal según BS.775	50
Tabla 5: Coeficientes de ganancia de los canales	51
Tabla 6: Características medidores de sonoridad	51
Tabla 7: Coeficientes del pre-filtro de entrada	58
Tabla 8: Indicador de las asignaciones de calidad del audio	59
Tabla 9: Valores de las mediciones experimentales	80
Tabla 10: Valores de sonoridad para distintos umbrales	81
Tabla 11: Corrección de los metadatos del ISDB-T	86

Visión Global de la Tesis

En este documento se trata el nuevo paradigma de medición del nivel sonoro originado en la necesidad de equilibrar el nivel de los programas de audio digital. Este cambio revolucionario consiste en dejar de nivelar el audio por la amplitud de sus picos, a favor de nivelarlo por su sonoridad. Esto involucra la utilización de un algoritmo que represente objetivamente la sensación de sonoridad *subjetiva* del ser humano, cuyo valor estandarizado permitirá obtener un nivel homogéneo en la difusión de los distintos contenidos. Si bien el tema de medir niveles de audio, ha pertenecido tradicionalmente al ámbito electrónico, el procesado digital del audio involucra el uso de tecnología informática en un área tan específica que se ha convertido en una especialización. Por lo cual, en este documento se juzga pertinente un análisis profundo de los conceptos básicos de medición del nivel del audio, con el siguiente desarrollo:

En el **Capítulo 1** se describen los objetivos de este trabajo, la metodología empleada para encontrar una solución, y las contribuciones que se esperan obtener.

En el **Capítulo 2** se desarrollan los conceptos fundamentales, exponiendo los antecedentes aplicados en las técnicas de medición tradicionales de los niveles de audio. Llegando al concepto de sonoridad, al funcionamiento de nuestro sistema de audición, y al análisis de los modelos de sonoridad contemporáneos. Incluyendo la descripción del funcionamiento del audio en una PC que se utilizará en las experiencias.

En el **Capítulo 3** se plantea la hipótesis de trabajo, partiendo de la definición del problema y en el contexto actual se considera oportuno presentar un algoritmo de evaluación de la sonoridad para mejorar su eficacia.

En el **Capítulo 4** se consideraran las propuestas de solución. primero se analizan los aportes de las distintas investigaciones realizadas sobre el tema, ya que cada documento analizado proporciona recursos al modelo utilizado en la solución planteada en esta tesis. Se evalúa el procesado del audio en el estándar de Televisión Digital Terrestre ISDB-T, donde este trabajo de investigación podrá ser aplicado. Se discuten también otras soluciones y se presenta el desarrollo de un modelo de medidor de sonoridad que contempla los nuevos componentes propuestos.

En el **Capítulo 5** se exponen las conclusiones arribadas en las experiencias de medición, así como las recomendaciones para la normativa del audio en televisión digital. Se mencionan las actividades llevadas a cabo a raíz del desarrollo de este trabajo. Y se presentan los posibles proyectos para ampliar estas investigaciones.

1 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 con el término Hi Fi ó Alta Fidelidad.

En este siglo la masiva utilización de tecnologías digitales, permitió llegar aun 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 comienzan a evaluar y a corregir.

Este trabajo intenta ser un aporte a la nueva medición de los niveles del audio digital. Las citas bibliográficas se indicaran utilizando la referencia numérica entre corchetes [].

1.1 Fundamentación

Desde el momento en que la tecnología permitió convertir la presión acústica de un **sonido** en una señal eléctrica (**audio**), se planteó la necesidad de controlar su nivel eléctrico, equivalente a su volumen sonoro. Tradicionalmente, esta magnitud se la ha visualizado y ponderado por medio de un medidor consistente en un instrumento de aguja ó galvanómetro d'Arsonval, y también por medio de diodos electroluminiscentes (leds), o gráficamente en una pantalla. [1]

En radiodifusión, las exigencias de controlar y homogeneizar los niveles de la señal, en su procesamiento y posterior emisión, hizo necesaria la estandarización del equipamiento de estudios, enlaces y planta transmisora, requiriendo de estándares que permitieran la industrialización de sus componentes 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 ó 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. [5]

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 día la radiodifusión y las compañías productoras de música usan una plétora 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 presentará el modelo de sonoridad obtenido de estudios y experiencias 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.

Para encarar este desarrollo, se considera necesario conocer ciertos conceptos ligados al tema específico de esta tesis, como ser, la características físicas del sonido, la fisiología auditiva humana, la metodología de medición del nivel sonoro, la métrica aplicada, los requerimientos en producción, los recursos informáticos utilizados y los algoritmos de medición de la sonoridad, que serán tratados en este trabajo.

La información obtenida de las fuentes de investigación consultadas permitió el abordaje de una visión de referencia histórica, a una solución de aplicación actual. De su análisis se obtuvo una perspectiva de los criterios, un panorama de las soluciones 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. A partir de diversos estudios metodológicos, se pudo obtener un modelo objetivo de la percepción subjetiva de la sonoridad, de manera que el estado del arte y el cambio

tecnológico en la utilización de los nuevos dispositivos informáticos de audio, se han orientado a obtener un perfil sonoro homogéneo (y satisfactorio) en toda distribución de material de audio. Lo que puede concluirse es que la investigación científica construye inferencias que enmarcan las observaciones específicas, y se materializan en normativas tecnológicas.

Este trabajo plantea la inminente 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 una nueva normativa en la medición y determinación de los niveles de sonoridad, con lo que se pretende la posterior determinación de una normativa correspondiente.

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]

1.2 Objetivos

Los objetivos generales de esta tesis son:

- Sistematizar los conocimientos en el tema.
- Proponer una mejora al algoritmo de medición de la sonoridad.
- Esbozar los lineamientos para una recomendación de calidad del audio digital.

Como objetivos específicos, se pretende:

- Crear un documento de referencia que contenga la problemática y el desarrollo de los medidores de nivel de audio digital.
- Diseñar y probar un concepto de medidor de sonoridad, realizando un prototipo.
- Analizar y evaluar las posibles mejoras al algoritmo de medición de la sonoridad.
- Proponer los lineamientos para una normativa sobre la especificación de niveles de audio en producciones audiovisuales.

Estos criterios se aplicarán en una experiencia de medición y correlación de los niveles de sonoridad en una PC, utilizando material de programa de audio y aplicando metodología cualitativa con una recolección de información sumamente abierta.

Se compondrá este documento de referencia integrando de manera sistemática el desarrollo histórico, considerando que será un elemento útil para posteriores innovaciones.

Esbozando los criterios a tomar en cuenta para una normativa sobre la utilización de estos parámetros en las transmisiones de Televisión Digital Terrestre en la Republica Argentina.

1.3 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 día tenemos la posibilidad de obtener registros sonoros prácticamente perfectos. La utilización de computadoras como estaciones de trabajo de audio digital (DAW), es hoy la manera común de resolver una producción. [8] Aunque existen fortalezas y debilidades de los sistemas informáticos cuando se intentan 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 ser la utilización de bases de datos, redes, edición no lineal, duplicación sin pérdidas, inclusión de metadatos, etc.

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.

1.4 Metodología

La metodología que se adopta para dar una solución al problema planteado es mostrar los marcos de aplicación en un análisis descriptivo.

En la exploración del material bibliográfico utilizado para determinar el avance de los estudios previos sobre el tema, se ha decidido utilizar una técnica cualitativa para integrar en profundidad las investigaciones realizadas.

El análisis descriptivo de los documentos involucró el examen de las métricas utilizadas para descubrir las variables inherentes a la subjetividad sonora con la que se trabajó en su elaboración, utilizando un enfoque holístico, permitiendo producir un conocimiento válido a partir de casos concretos.

De acuerdo a esto se propuso el siguiente esquema de investigación:

1. **Analizar** el fenómeno de la sensación de sonoridad subjetiva.
2. **Justificar** las razones de porque es necesaria su medición con exactitud.
3. **Analizar** las soluciones existentes y las exigencias en el mundo digital.
4. **Proyectar** el futuro de los medidores de sonoridad multicanal.
5. **Proponer** el uso de medidores de sonoridad normalizados, realizando un acercamiento normativo.

Produciendo un documento de tesis que muestre el estado actual y las tendencias en el área, y describa los requisitos necesarios para una medición objetiva de los parámetros de sonoridad del audio digital.

1.5 Contribuciones

Este documento permitirá la difusión de la nueva medición de la sonoridad en idioma castellano (escasa ó inexistente). Siendo una recopilación sistemática de las investigaciones que han llevado al cambio de paradigma en la medición de los niveles de audio. La descripción de los requerimientos en la producción de material de audio digital y su post-producción aplicando tecnología informática podrá impulsar el desarrollo de hardware y software específicos.

La obtención de un panorama de las necesidades y normativas, ha posibilitado conocer dónde se puede desarrollar 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, 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]

Al realizarse esta Tesis dentro de un determinado plazo temporal y estando este tema en constante desarrollo, estos estudios podrán ciertamente quedar escasos o incompletos como instrumento para su aplicación, gracias a las constantes mejoras de los estándares vigentes.

Por lo que solo se intentará definir el concepto aplicado y la descripción de las experiencias. Se realizará un prototipo del concepto que cumpla con las distintas posibilidades de operación y su prueba de desempeño estará restringida al diseño de un dispositivo conceptual.

2 DESARROLLO TEÓRICO

2.1 Antecedentes del tema

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. Una buena operación técnica debe hacer que los relieves sonoros sean articulados como un arte, no siempre bien aplicado.

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

Investigando como implementar un sistema de medición de audio utilizando una PC, se ha visto que la misma puede transformarse en una estación de trabajo de audio. Siempre y cuando posea la potencia de cómputo, el hardware y el software dedicados, lo que no es siempre posible realizar con computadoras para usos generales, ya que hay parámetros específicos que deben poseer para trabajar en audio profesional como las denominadas “Digital Audio Workstation”, según estándares de la Sociedad de Ingeniería en Audio (AES), Intel y Microsoft. [6, 8, 9, 12, 13]

De los documentos científicos publicados recientemente, se reconoce que varias organizaciones internacionales, como ser: la Unión Internacional de Telecomunicaciones (UIT), la Comisión Electrotécnica Internacional (IEC), la Organización Internacional de Estándares (ISO), el Instituto de Ingenieros en Electricidad y Electrónica (IEEE), la Sociedad de Ingenieros en Audio (AES), etc., 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 (SCT), el Instituto de Radiodifusión Sonora (IBS) de Inglaterra, la Unión Europea de Radiodifusión (EBU - EUR), la Asociación Europea para la Coordinación en la

Estandarización, (ANEC) - Bélgica, entre otros, 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.

A partir del material de base de las especificaciones se definieron los parámetros a medir, realizando el protocolo de mediciones utilizado durante las experiencias.

Aplicando también los trabajos relacionados a la especificaciones de calidad que deben cumplir los equipos PC en audio digital. [12, 13, 15, 16]

2.2 Medición del nivel de la señal de audio

Con el termino ***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*** (ó *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 deberá 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

Las vibraciones en el aire se propagan como una modulación de la presión atmosférica. El rango de variación en la presión sonora que el oído humano maneja va desde 0,0002 dinas/cm² (20μPa), hasta 200 dinas/cm² (20 Pa). Esto corresponde a un rango de un millón de veces. El que trasladado al terreno electrónico, por medio de un transductor, corresponderá a una señal de audio (medida en tensión o potencia eléctrica) que tendrá el mismo grado de niveles. Cabe mencionar que recién con la llegada de la tecnología digital fue 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*, nombre que se escogió 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 del Plata, Argentina. Aunque no es una unidad ISO nativa, su uso es aceptado. [17, 28]

El Bel acústico, entonces, es una unidad cuya referencia es el umbral promedio de audición humana (20μPa). Para una mejor representación de los valores, se utiliza una unidad 10 veces menor al Bel. Este se denomina decibel (dB), que minimiza el uso de números con decimales, y se define con la formula:

$$\text{Nivel de Presión Sonora [dB]} = 20 * \log\left(\frac{Pres_x}{Pres_{ref}}\right) \quad (1)$$

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 de referencia, el valor resultante será negativo. Estos no se utilizan en acústica ya que los valores negativos indican niveles no audibles.

Presión [Pa]	dB	Sensación	Ejemplos
20	120	Aturdimiento	Turbina jet, recital de rock, explosión.
2	100	Alto volumen	Discoteca bailable, música electrónica, despertador.
0,2	80	Sonido fuerte	Sonido en automóvil, audición 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.

Tabla 1: Rango de audición

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 lo muestra la Tabla 1. Donde se observa que el nivel de referencia de 0 dB, solo 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 todo caso a la disminución de la capacidad auditiva. [18]

Pero la sensación de sonoridad no depende solo del nivel absoluto de un sonido, sino de su conformación espectral, temporal y del entorno sonoro, incluyendo un gran componente psicológico. Debemos hablar entonces de una sensación psicoacústica, más que de un valor absoluto, cuando nos referimos al nivel de sonoridad (o volumen sonoro) de un sonido dado, además de nuestros oídos utilizamos nuestra psiquis. [10]

Para poder definir la sonoridad, este comprende un análisis más exhaustivo de nuestro sentido del oído, no solo del funcionamiento fisiológico del aparato auditivo sino de su interpretación en el cerebro.

2.2.2 Los 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*, mejor conocido como vúmetro. [3]

La potencia de 1mW disipada sobre una resistencia de 600 ohm requiere de un voltaje de 0,77459 V eficaces ó RMS (redondeado comúnmente a 0,775V). Por lo que esta potencia se definió como el valor de referencia eléctrico y en una escala graduada en decibeles correspondió a 0 dBm, (la tercera letra indica la referencia, en este caso 1 mW).

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 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]

La potencia eléctrica expresada en decibels se calcula con:

$$N [dB] = 10 \log (P / P_{ref}) \quad (2)$$

Donde: N = es la cantidad de decibels expresados en potencia

P = es la potencia medida

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

Esta fórmula puede ser expresada en términos de tensiones como:

$$N [dB] = 20 \log (V / V_{ref}) \quad (3)$$

Donde: N = es la cantidad de decibels en tensión

V = tensión medida

V_{ref} = Tensión de referencia (= 0,775 V para dBm ó dBu)

Asumiendo en este cálculo que la resistencia interviniente 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 medio de programa, medido con un *medidor de unidades de volumen* estandarizado (el vúmetro). [11, 20]

2.2.3 El vúmetro

Como fue expresado anteriormente, este se desarrolló inicialmente como un indicador de volumen estándar (*SVI* ó *Standard Volume Indicator*), para controlar y monitorear los niveles de audio. Aunque sus especificaciones reflejan la filosofía de los años 30, recién fue estandarizado en 1942, corregido en 1953, redefinido como estándar IEC N° 268-17/1990 y IEEE/ANSI N° 152/1991, y actualizado finalmente en 1999. Es esencialmente un galvanómetro pretendiendo medir el valor eficaz de la señal de audio, rectificándola y obteniendo el valor medio. [3,11, 20]

Su escala cubre de -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 1.

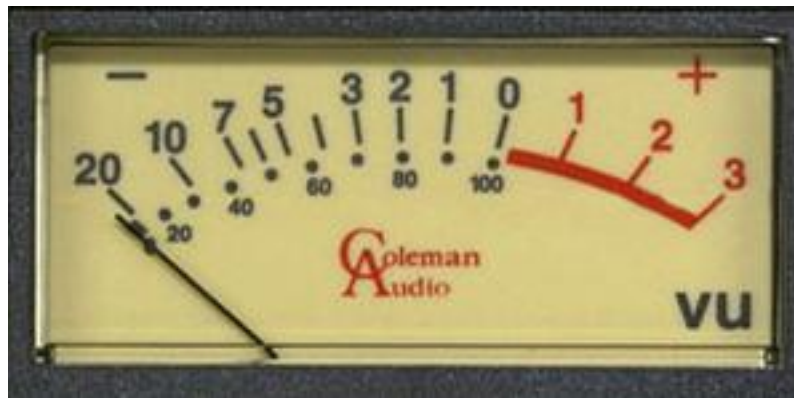


Figura 1: Escala de un clásico instrumento Vúmetro

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 segundos en estabilizar su deflexión, con un amortiguamiento crítico. [22, 26]

La señal que produce este efecto está comprendida entre 40 Hz y 10KHz, con un tiempo de crecimiento de 0,3 seg. 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]

Una señal instantánea de la voz o de la música puede tener en realidad más de 10 VU por encima del valor medido. Como resultado, los equipos de grabación y los de distribución en una cadena de audio necesitan tener un techo o margen superior muy elevados para evitar recortes (*clipping*) en los picos del material de programa, cuando se presentan ráfagas de alto volumen de corta duración. Por lo que un equipo que maneje +8 dBm como valor *SOL*, debe tener un techo mínimo de +18 dBm a una muy baja distorsión armónica total (THD) de menos del 1%. [2, 11]

Estas crestas de señal pueden pasar inadvertidas para un operador que está observando un vúmetro y las mismas llegarán al transmisor o grabador sin limitaciones, sobrecargándolo. La situación es aún más grave en transmisores de FM donde se utiliza un preénfasis de 75 μ seg resultando en una sobrecarga de +14 dB a 10 kHz. De manera que para evitar la sobremodulación se utilizan varios tipos de combinaciones de limitadores/compresores de nivel, llamados procesadores de audio. [2, 4, 5]

2.2.4 Niveles de referencia

Como se mencionó anteriormente, el nivel original fue el mW (0 dBm) y el nivel de salida para 0 VU de +4 dBm ó +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 0,775V 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 medio de programa, pero no garantiza la compatibilidad con otros equipos. Entonces, la mala interpretación de este valor lleva comúnmente a que los usuarios saturen la señal por exceso de amplitud, resultando en un recorte [23]. En la Figura 2, se observa en la pantalla de un osciloscopio una señal recortada en amplitud.

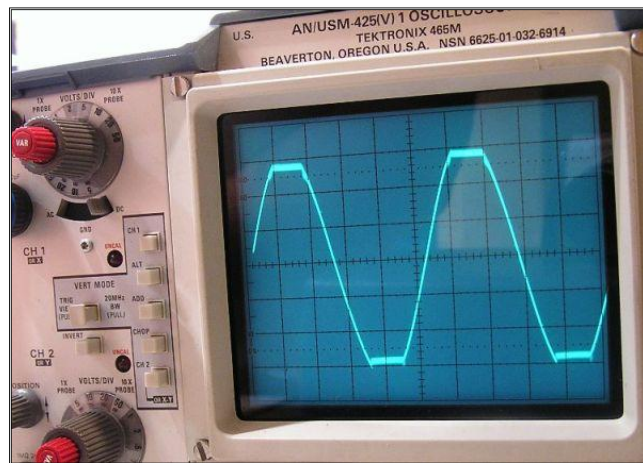


Figura 2: Señal recortada en amplitud

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 dependerá de los conversores A/D y D/A utilizados.

En el dominio digital este máximo valor debe ser “inalcanzable” por las crestas del material de programa, ya que tiene mayores consecuencias que en el dominio analógico.

Se presenta aquí 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 desde 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 18 dB 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.

Por lo tanto, en un equipo digital (como una computadora PC), como todos los valores medidos estarán por debajo del FS, se medirán en decibeles negativos.

El valor de referencia de 0 VU, fijado en Europa está en -18 dB del valor FS. (EBU R68).

La tendencia en los EEUU es definir este 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 aún mayor, ya que los +18 dB son referidos a una onda senoidal, y no al valor de 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]

En la Figura 3 se observa que al sobrepasar la señal los máximos valores posibles, el sistema recorta los picos agregando una cantidad de armónicos impares de alta frecuencia.

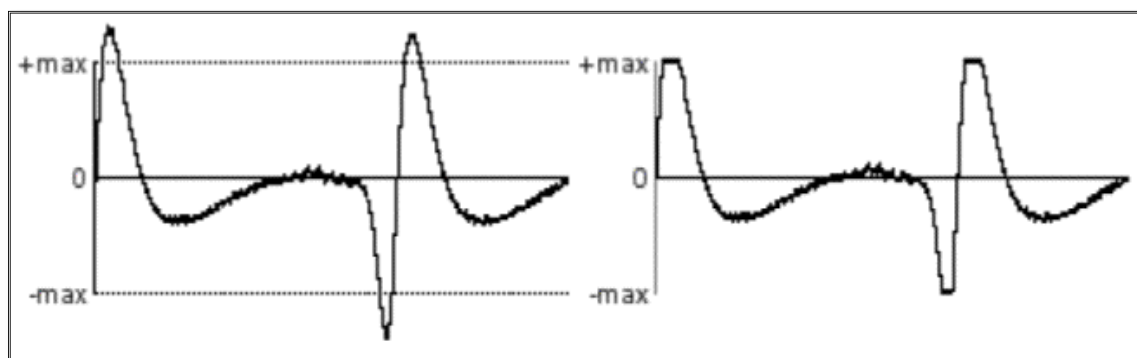


Figura 3: Recorte de los picos de señal por falta de margen

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

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

Este tema de la definición de niveles será tratado en el Capítulo 4, **La sonoridad en la Televisión Digital.**, en el apartado 4.2.1.

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

2.2.5 El medidor de crestas ó 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. Como se ha mencionado antes, 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.

Pueden 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 los 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 / limitación puede, en función de este valor, variar su ganancia y así evitar la saturación del equipo. [11, 25]

Esto ha permitido excitar al máximo los moduladores de AM y FM sin llegar a la sobremodulación. Pero un PPM no dice nada respecto a cuan fuerte puede sonar una porción específica de un programa. Y peor aún, muchos dispositivos traen medidores de crestas que son interpretados por los usuarios no expertos como vúmetros, llevando inexorablemente sus producciones a la saturación. [4, 9, 22]

En Europa se ha difundido el uso del *QPPM* (Quasi PPM) como un medidor de compromiso entre el Vúmetro y el PPM. El problema aquí, es que se usan diferentes estándares; por ejemplo en Inglaterra utilizan el estándar IEC 268-10 tipo IIa, en Alemania la escala DIN45406, en los países nórdicos el IEC 268-10 tipo I, en los EEUU se emplean indicadores tipo NBC. [11, 24]

Muchos fabricantes de equipos definen sus propios estándares y sus escalas están relacionadas a un producto específico, por lo tanto las escalas pueden variar. [24, 25]

Cuando se convierte una señal analógica a una digital, es esencial conocer como 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 R68, y la diferencia entre ellos es de 6 dB. [19, 22]

Como el factor de cresta en música puede superar fácilmente los 10 dB (3 veces), como fue mencionado. En los primeros conversores A/D si la señal analógica superaba la máxima escala del conversor, aparecían números binarios aleatorios. Hoy día los conversores pueden manejar esto manteniéndose, durante la sobrecarga, al máximo 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]

La manera tradicional de medir esta saturación es detectar los valores binarios de la señal. Si el valor muestreado coincide con el máximo binario estamos ante un valor peligroso y si esto se mantiene por dos muestras ó más indica una posible saturación (*clipping*). [20, 24]

Sin embargo, aunque se registren en las muestras valores por debajo del máximo, esto no garantiza que la señal no esté saturada cuando se convierta al dominio analógico.

Este problema ocurre porque los potenciales valores pico de la señal muestreada usualmente ocurren entre los espacios del muestreo, más que precisamente en el instante del mismo. Este tipo de medición de los valores máximos, nos da lecturas inconsistentes, manifestándose como sobrecargas no esperadas, sub-lecturas y posibles batidos en la medición de tonos puros.

2.2.6 El gálibo y la franquía de una señal de audio

En español, podríamos utilizar el término **gálibo** de origen estructural, y utilizado también en radiofrecuencia, que identifica a las cotas que señalan las dimensiones máximas permitidas, para el paso seguro por un túnel o debajo de un puente, ó 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 R128). 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 termino ‘*Peakroom*’ del inglés.

En la Figura 4 se muestra el Galibo y la Franquía de una señal de audio.

En el caso de equipos digitales, suponiendo la utilización de 16 bits en la cuantificación de las muestras, el valor FS correspondería a máximo valor pico binario 7FFFh, mientras que el valor de -20 dB estaría a un 10% o sea a 0CCCCh. Entonces el gálibo sería de 7333h. Y si esta señal tiene picos de +18 dB (65BCh), correspondería a una franquía de 0d77h.

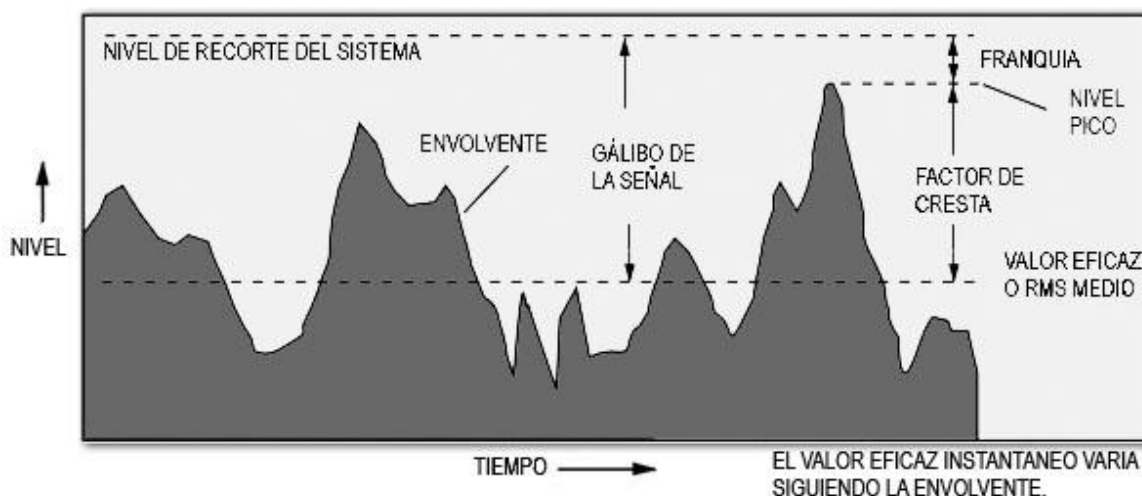


Figura 4: Galibo y Franquía de una señal de audio

2.3 La sonoridad.

Según la definición tomada del American National Standards Institute (American National Psychoacoustical Terminology, S3.20 de 1973), la sonoridad es una medida subjetiva de la intensidad del sonido percibido por el ser humano. Es decir, la sonoridad es el atributo que nos permite ordenar sonidos en una escala que va del más débil al más fuerte. [26]

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. Es más, seguro que esta variará en una cadena de audio.

Por lo que obtener una medición confiable se hace primordial para poder normalizar los niveles de una señal. El interrogante entonces es, si se puede medir la sonoridad 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 día gran parte o toda la cadena que va desde el artista ó 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. Y puede ser 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 mas volumen y señales mas procesadas reduciendo su rango dinámico, dentro de una cadena digital. Como consecuencia el oyente tiene cada vez menos control. Esto es lo que se llama “*Loudness War*”. [4, 22, 36, 38, 44]

Los productores de contenidos, especialmente de publicidad comercial, utilizan la técnica de sobre-comprimir los niveles para que sus publicidades suenen más fuertes que la competencia. Para lograr esto se utiliza un procesamiento en multi-frecuencias que modifica las componentes del sonido, dando como consecuencia un sonido “irreal”.

Cabe aclarar que el mundo de la radiodifusión analógica (AM y FM) tiene un párrafo aparte por su exiguo rango dinámico y el comportamiento de su cobertura. Aquí se utilizan procesadores para maximizar la modulación y por lo tanto la sonoridad. Esta nueva pauta de nivelación por sonoridad debería utilizarse en la etapa de creación y producción dejando que el operador y el procesador de salida del transmisor sea quien, en definitiva acomode la señal a las necesidades técnicas y estéticas de una emisora analógica.

Esto ha despertado un interés creciente en medir la sonoridad como parámetro útil para cuantificar los niveles de sonido percibidos.

En los trabajos de investigación se utiliza comúnmente el término “sonoridad percibida” ó “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 pleonismo, que en ocasiones, puede tener la función instructiva de reforzar la memoria de un tema. En este trabajo usaré el término sonoridad, sabiendo que su evaluación siempre será subjetiva por parte del oyente.

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 una cierta 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 requieren 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 diferencia. La capacidad de notar diferencias tonales depende también de la duración de los tonos, y es más o menos 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 resultan 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 exciten 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, o sea 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 5 se muestran las curvas isofónicas, que corresponden al nivel de presión sonora versus la frecuencia para un tono sinusoidal.

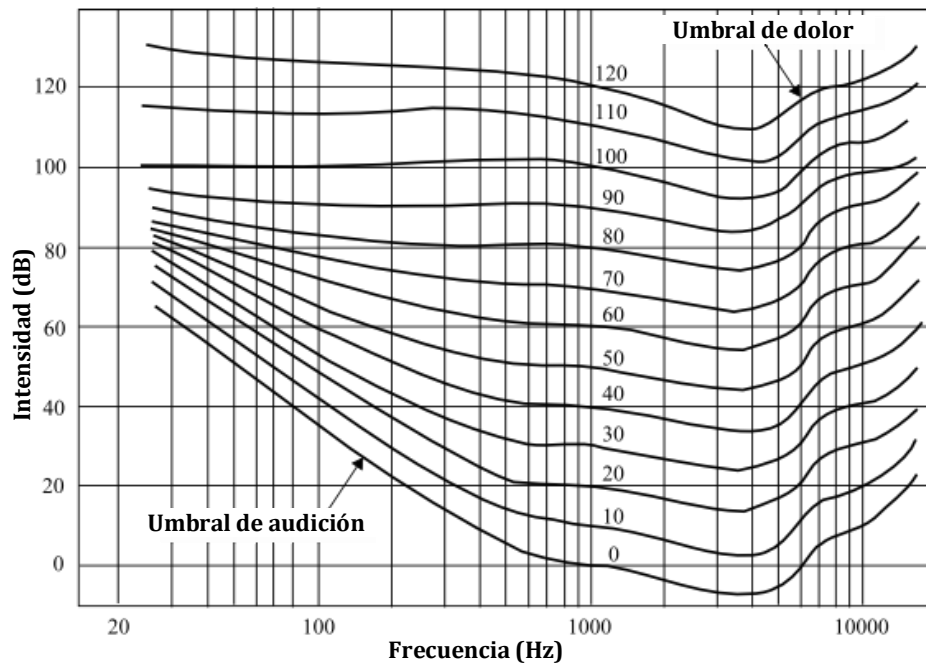


Figura 5: Curvas Isofónicas (ISO 226)

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_{SPL}.

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:

$$S = 10 * \log_{10} \left(\frac{I}{I_0} \right) \text{ Fones} \quad (4)$$

Donde: S = sonoridad en Fones

I = Intensidad medida a una frecuencia

I₀ = 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 la sonoridad de un determinado tono. Esto se debe a que la escala en Fones está directamente relacionada con la escala isofónica logarítmica, mostrada en la Figura 6.

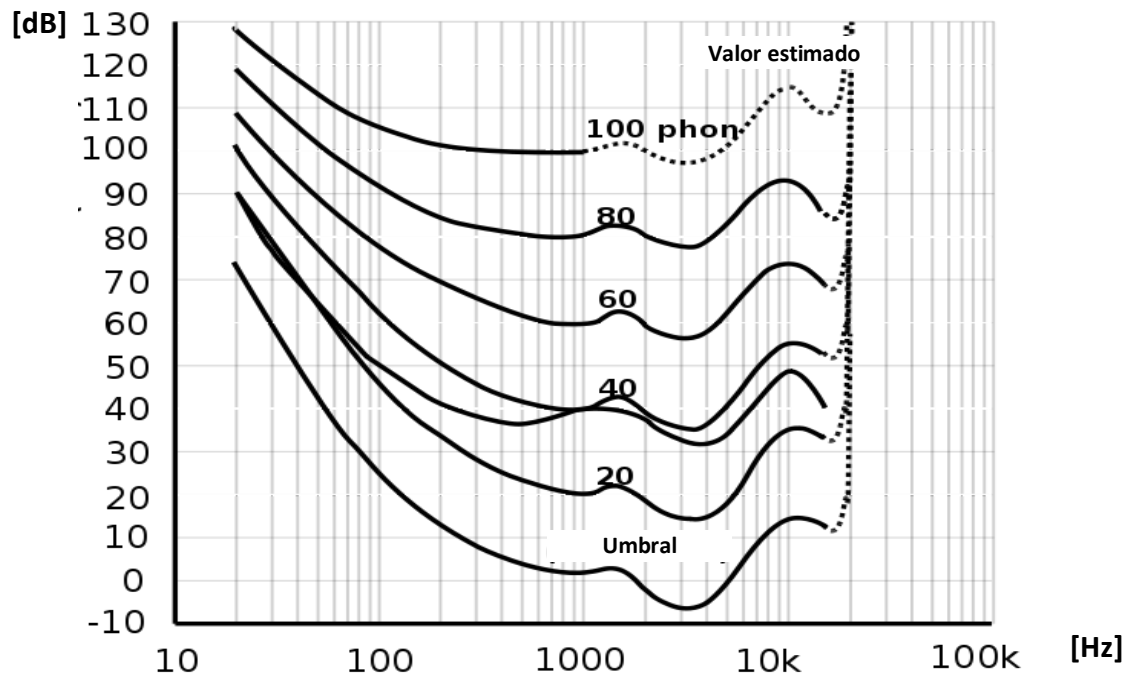


Figura 6: Intensidad expresada en Fones

La creación de estas curvas está basada en la respuesta subjetiva a tonos senoidales presentados frontalmente al oyente y solo 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.

A medida que aumenta la intensidad sonora, estas curvas se hacen cada vez más planas, mostrando que la dependencia a la frecuencia es menor a medida que aumenta el nivel de presión sonora. Por otro lado, si disminuye la intensidad sonora los primeros sonidos en desaparecer serían los graves, o sea las bajas frecuencias.

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).

La función de transferencia de estas ponderaciones se muestra en la tabla 2:

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

Tabla 2: Transferencia de las distintas ponderaciones

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 presión sonora se expresan en dBa. [1, 10, 18]

La ponderación B es la que corresponde a intensidades medias de sensibilidad del oído, por lo que es la que se toma como referencia para las mediciones de música. (dBb) [1, 10, 18]

La ponderación C modela la respuesta del oído ante sonidos de gran intensidad. Tiene una respuesta muy plana por lo que se utiliza para la medir equipos de sonido y para la evaluación de sonidos de baja frecuencia. (dBc) [1, 10, 18]

La Figura 7 muestra la curva de ponderación B utilizada en la sonoridad.

Ponderación B

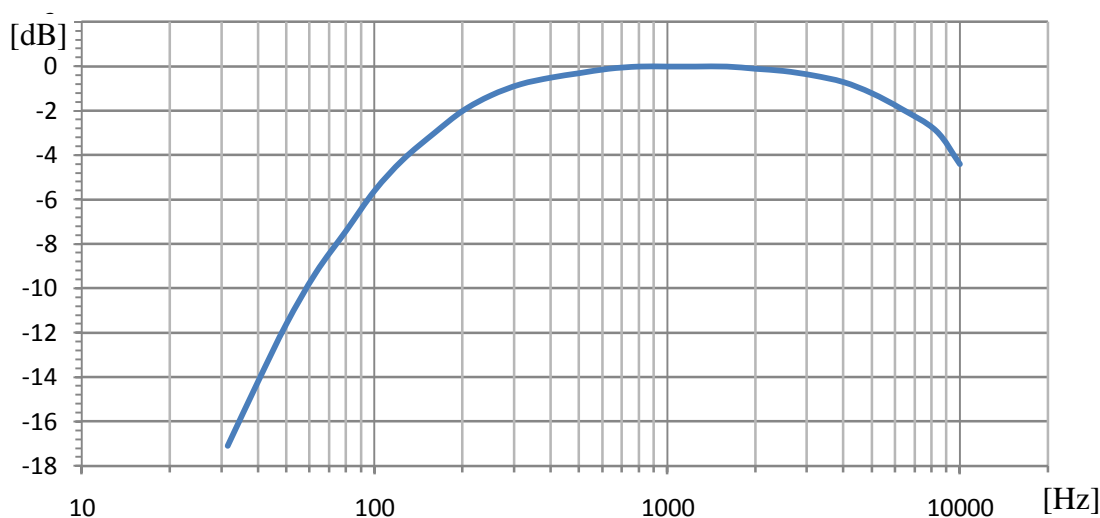


Figura 7: Curva de Ponderación B utilizada en la sonoridad

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. Chinn *et al.* se ve el primer diseño de un medidor de este tipo, que mide el valor medio 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 que los equipos electrónicos de audio tienen un máximo de señal manejable, independientemente de la frecuencia, todas las crestas deben ser medidas, aun 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 1 y en la Figura 4 el rango de audición normal del oído humano está entre los 20 dB $_{spl}$ y los 100 dB $_{spl}$, rango imposible de alcanzar por medios emisores analógicos, cuando se fijaron las normativas en radiodifusión en 1939.

Se logró tener un rango técnicamente optimizado de solo 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 de la Sonoridad

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 8, 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 está 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].

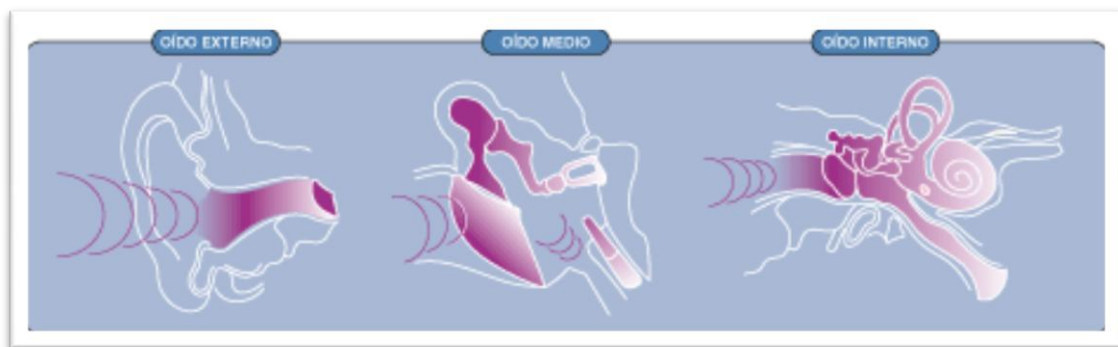


Figura 8: Partes del oído

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 solo. 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 en 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 primariamente 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 9.

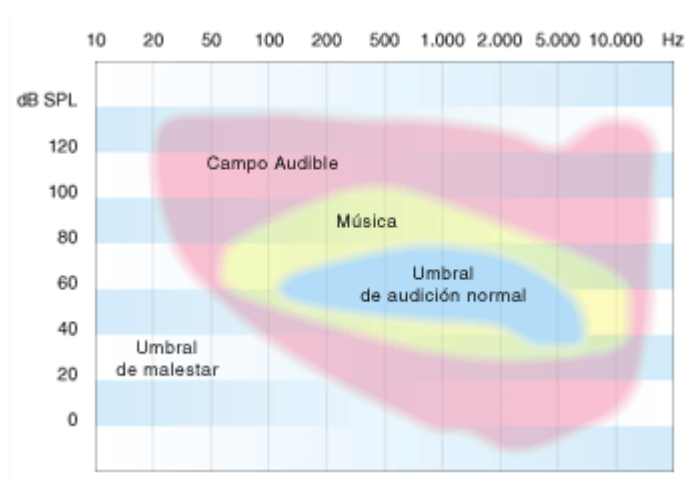


Figura 9: Rangos de sonoridad

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 como ser 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 radio, televisión o distribuido en CD ha sido dinámicamente y espectralmente procesado de manera de 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 la 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 un modelo perfecto que mida exactamente la sonoridad de un material de audio, sea este música, textos, o comerciales emitidos por los canales de radio y televisión. [10]

Por lo tanto, en conexión con las producciones de radiodifusión las aplicaciones de los medidores 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 definidos así 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, se han encontrado divergencias en las mediciones por lo que la mayoría de los estudios se realizan a largo-plazo. [10]

2.4.2.1 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, en la que las características de la función *potencia* se aplican a la magnitud física Intensidad (I) para obtener la magnitud Sensación (S).

$$S = K * I^n \quad (5)$$

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 algoritmos:

$$\text{Log } S = n \log I + \log K \quad (6)$$

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 decibels.

2.4.2.2 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 proporcionados 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 más conveniente, o precisa, para indicar diferencias de niveles de sonoridad.

2.4.2.3 Propiedades espectrales

Existe una estrecha correspondencia entre la frecuencia del sonido y su sonoridad. La fuerte dependencia de ésta función con la frecuencia es expresada comúnmente como *contorno de igual sonoridad*. (Estándar ISO: 226, 1997). (Fig. 4)

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 700Hz y de 100Hz por debajo de los 700Hz. [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. Se produce aquí una integración en el eje de la frecuencia, especificada 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 solo 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 decibels y fones con tonos puros las escalas coinciden, pero con ruido se desvían por causa de esta característica de la propagación. [30]

2.4.2.4 Propiedades temporales

Nuestra audición no reacciona instantáneamente al sonido, debido a que la sensación sonora requiere un tiempo. Dentro de ciertos 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.

2.4.2.5 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]

2.4.2.6 Señales de prueba

Para obtener una noción de los mecanismos psicoacústicos del sistema auditivo, se han empleado en las experiencias, señales de audio con ciertas propiedades analíticas.

Como estímulos comúnmente se utilizan ondas senoidales constantes o pulsantes de amplitud variable, 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.

La 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]

2.4.2.7 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 solo 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 ó *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 modelo de sonoridad solo se toma un tipo de entrada, el modelo se puede simplificar dramáticamente. Si, por ejemplo, el análisis se limita solo 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 cierto 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 descriptos completamente por su contenido espectral, tal como ruidos y tonos complejos (modelo de Zwinkel). 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 ó 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 solo 6 dB. El psicólogo John G. Neuhoff 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 notable que uno decreciente. Se ha sugerido que la escala en Sones manifiesta una polarización experimental, y por lo tanto no representa una relación fundamental entre estímulo y sensación sonora. [33]

2.5 El sistema de audio de la PC

El subsistema de audio de una computadora PC no está pensado para tareas de audio profesional, sino para las aplicaciones multimedia generales, que incluyen juegos, sonidos de alerta, sonido de videos, reproducción de música y algunas veces tareas de generación de audio. [8] Este subsistema ha evolucionado para llegar a ser complejo, con numerosos métodos para que las aplicaciones de software interactúen con las distintas fuentes de sonidos y destinos. Una aplicación puede solicitar al sistema operativo un archivo de audio, o poner directamente en memoria una onda de audio y solicitar al dispositivo de audio, por medio de su driver, que lo ejecute. [34]

Las aplicaciones pueden producir música usando un lenguaje abstracto llamado MIDI que soporta comandos de interpretación musicales. Y estos pueden generarse en un teclado de control externo conectado a la computadora. Además de estas fuentes hay otros dispositivos como ser lectores de CD-ROM, DVD y micrófonos que también producen audio, que puede ser oído y almacenado como un archivo de datos. [34]

Como respuesta a la demanda de mejores periféricos para manejar señales de audio, se han desarrollado soluciones de hardware específicas, como ser “placas de audio”. La compañía Creative por ejemplo, desarrolló el procesador de audio EMU10K1, y luego el EMU10K2, que posee un sintetizador de música de alta calidad, los conversores A/D y D/A, y un procesador de efectos de audio (DSP) en el mismo chip (detallado en el ítem 4.4). [ref. 34]

La performance de un equipo de procesamiento de audio está dada por el cumplimiento de especificaciones. Lo importante aquí es cómo estas están definidas de manera que puedan ser verificadas. Algunas mediciones son idénticas a las que se realizan en los equipos de audio analógicos mientras que otras solo involucran a equipos digitales. [6, 12, 15, 16]

La calidad de audio esperable en una PC para formar un Digital Audio Workstation – DAW- esta especificada en documentos de información técnica de manera de estandarizar los parámetros y permitir la interconexión y la comparación entre especificaciones de distintos proveedores, homogeneizando el mercado del audio en PC. [9]

Con respecto a sus aplicaciones en el audio profesional, en su trabajo, *Sutcliffe y Yonge* presentan una temprana visión de las ventajas y desventajas en la utilización de estaciones de trabajo para audio (DAW) en PC, mostrando que la trascendencia que tienen las computadoras en unas áreas no reemplaza a los equipos de hardware dedicado en otras. [8] Entonces en el audio profesional el equipamiento puede ser dividido en dos tipos, consolas de mezcla y procesamiento de audio en tiempo real, y sistemas de post-producción de audio digital.

Estos dispositivos se utilizan en estudios de grabación y en radiodifusión, en un gran rango de aplicaciones. En estas, el operador usa una consola para controlar simultáneamente las fuentes de señal, la mezcla, el procesamiento de efectos y el ruteo de señales. Para trabajos en post-producción los sistemas de edición no lineal son utilizados principalmente en audio y video. [8]

Algunas de las ventajas de usar computadoras PC en la producción de audio se pueden encontrar en los desarrollos de software específico, como los producidos por las empresas Adobe, Audio Design o Steinberg, entre otras. Estos ofrecen una interface familiar dentro de un sistema operativo como Windows ó Mac OS. [8]

Sobre las ventajas que poseen estos sistemas vale la pena enumerar algunas desventajas:

Las soluciones dedicadas siempre tienen ventaja sobre las de uso general. Esto es para cumplir con los requisitos necesarios para la mezcla y control de un gran número de fuentes de audio de calidad profesional.

Cuando se intenta implementar un DAW con estaciones de trabajo basadas en PC estándar orientadas al entorno de negocios (procesamiento de textos), se presentan graves inconvenientes. Una estación de trabajo de este tipo raramente tiene la gran capacidad necesaria para el procesamiento de audio digital multicanal. [8]

Los requerimientos de producción en radiodifusión son distintos a los de post- producción. Los principales requerimientos de las consolas procesadoras de audio para difusión pueden ser resumidos como sigue: manejo de muchas entradas, operación libre de fallas y controles de fácil acceso. Para post-producción de audio los principales requerimientos son, velocidad, capacidad de control remoto, gran capacidad de almacenamiento y una operación multiusuario y multicanal. [8]

La segunda razón de por qué las computadoras no han tomado el lugar de las consolas de mezcla esta atribuida a la ergonomía y a la baja resolución de las pantallas de video.

Es más fácil para el operador mover un potenciómetro de una consola para variar un nivel que efectuar esa función moviendo el mouse. Representar los cientos de controles en una pantalla de computadora no es práctico. Y aunque se muestren los paneles por sección, en diferentes páginas, esto impide el acceso y no es aceptable para la operación en vivo. Aún en la operación *off-line* los operadores tienden a desarrollar lo que se llama relación *mente-mano* con su consola. Ellos ajustan los controles sin mirarlos, lo que no los distrae de su concentración en la música. En cambio, usar un mouse o el teclado para acceder a las funciones y controlar los parámetros lleva a una falta de concentración por parte del operador. [8]

La resolución de los displays está limitada por la tecnología de los monitores y la disponibilidad de grandes modos de resolución. Actualmente, se llega en el formato experimental WHUXGA, a 7680×4800 píxeles por líneas activas (16:10), en el mercado se consiguen pantallas de hasta 2560×1600 (16:9), en el formato WQXGA. [35]

Para sistemas multiusuario y múltiples señales, el procesamiento requerido aumenta dramáticamente sobre lo exigido para la operación de un solo usuario. Una estación de radiodifusión o una facilidad requieren compartir valiosos recursos entre los usuarios. Esto requiere una central de procesamiento que almacene y distribuya audio a múltiples usuarios. Algunas aplicaciones multimedia permiten compartir archivos de audio pasándolos a través de redes convencionales. Pero estas tienen un ancho de banda muy restringido para operaciones de trabajo pesado donde múltiples usuarios requieren múltiples canales de audio de calidad profesional bajo demanda y en tiempo real. [8]

Se puede argumentar que el crecimiento en el procesamiento cliente/servidor en la industria de las computadoras apunta en este sentido, ya que se puede usar la topología de red Gigabit Ethernet, optimizada para las necesidades del audio profesional. [8]

Un sistema que pueda combinar hasta 2000 fuentes y almacenar más de 72 horas de audio en 128 pistas, no se puede implementar con una sola computadora de tecnología tradicional, sino que se requiere un servidor especializado. Una computadora DAW solo puede usarse como interface hacia el núcleo para realizar trabajos intensos. [8]

2.5.1 Situación Actual del audio en PC

Al analizar las características que debe cumplir una estación de trabajo en audio digital, se observa que los estándares están orientados a la obtención de la máxima calidad en las señales digitales y cubrir las necesidades de los usuarios. [6]

Esto permite reemplazar el equipamiento tradicional de laboratorio de audio, con una PC de alta gama, dedicada a las mediciones de audio. [6, 12, 15]

Por lo que se debe desarrollar una metodología de implementación, y determinar la confiabilidad de las mediciones realizadas en una PC, compatibilizando las mediciones de señales digitales con las analógicas. [9]

La especificación de Audio de Alta Definición de Intel “AZALIA”, recientemente publicada, describe una arquitectura e infraestructura para soportar implementaciones de audio de alta calidad en PCs. La especificación define la interface a nivel de registro, las características de los enlaces físicos, el modelo de programación y los componentes arquitectónicos de los Codecs. Esta especificación está pensada para los diseñadores de componentes de hardware, constructores de sistemas y desarrolladores de dispositivos de manejo de software (Drivers). [7]

La especificación del audio de una computadora PC está actualmente determinada por la normativa PC2001. [7, 9, 13]

En aplicaciones profesionales, donde son necesarios varios canales de entrada y salida, se hace imperativo el acceso directo a memoria que permite la rápida transferencia en tiempo real. Las limitaciones del audio digital son básicamente dos; una es el manejo de la resolución de la señal (bits por muestra) y la otra es la cantidad de datos (bytes de almacenamiento). [6, 9]

El Bus PCI parece proveer un ancho de banda adecuado para el audio, pero comienza a ser un cuello de botella. [34] Para ello se ha implementado el BUS PCI-E, PCI-Express de tercera generación, como una evolución del bus PCI que usa los conceptos de programación y estándares de comunicación ya existentes, pero se basa en un sistema de comunicación serie mucho más rápido. Este sistema es apoyado principalmente por Intel. En el PCI-E 1.1 (2005) cada enlace transporta 250 MB/s en cada dirección (4GB/s en x16). PCI-E 2.0 (2008) duplica esta tasa a 500MB/s (8GB/s en x16) y el PCI-E 3.0 (2011) la vuelve a duplicar a 1GB/s (16GB/s en x16).

2.6 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 necesidad de potencia computacional es “simplificar” las señales a procesar, y 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 mantiene 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 ó 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 estimación 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 mucho la conducta del Compresor o Encoder. Los diseños actuales están incluyendo una transición suave entre los segmentos, conocido como “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 10.

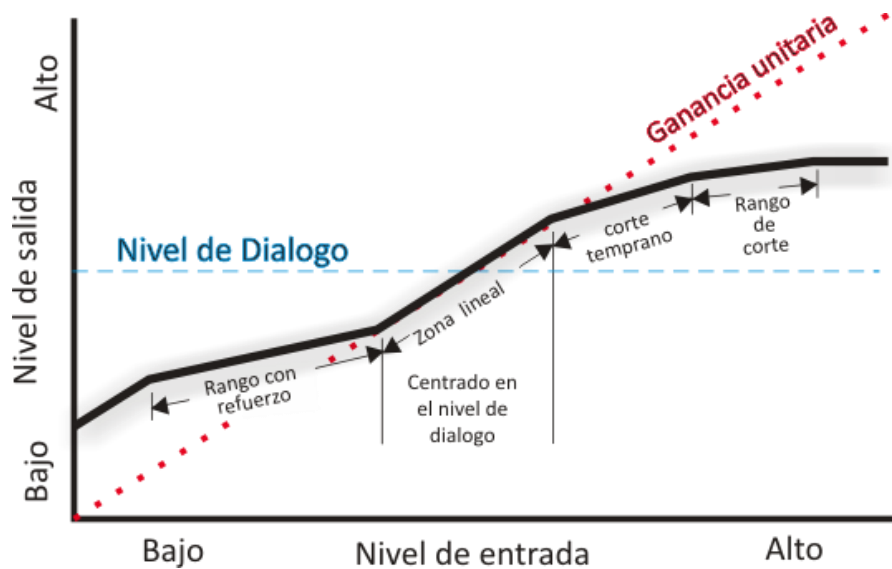


Figura 10: Curva de Compresión Dinámica

Los Procesadores de Sonido Digital ó 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 la 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 es trabajar con una estructura de 28 canales, con filtros de fase lineal con ceros dentro del círculo unitario en la respuesta. Para derivar la señal de control temporal se introducen 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 del 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 10, 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 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 11 se aprecian los distintos rangos dinámicos en función de las distintas condiciones de audición.

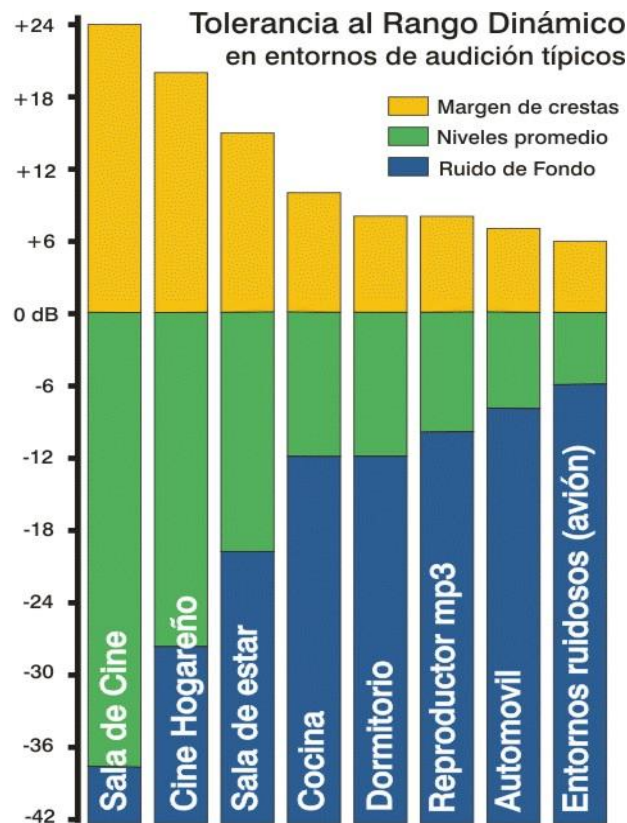


Figura 11: Rangos dinámicos típicos

2.6.1 El sobre-rango o saturación digital

El concepto de “over” (over-range ó 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 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 75dB y la relación señal-ruido (S/N) está en el orden de 96 dB.

Cuando se produce o graba audio en el mundo analógico, se incluye un cierto 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.

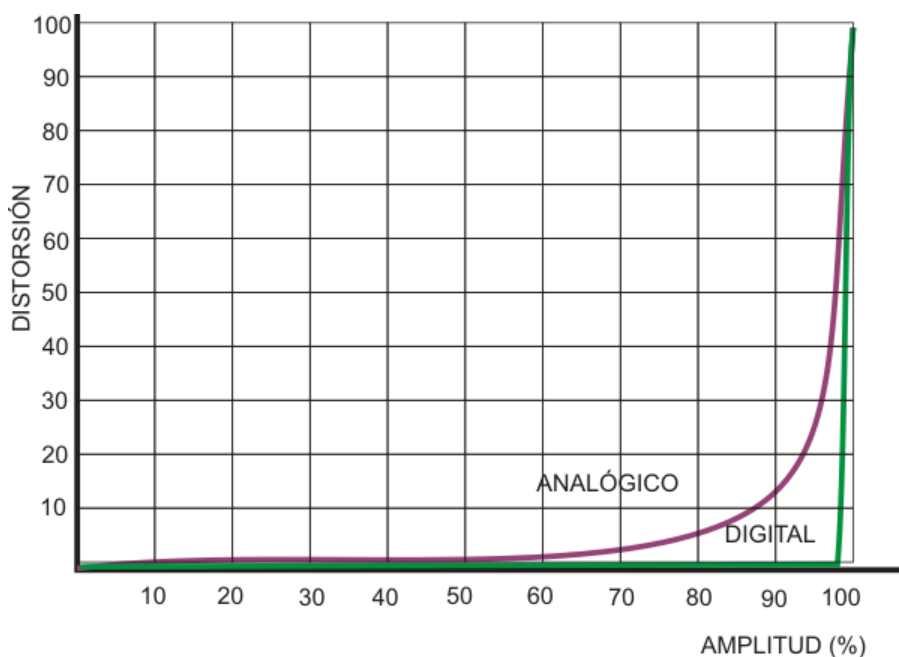


Figura 12: Comportamiento aproximado de la distorsión

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 esta 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 12 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, solo 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, como responderá un dispositivo a una sobrecarga depende del diseño y la tecnología del mismo. Algunos lo harán suavemente y otros podrán distorsionar severamente al recortar la señal.

En el audio digital, y esto es muy importante, 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). Y 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 ocasional 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 y 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]

2.6.2 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 este gran rango dinámico. Los grandes bafles de los ochenta se reemplazaron por pequeños bafles exigiendo mayores potencias 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 128kbps suena casi como un casete. Los reproductores de música digital con este formato, tienen la gran ventaja de que ya no poseen partes móviles, la música está almacenada digitalmente en una memoria flash, y se transfiere directamente desde un conector USB, sin copia degenerativa alguna.

Estos equipos evolucionaron hasta ser hoy día incluidos en los teléfonos y demás dispositivos portátiles. Pero la calidad de escucha es muy mala. Por un lado el ruido 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].

3. HIPÓTESIS DE TRABAJO

Mejorar la medición de la sonoridad

La exigencia de los oyentes de Televisión Digital es no tener que estar constantemente ajustando el volumen de sus receptores. Esto exige mantener un volumen constante en todo el material emitido. Para cumplir con esto es necesario poder medir y evaluar la sonoridad del material antes que éste sea transmitido. Entonces, si se puede disponer de un sistema de medición objetivo y de una técnica de aplicación correcta, se podrá corregir este defecto.

La merma del rango dinámico en las grabaciones contemporáneas en busca de una mayor atención del oyente se ha visto que es perjudicial para la salud. Por lo que se propone presentar un cambio paradigmático en la manera de medir los niveles de señal y no nivelar el material de audio digital por los picos, sino por su sonoridad.

Es muy necesaria la medición de los valores de amplitud de las señales de audio procesadas por medios informáticos para evitar la saturación binaria. Y como se ha comprobado, la medición según las técnicas utilizadas en el mundo analógico no es representativa de los niveles de sonoridad.

Primeramente se propone el estudio de un algoritmo que represente fielmente esta característica del oído, partiendo de la solución pautada en la recomendación ITU BS.1770. Se diseñará y construirá, bajo este concepto un medidor que cumpla con estos parámetros. Proponiendo que el nivel de referencia este a -23 dB por debajo del máximo.

Para comprobar su desempeño se realizarán una serie de mediciones de comparación. Aplicando técnicas estadísticas para determinar su desempeño y exactitud.

Obteniendo los valores más apropiados para la realización de un algoritmo de medición confiable.

Se podrá aplicar este trabajo como guía a una propuesta de recomendación de normativa en el ámbito nacional.

4. PROPUESTA DE SOLUCIÓN

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 que intentan presentar una solución al problema planteado. El mismo permite conocer fundamentalmente los modelos para implementar el sistema de medición y define los parámetros a considerar.

4.1 Aportes de las distintas Normativas

4.1.1 Recomendación ITU BS.775 [37]

Las importantes limitaciones inherentes a los sistemas de sonido multicanal domésticos y la necesidad de una presentación optimada, 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 domestica, 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 que se recomienda la adopción de un sistema multicanal universal, mostrado en la Figura 13. 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 se muestran los niveles de matrización, según los canales utilizados.

Formato	L	R	C	LS	RS
Estéreo L R	1,000 0,000	0,000 1,000	0,707 0,707	0,707 0,000	0,000 0,707
3/2	1,000	1,000	0,5	0,707	0,707
2/2	1,0 + 0,7 C	1,0 + 0,7 C	0,000	0,707	0,707

Tabla 3: Ecuaciones de matrización de codificación y decodificación

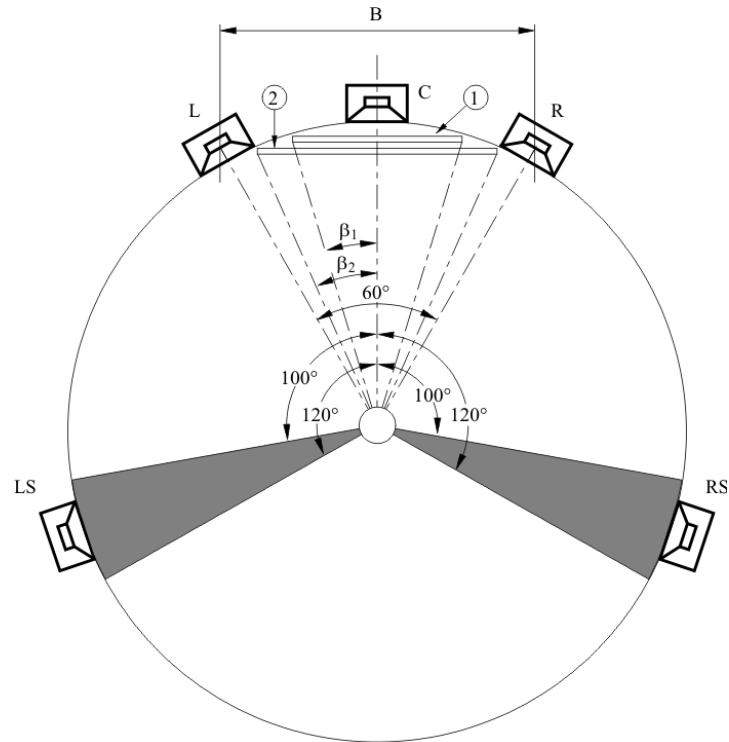


Figura 13: Disposición de los altavoces L, R, C y LS, RS

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

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

H = Altura de la pantalla

B = Distancia de los Altoparlantes

En la Tabla 4 se muestran las posiciones de los parlantes en un sistema multicanal

Altoparlante	Angulo Horizontal	Altura (m)	Inclinación (grados)
<i>C</i>	0	1,2	0
<i>L, R</i>	30	1,2	0
<i>LS, LR</i>	100 ... 120	$\geq 1,2$	0 ... 15 hacia abajo

Tabla 4: Posición de los parlantes multicanal según BS.775

4.1.2 Recomendación UIT BS.1770-1 [38]

Algoritmo de la sonoridad de los programas de audio y niveles de pico real.

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 el valor pico y el nivel de recorte. Se recomienda el uso del algoritmo desarrollado para permitir tanto la ecualización 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 14) indica el algoritmo de medición:

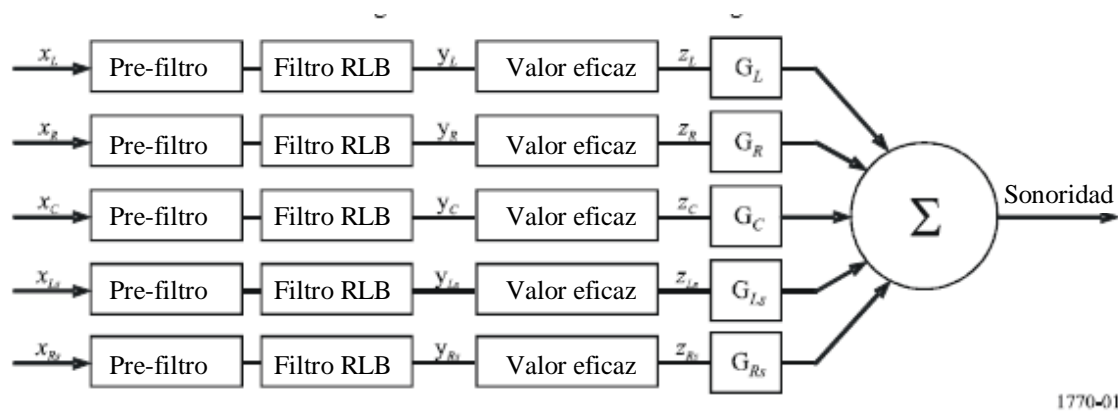


Figura 14: Algoritmo de sonoridad

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:

		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:

		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 formula:

$$z_i = \frac{1}{T} \int_0^T y_i^2 dt \quad (9)$$

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

Luego se suman los canales, dando una fórmula:

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

Los coeficientes a aplicar a cada canal se muestran en la Tabla 5:

Canal	Izquierdo (G_L)	Derecho (G_R)	Central (G_C)	Posterior Derecho (G_{RS})	Posterior izquierdo (G_{LS})
Ponderación, G_i	1.0 (0 dB)	1.0 (0 dB)	1.0 (0 dB)	1,41 (~1,5dB)	1,41 (~ 1,5dB)

Tabla 5: Coeficientes de ganancia de los canales

4.1.3 Recomendación UIT BR.1771-2 [39]

Requisitos de los medidores de sonoridad y de cresta real.

Los vúmetros y los medidores convencionales de nivel de cresta de programa no proporcionan 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.1770 especifica un algoritmo de sonoridad y de niveles de cresta reales que cumple con los requisitos de medición en 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 6.

Unidad de sonoridad (LU):	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.
Presentación visual de Tipo I:	Presentación visual electrónica con resolución de uno o más segmentos para 3 unidades de sonoridad.
Presentación visual de Tipo II:	Presentación visual electrónica con resolución de uno o dos segmentos por cada unidad de sonoridad.

Tabla 6: Características de los medidores de sonoridad

Como ejemplo se puede decir que una señal de referencia es una onda sinusoidal continua con un nivel de presión sonora (SPL) 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 objeto 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_{SPL} 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), y se utilizan para distintos objetivos:

- **El modo rápido** se utiliza en la producción, posproducción 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 la 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 a 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_{SPL}.

El modelo perceptivo para la sonoridad es una función no lineal de la amplitud, la frecuencia y al 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 las 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 el medidor pueda indicar, cuantos decibeles (en ganancia ó 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 ó pérdida en decibeles que debería aplicarse a una señal para llevarla a la sonoridad de referencia. Las unidades de sonoridad tienen 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.

4.1.4 Documento A/85 del ATSC [40]

Técnicas para Establecer y Mantener la Sonoridad del Audio en la Televisión Digital

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 la mejor calidad en las 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 post-producción y en los métodos que permitan un control efectivo de la sonoridad en la distribución e intercambio de contenidos. Adicionalmente, recomienda métodos para un control de la 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 se refiere al nivel del dialogo, ó en su defecto al nivel medio del programa, que debe configurarse correctamente, para prever 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 dialogo, consta de 5 bits sin signo, indicando cuan debajo está de 0 LKFS. Los valores validos van desde 1 a 31, siendo el “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 dBspl, según el tamaño del recinto de control, para una señal de -24 LKFS.

Define además que el valor 0 dBFS es para una onda senoidal, según lo establecido en AES17-1998/R2004.

4.1.5 Documento R 17 / 19 de la CST [41, 42]

La CST (Commission Supérieure 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 -83 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 1KHz, dando -9 dB en PPM DIN 45406 y 0 VU en un vúmetro.

El máximo valor pico debe no exceder 0 dB *PPM*, DIN 45406, y -3 dBFS.

La dinámica según la norma RT-19 debe estar centrada en el nivel de dialogo. Se definen las mediciones en unidades LKFS, que es el valor de sonoridad (el nivel) expresado como una medición ponderada en K (*Leq* (R2LB)) a plena escala.

La escala LKFS esta graduada en pasos de 1 dB. El nivel de dialogo puede tener un piso 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 +/- 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 dialogo, valor instantáneo.
- +/- 12 dB para todo sonido, valor instantáneo.

Para medir la dinámica utilizar una herramienta que permita ver el nivel de energía durante 10 segundos en una ventana deslizable en LKFS (ITU BS.1770). Para cumplir con 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).

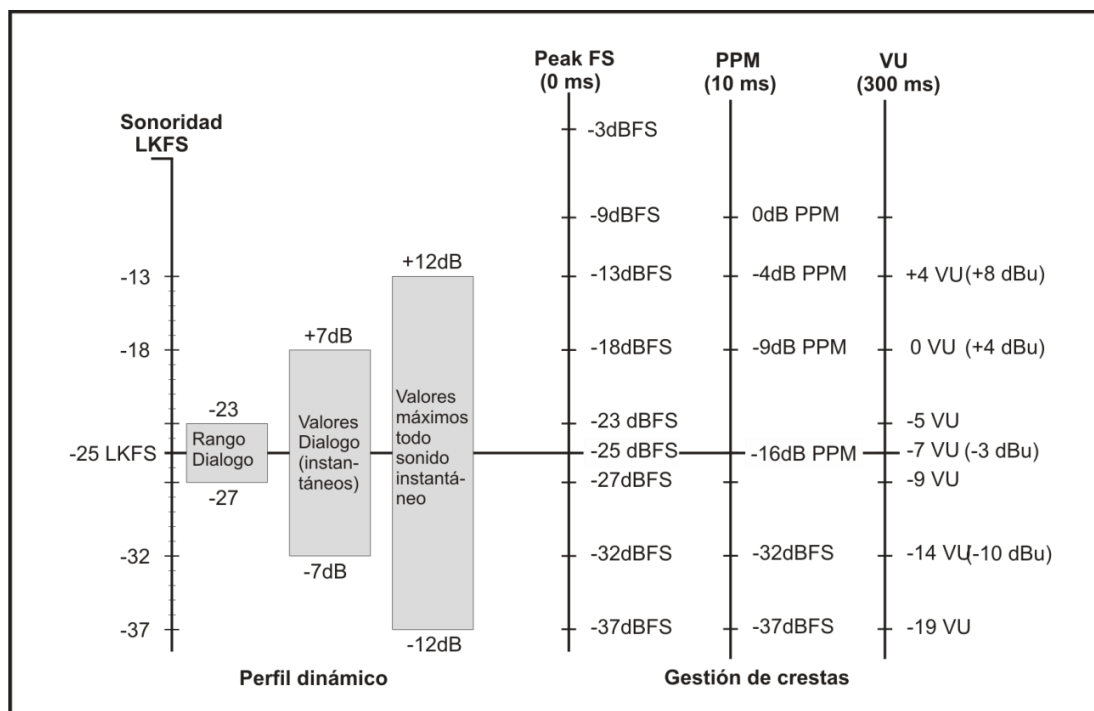


Figura 15: Escalas de los distintos niveles según la norma RT-17

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 conversores entre 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 matizado 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 dialogo (*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.

4.1.6 Documento ARIB TR-B29 [43]

El estándar de Televisión Digital Terrestre adoptado por nuestro país es el ISDB-T (Integrated System Digital Broadcasting – Terrestrial) desarrollado por 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:

Sintaxis	Número de bits	Mnemónico
Audio parameters () {		
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, bit mas 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 diagnostico 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:

$$Aii = INT \left[\frac{1}{8} \left(\frac{1}{2N} \sum_{i=0}^{N-1} abs(X(i) + Y(i)) \right) \right] \quad (12)$$

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 los 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 de audio fuera de fase entre dos canales AES, según lo calculado por:

$$Aoi = INT \left[\frac{1}{8} \left(\frac{1}{2N} \sum_{i=0}^{N-1} abs(X(i) - Y(i)) \right) \right] \quad (13)$$

Descripción de las variables, igual que arriba.

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 formula 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:

$$AMI = INT \left[\frac{1}{8} \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} X(i)^2} \right] \quad (14)$$

Descripción de las variables, igual que arriba.

Se aplica un pre-procesado a la señal de audio con un filtro de 20 Hz, mostrado en las Figuras 16 y 17, estructurado como dos filtros de 2º orden en cascada, con los siguientes coeficientes:

		$b0$	0,9981318
$a1$	-1,9962602	$b1$	-1,9962636
$a2$	0,996267	$b2$	0,9981318

Tabla 7: Coeficientes del pre-filtro de entrada

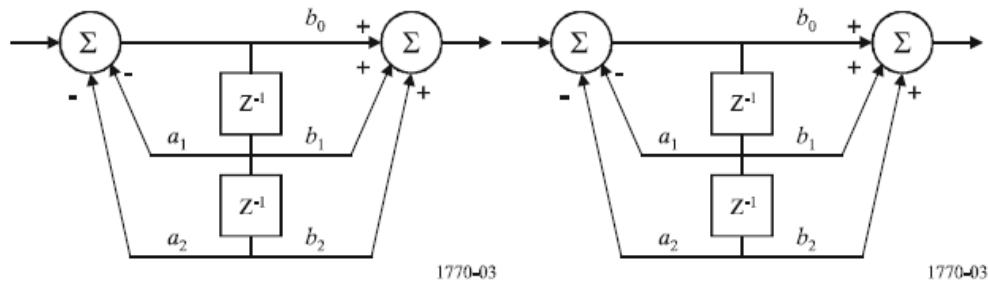


Figura 16: Diagrama de flujo de la señal como un filtro de 4º orden

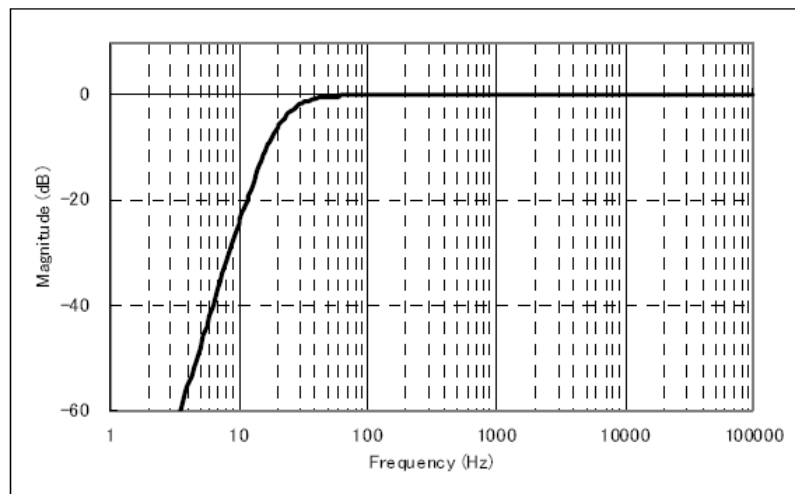


Figura 17: Respuesta en frecuencias del pre-filtro

Indicador de calidad	Nombre del modo	Criterio de calidad	detalles
00	reservado		
01	Modo 1	Modo estándar de televisión terrestre	Bit-rate 192 a 256 kbps / estéreo o mas (2)
10	Modo 2	distinto al 1, y no calificado como 3	Otro modo no especificado
11	Modo 3	audio con calidad limitada < 192kbps	Los teledifusores elegirán este modo a su criterio.

Tabla 8: Indicador de las asignaciones de calidad del audio

La Tabla 8 muestra que el Modo 1 es el mejor y es el estándar, ya que cuanto mayor es el bit-rate del flujo, mayor es 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 brasileiro 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.

4.1.7 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 medio de la sonoridad de un programa (*'Programme Loudness'* ó *'Sonoridad de Programa'*), los descriptores *'Loudness Range'* (*'Rango de Sonoridad'*) y *'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 Tech Doc 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'* se utilicen para caracterizar una señal de audio.

Que el *'Programme Loudness Level'* sea normalizado a un *'Target Level'* (nivel índice) de -23 LUFS, con una desviación máxima permitida de ± 1 LU.

La medición sea hará con un medidor de sonoridad en conformidad tanto al documento ITU-R BS.1770, como al EBU Tech Doc 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 Doc 3341,

El '*Loudness Range*' se medirá con un medidor conforme al documento EBU Tech Doc 3342 [5]. El '*Maximum Permitted True Peak Level*', durante la producción de un programa, será de -1dBTP (dB True Peak), medido con un medidor conforme al documento ITU-R BS.1770 y al EBU Tech Doc 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 Tech Doc 3343 [6],

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

Y este documento define que:

Un **Programa** es un elemento audiovisual o solo de audio autocontenido e individual, para ser presentado en radio, televisión u otro medio electrónico. Un anuncio comercial, un tráiler, un elemento promocional ('promo'), 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 del programa.

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

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 Tech Doc 3341 [4], recién habían comenzado a estar disponibles.

4.2 La sonoridad en el audio digital

4.2.1 Norma Brasileña ABNT NBR 15602-2 [44]

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 ó 5.1) y su necesaria homogeneización 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 (p ej. 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 uPa) y referido a intensidad sonora serían unos 10^{-16} W/cm².

- 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 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 estén 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, solo lo 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 por lo menos un factor de cresta de 6 dB en potencia, lo que es el doble en tensión (+6 dB). Este criterio, se entiende, limitaría la posibilidad de sobrecomprimir 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

Indicador 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. 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 tornar 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

Dynamic Range 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

4.2.2 El caso argentino

En la Argentina, al elegir el estándar ISDB-T para la transmisión de la Televisión Digital Terrestre no se han definido aún 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.

4.2.3 Análisis de otras propuestas de medición de la sonoridad

Una cantidad de mediciones de la sonoridad a largo plazo, se han desarrollado para diversos propósitos, como ser medidores de nivel, de control, y para estimación y predicción de la percepción de la sonoridad. [10]

Ya en el año 1997 Ioan Allen expuso en la conferencia de la SMPTE que la métrica del *Leq* (Loudness equivalent), que representa la integración de la energía a largo plazo, midiendo la potencia equivalente sobre el tiempo, puede ser usada como un estimado de la sonoridad. En el mencionado artículo denominado “Are movies too loud” Allen concluye que para estimar la sonoridad subjetiva a largo plazo, el *Leq* se comporta mejor que la medición de picos con una ponderación C propuesta.

$$Leq_{(dB)} = 10 \log_{10} \left[\frac{1}{T} \int_0^T \left[\frac{P_w(t)}{P_{Ref}} \right]^2 dt \right] \quad (15)$$

Ecuación de cálculo del *Leq*

También se ha sugerido el uso de ponderaciones de frecuencia alternativas. Earl Vickers propuso en la AES 117 del 2001 usar un corte de 200 Hz de primer orden como aproximación al corte de baja frecuencia de la curva de ponderación B. Esta aproximación se justifica por el bajo costo computacional y por el hecho de que las señales musicales típicamente tienen poca energía a altas frecuencias. La frecuencia de corte se eligió para evitar la modulación en baja frecuencia y pre-énfasis en compresores de banda única. [51]

Soldroude y Norcross investigaron las mediciones objetivas de la sonoridad, incluyendo la propuesta anterior a la que se refirieron como *Leq*(Bhp) (por Butterworth high-pass), y realizaron otras mediciones utilizando las curvas de ponderación A, B, y C, la ponderación de Terhardt basada en el umbral de audición absoluto (referido como *Leq*(ATH)) y la nueva ponderación *Leq*(RLB) que es una revisión de la ponderación B. Sus estudios

mostraron que una atenuación en altas frecuencias degrada la performance de la estimación de sonoridad. Soldroude y Norcross encontraron que la ponderación *Leq*(RLB) fue la mejor de las evaluadas. [52]

La ponderación *Leq*(RLB) es bastante plana, corresponde a una medición RMS con una caída de -1 dB en 100 Hz y -3 dB a 60 Hz. Se encuentra entre las ponderaciones B y C, de donde se desprende que es apropiada para mediciones de alto nivel (70 a 100 dB).

A pesar de que las ponderaciones A, B y C invertidas se aproximan pobremente al contorno de igual sonoridad, sorprende cuanto se utilizan en las pruebas de audición.

Stevens y Zwicker desarrollaron un modelo de sonoridad multibanda que fue incorporado en los estándares ISO 532-A y 532-B. [26]

Moore y Glasberg propusieron mejoras adicionales incluyendo un método de cálculo de la sonoridad con sonidos fluctuantes. [53]

Seefeldt *et al*, desarrollaron una nueva medición objetiva de la sonoridad basada en modificaciones a los modelos psicoacústicos previos y encontraron que se desempeñan significativamente mejor que otros métodos probados con ejemplos tomados de una base de datos de audio. [53]

En el 2004 Skovenborg y Nielsen escribieron un extenso trabajo donde evaluaron doce modelos de percepción de la sonoridad a largo plazo, incluyendo varios modelos *Leq*, dos variantes del método Zwicker y dos modelos propuestos por TC electronic. Para el estudio utilizaron un nivel de referencia centrado en 70 dB SPL, como un compromiso entre el volumen de un televisor y el cine. Encontraron que los modelos TC dieron la mejor performance seguidos por el *Leq*(RLB). [10]

A pesar de la gran sofisticación de los modelos multibanda, los modelos *Leq* tuvieron una sorprendente performance en las pruebas de audición. Sin embargo los descriptores que utilizan una sola banda están limitados por el compromiso de que la sonoridad está compuesta por una serie de curvas con distintos contornos que no puede ser duplicada con un solo filtro. [51]

Además estos métodos no pueden modelar la sumatoria espectral de la sonoridad [10], y comúnmente asumen que las señales dinámicas y estacionarias, con la misma distribución de energía, tienen la misma sonoridad. [51]

En 2007, el *Leq*(RLB) llegó a ser la base de la recomendación ITU-R BS.1770, junto con un prefiltro de +4 dB para compensar los efectos acústicos de la cabeza.[38].

Tiene la ventaja de ser simple, computacionalmente económico, muy aceptado y razonablemente robusto para contenidos de audio típicos.

También es la base de la recomendación R128 de la EBU, ya que toma a la norma BS.1770 como punto de partida para la revolución en la medición de los niveles de sonoridad.

El factor de cresta es la relación entre el valor medio y el valor pico, o sea, es el valor pico de la amplitud de la señal dividido por su valor RMS. Una métrica perceptual referenciada al valor de cresta, tendrá el error típico de definir el valor medio desde el valor pico. Es más la integración en el tiempo de estos valores no dará la misma impresión que el valor RMS. Por lo que su utilización está muy comprometida en las especificaciones de los nuevos medidores de sonoridad.

Un concepto que se está introduciendo es el del factor de cresta de la sonoridad (LCF). Esto es el máximo valor de la señal con respecto al valor medio de su sonoridad medidos en LK. De esta manera su cálculo es:

$$LCF = \frac{|x|_{pico}}{10^{0,05 (BS.1770)}} \quad (16)$$

Ecuación de cálculo del factor de cresta

Se introduce en el cálculo del valor de sonoridad de la norma BS.1770 el factor de corrección 0,691 para compensar la ganancia total de los filtros para obtener una ganancia de 0 dB a 1 Khz.

Otra solución propuesta es la presentada en el proyecto Replay Gain, la que computa la sonoridad de la señal contra la sonoridad de un ruido rosa de -20 dBFS, cuando se reproduce a 83 dB SPL (con una ponderación C, en modo Slow). Se ha intentado embeber el valor del Replay Gain dentro de los metadatos de los archivos de música usando el estándar ID3v2.

El procesamiento básico consiste en aplicar una compensación de frecuencia, computar su valor RMS sobre segmentos de 50 ms y usar un histograma para seleccionar la sonoridad excedida en solo el 5% de los bloques en el programa. [54]

Se ha presentado también el problema de los silencios o partes de poco volumen. Con los métodos de cálculo del promedio un silencio relativo puede causar una reducción no deseable de la sonoridad estimada. Dicho más genéricamente, poco se conoce sobre cómo las personas estiman la sonoridad a largo plazo en material con intensidad fluctuante. Si es por la porción silenciosa o por la porción más sonora. [53]

Como ha presentado Vickers [51], hablamos de “sonoridad” refiriéndonos al nivel de un sonido, no del silencio o “quietud” del mismo. Esto sugiere la posibilidad de que evaluemos más las partes fuertes que las tenues, cuando juzgamos la intensidad sonora a largo plazo en una señal con intensidad fluctuante. Zwicker y Fastl definieron que en muchos casos la sonoridad a largo plazo de sonidos con variaciones dinámicas importantes está bien caracterizada por su nivel N_5 , es decir por el valor de sonoridad que exceda el 5% del tiempo. [30].

Un estudio llevado a cabo por Zhang y Zeng encontró que para dos estímulos con el mismo nivel RMS, el que tiene mayores fluctuaciones temporales es muchas veces percibido como más fuerte. [55] Mientras otro estudio realizado por Johnston también sugiere que las señales con más dinámica se perciben como más sonoras que otras más estacionarias, teniendo ambas la misma energía. [51]

Otra conclusión opuesta se podría encontrar en los estudios llevados a cabo por Moore *et al*, donde encontró que la palabra comprimida parecía tener más volumen que la misma sin compresión. Pero aquí el tema es que la disminución de las crestas permite aumentar el volumen del material, dando en definitiva un mayor valor de sonoridad. Entonces si aumentamos los pasajes más suaves esto parece tener más influencia que la atenuación de los pasajes fuertes. Esto es un tema que requiere más estudios. [52]

Grimm *et al*, analizaron que las partes silenciosas podían ser descartadas del cálculo de sonoridad a largo plazo, por medio de un sensado del nivel, muy útil en películas con relativos largos segmentos silenciosos, y sugirieron el agregar algún control de nivel en el estándar BS.1770. [56]

El grupo de trabajo P/LOUD está finalizando la redacción de la recomendación R128, y sus reglamentaciones, que propone abandonar la normalización por picos a favor de la normalización por sonoridad, usando el descriptor de la sonoridad de programa (LU) computado aplicando el algoritmo recientemente presentado y que ofrece dos descriptores adicionales, el Loudness Range y el True Peak Level. [44]

También la Fundación “Pleasurize Music”, provee un medidor de rango dinámico con tecnología Tishmeyer. Este mide la diferencia acumulativa entre los valores pico y RMS durante la duración de una canción o todo un álbum, tomando en cuenta solo el 20% más sonoro. El valor resultante DR, se presenta como un número en dB. [57] Esta medición no incluye ninguna ponderación en frecuencia y parece (por las mediciones efectuadas con él) que responde a los valores de cresta más que a los valores RMS. Es evidente que el uso de la terminología “Rango Dinámico” es potencialmente confuso.

Otro acercamiento al tema fue desarrollado por Bob Katz con el sistema K-SYSTEM. Este no es solo una escala de medición, sino que involucra todo el sistema de edición. Consiste en comparar el valor RMS de la señal contra el nivel de un ruido rosa a -20 dBFS y este valor calibrado para una reproducción a 83 dBspl, ponderación C. Se dispone de tres escalas denominadas K-20, K-14, K-12, que presentan el valor de 0 dB a -20 dBFS, -14 dB y -12 dBFS de manera de permitir a los distintos entornos de reproducción una representación centrada en la escala del medidor. Su diseño es plano en frecuencias, por lo que recientemente se ha visto una versión ITU, que incluye la corrección de frecuencias propuesta por el documento BS.1770.

También se analizó el sistema Dorrough, que consiste en un medidor de gran escala con 40 leds, que mide simultáneamente el valor medio y el valor pico. Su desarrollo data de los años 80. Este no es un sistema, solo un medidor que no mide RMS verdadero, ni está ponderado en frecuencias. Su desempeño es mejor que un vúmetro de aguja. Y aunque se comercializa como un medidor de sonoridad no cumple con los estándares que actualmente se están proponiendo. [24]

4.2.4 Análisis de la arquitectura del Audio en PC

Como se utilizó una PC para las mediciones del prototipo, la necesidad de garantizar la fiabilidad de los resultados y la exigua información específica sobre algunas características eléctricas de las placas de audio, concluyó en encarar su estudio en detalle para conocer sus prestaciones de manera de avalar que las mediciones están dentro de los parámetros. Y en su defecto reconocer las limitaciones inherentes a su aplicación en las mediciones.

Como se ha visto el sistema de audio de la PC consta de dos importantes sectores, el hardware y el software. Como es lógico suponer las características particulares de un hardware no están necesariamente reflejadas en el software de control o driver. Estos son genéricos por fabricante, pudiendo encontrar importantes diferencias entre dispositivos aún de la misma familia de chips.

De manera genérica, se definen las facilidades de audio de entrada / salida, como se observa en la Figura 18 donde el mezclador de entradas acomete al conversor A/D y de allí sale al BUS PCI. La salida de datos desde el BUS PCI pasa por un conversor D/A y de allí al distribuidor de salidas de línea, esto se repite para cada canal de entrada y salida.

También están las entradas y salidas digitales serie a través de las conexiones S/PDIF o AES/EBU conocidas como AES 3.

En el caso de la interface utilizada, una placa Sound Blaster Audigy 2ZS, con un DSP EMU10K2. La arquitectura utilizada consta de una interface bus máster PCI, un sintetizador de ondas de 64 canales, un procesador de efectos de calidad profesional, conversores de tasa de muestras de alta calidad, con un oversampling de 92KHz y receptores y transmisores de audio digital de 24 bits. [34]

El sintetizador de ondas consiste en 64 osciladores de interpolación, un filtro digital resonante pasa-bajo de dos polos, tres generadores de envolvente y un amplificador mezclador con varias salidas seleccionables. El procesador de efectos permite rutear la salida sobre un sistema de múltiples parlantes (5.1 ó 7.1), satisfaciendo los requerimientos del entorno de una PC con capacidad para 31 entradas y 32 salidas. Al manejar 24 bits de resolución de cuantificación tiene un rango dinámico teórico de 144 dB, suficiente para cubrir todo el espectro del sistema auditivo humano. Las instrucciones internas usan 32 bits para tener margen en las operaciones de punto fijo, soportando la precisión adicional requerida para las operaciones complejas tales como el filtrado recursivo.

Los receptores asincrónicos permiten recibir audio directamente desde dispositivos SPDIF. Si la frecuencia de muestreo de estos dispositivos es de 44.1 KHz y como el EMU10K2 opera a 48 KHz, estas se procesan con la implementación de un conversor del tipo *Smith-Gosset* de 16 puntos. Lo que le permite soportar hasta tres flujos de datos estéreo asincrónicos simultáneos, usando una conversión asincrónica de alta calidad. [34]

La grabación de audio digital puede requerir distintas velocidades o frecuencias de muestreo para optimizar la resolución de bits y la capacidad de almacenamiento requerida, Las frecuencias estándar son 8 KHz; 16 KHz; 22,25 KHz; 32 KHz; 44,1 KHz, 48 KHz y 96 KHz. El método convencional requiere el uso de conversores A/D y D/A separados y de filtros analógicos más caros. El EMU10K2, en cambio utiliza conversores de tasa de bits de alta calidad de 64 bits, lo que permite el uso de chips Codecs estándar y filtros analógicos anti-aliasing sencillos. [34]

La grabación requiere muy poco uso de la CPU, al usar DMA del tipo Bus Master que graba directamente en RAM. Se pueden utilizar así los 32 canales de salida para grabaciones multipista, habilitándolo para ser usado en el entorno de un pequeño estudio.

En la figura 18 se presenta un diagrama en bloques de la sección de audio de una PC.

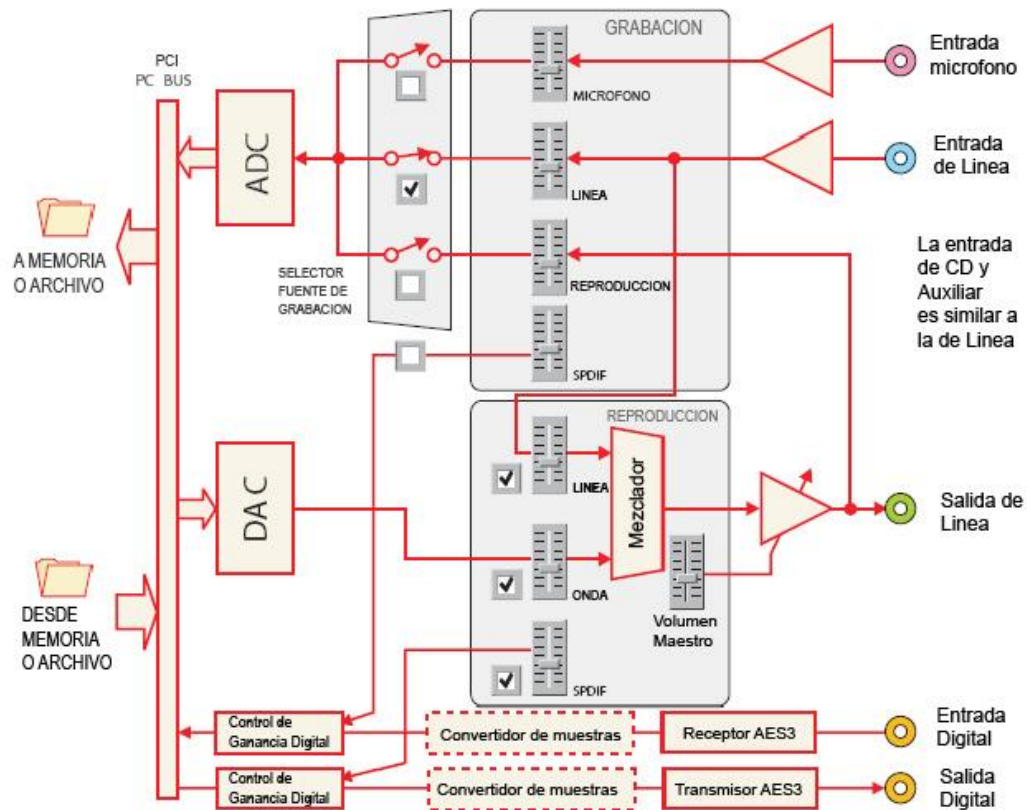


Figura 18: Diagrama de la sección de audio en una PC

Si se muestrea a 48 KHz con un filtro anti-aliasing de 24 dB/octava y un ancho de banda efectivo de 20 KHz y utilizando un rango de cuantificación de 16 bits para cada señal entrante, para una señal estéreo tendríamos: $16 \text{ bits} \times 48 \text{ KHz} \times 2 \text{ canales} = 1.536 \text{ Mbps}$. Esta velocidad no es muy alta para una PC y puede manejarla sin problemas, siempre que no haya otra rutina que la interfiera mientras está tomando las muestras cada 0,651 μseg .

En la placa utilizada las mediciones realizadas aportaron:

- Máxima salida de señal: 2,5V rms
- Ruido de fondo: -100 dB (ponderación A)
- Distorsión THD (<20Khz): 0,0001%
- Ancho de Banda (-1 dB): 20 a 16 KHz.
- Impedancia de entrada de línea: 10 Kohms.
- Impedancia de salida en Línea 1: 100 ohms

Se consideraron no significativos ni la distorsión ni el ruido. Se definió la salida de nivel correspondiente a 0 dBm a 0,775V rms, dando una salida máxima de +10 dBm.

Calibración:

Se calibró la salida para entregar 0,775V eficaces reproduciendo una onda senoidal de máxima amplitud en el *Adobe Audition*. Esto es, con sus picos coincidentes con el 0 dBFS. Este programa como muchos otros en el ámbito de la PC, miden los valores pico, y no los valores RMS. Por lo que el uso de los medidores incluidos en él, suele traer confusión a los usuarios no experimentados resultando en saturaciones. Personalmente he visto que el material de audio disponible en Internet está usualmente saturado.

Este programa trae la opción de calcular el valor rms de la señal lo que fue utilizado en las pruebas para compatibilizar los valores obtenidos en el medidor.

Una vez calibrados los controles del mezclador de Windows, estos no fueron modificados en ningún momento, de manera de garantizar una salida calibrada y confiable.

4.3 Implementación de un medidor de sonoridad

4.3.1 Criterios de diseño

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 software específico de edición de audio. Se estableció que se deberían 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 instantáneo y de corto plazo del medidor de sonoridad. Ya que su utilidad sería el reemplazo de un vúmetro, para lo que se diseñó un medidor autónomo conceptual. Primeramente 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 utilizaron las siguientes aplicaciones diseñadas para la edición de audio, el *Adobe Audition* v3.0., para la medición de los parámetros, se utilizó 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 utilizó el prototipo del **Loudness Meter** V1.02 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 y procesamiento digital se utilizó el **MATLAB** v7.10.0.499 (R2010a).

Un criterio importante en el desarrollo del prototipo, fue obtener un dispositivo de bajo costo, (indicado para nuestra región) y realizable con tecnología disponible en nuestro mercado, para lo que se comprobaron diversas configuraciones circuitales.

Se evaluó luego, 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, primeramente 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.

4.3.2 Diseño conceptual

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 tres 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 seg 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 trabajo.

Analizando los valores definidos como objetivo de la normalización se encontró que, el valor de referencia ó ‘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 EBU es de -23 LUFS +/- 2 LU, y el propuesto por la CST es de -25 LKFS +/- 2.

Esto llevo a considerar previo al diseño, un análisis de los valores propuestos.

El valor de -20 dB es para señales RMS senoidales, por lo que hay que agregar 3 dB al valor del fondo de escala digital, dando un valor de -23 LKFS.

Como se estimó además que la incorporación de una compuerta de nivel, dará un corrimiento de +/- 1 LKFS, se eligió el valor de -23 LKFS como referencia en este trabajo. Viendo que -24 LKFS (propuesto por algunos autores) están dentro de dicho margen, al igual que -25 LKFS +/-2.

Se define entonces el 0LU en el valor de -23 LUFS 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 el 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 matrizado 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 vúmetro, midiendo el valor medio. Esto permitió comprobar el funcionamiento de la matriz logarítmica de 30 diodos led y el concepto básico de disponer de un medidor con un paso de 1 dB por led.

El segundo prototipo integró la incorporación del filtro Head, el filtro RLB y el conversor de RMS. Con una presentación simultánea de ambos valores, mediante un sistema de multiplexado, permitiendo seleccionar y visualizar el valor pico, el de sonoridad ó ambos simultáneamente. Además de tener dos constantes de tiempo para la integración del valor de sonoridad.

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

Se utilizó además para comprobar la lectura de ambos parámetros, a fin de apreciar las velocidades de integración del instrumento, la relación de los niveles de sonoridad a los

picos de la señal y los valores de señal 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 ponderación RLB modificado, Leq (RLB).
3. Se calcula el valor RMS con distintos valores de integración. (RMS)
4. Se convierte la salida en una escala logarítmica en decibeles. (dB)

El algoritmo para el valor pico resulta:

1. Rectificación completa de la señal.
2. Filtro pasa-bajos asimétrico. El tiempo de ataque es menor a 10 μ seg y la descarga tiene una corriente constante de 0,27 μ A dando una pendiente exponencial.
3. Integra el valor de la envolvente.
4. Convierte la salida en una escala logarítmica en decibeles.

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

La Tabla 9 muestra los valores experimentales.

Nº 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

Nº 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
9	-19	-19,2
8	-20	-20,05
7	-21	-21,03
6	-22	-22,17
5	-23	-22,8
4	-24	-23,7
3	-25	-25,21
2	-26	-25,89
1	-27	-26,85

Tabla 9: Valores experimentales del prototipo de medidor

En la Figura 19 se puede apreciar el gráfico de la regresión lineal que muestra el error experimental.

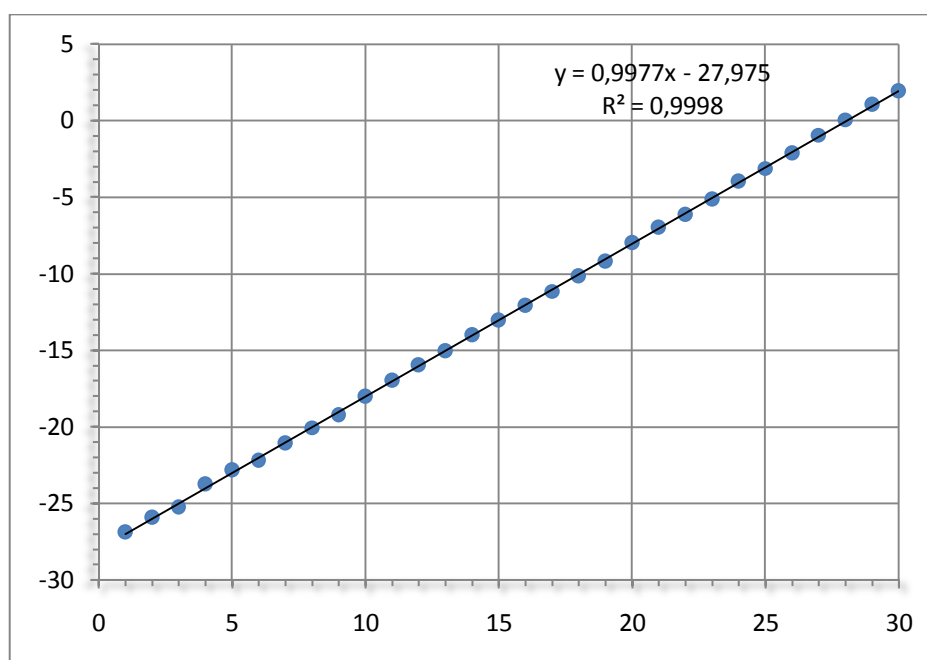


Figura 19: Regresión lineal de los valores del prototipo de medidor

El error del controlador de leds utilizado en el prototipo, según el fabricante es de 1,5 %, +/- 1 Led, dando en 30 dB, 0,45 dB de error, el conversor de RMS tiene un error máximo del 1%, por lo que la propagación de los errores da un resultado final de un error de +/- 1 LU, aceptable, ya que la correlación de 0,9998 confirma la linealidad del medidor experimental.

Por ser el medidor autónomo un instrumento de referencia visual incremental, el operador ve la diferencia (la velocidad de variación), con pasos de 1 LU, concluyendo que el instrumento tiene el error de un paso.

4.3.3 Experiencias

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 presentes grandes silencios, se propone incluir una corrección de la 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 solo 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 la 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 esta generada a partir del filtro RLB utilizado en la ponderación de frecuencias, cuya fórmula es:

$$h(z) = \frac{1 - 2z^{-1} + z^{-2}}{1 - 1.99004745483398z^{-1} + 0.99007225036621z^{-2}} \quad (17)$$

Además se incluyó el filtro de pre-énfasis a 1KHz de +4 dB, dado por el efecto de la cabeza. La curva de respuesta en frecuencia resultante se observa en la Figura 20.

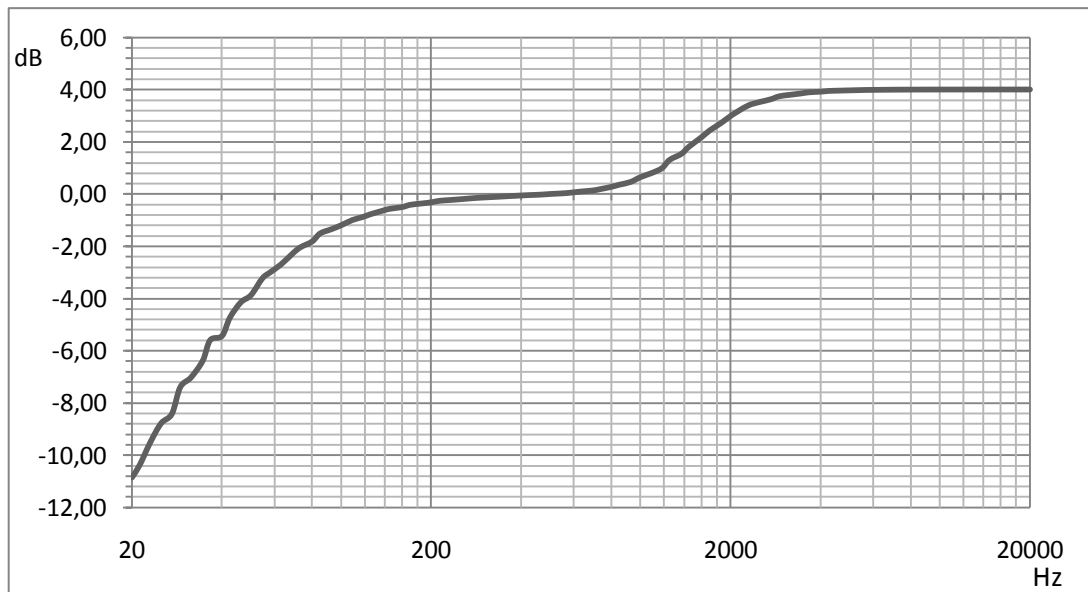


Figura 20: Curva de ecualización aplicada (real).

En la Figura 21 se muestra el diagrama en bloques del algoritmo utilizado en el prototipo.



Figura 21: Diagrama del algoritmo desarrollado en el prototipo

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 seg. Los archivos de música se tomaron de CDs y se ajustaron a 48 KHz.

Se desarrollaron luego diferentes algoritmos para acercarse al valor final, como 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 último 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:

$$Z_i = \frac{1}{T_c} \int_{T_{cj}}^{T_c(j+1)} y_i^2 dt \quad (18)$$

Donde Z_i = valor de la sonoridad del bloque

T_c = tiempo asignado al bloque

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

y_i = valor de nivel medido

Para un umbral de compuerta asignado por un valor Γ , 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:

$$\text{Sonoridad controlada} = -0,691 + 10 \log \sum G_i \left(\frac{1}{|j_g|} \cdot \sum_{j_g} Z_{ij} \right) \quad (19)$$

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 medio.

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 22 se observa el diagrama del algoritmo desarrollado en el MATLAB.

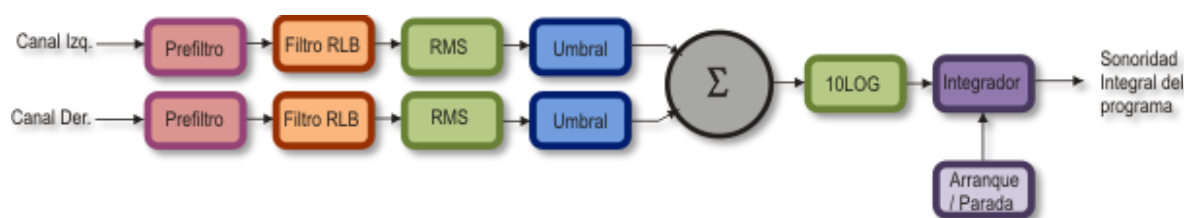


Figura 22: Diagrama del algoritmo desarrollado en software

4.3.4 Mediciones objetivas

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 9 muestra los niveles de corrección de los 16 archivos utilizados en la experiencia.

Archivo	Nombre	Sonoridad medida		
		Audition	Algoritmo	audición
1	Fantasía.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	-

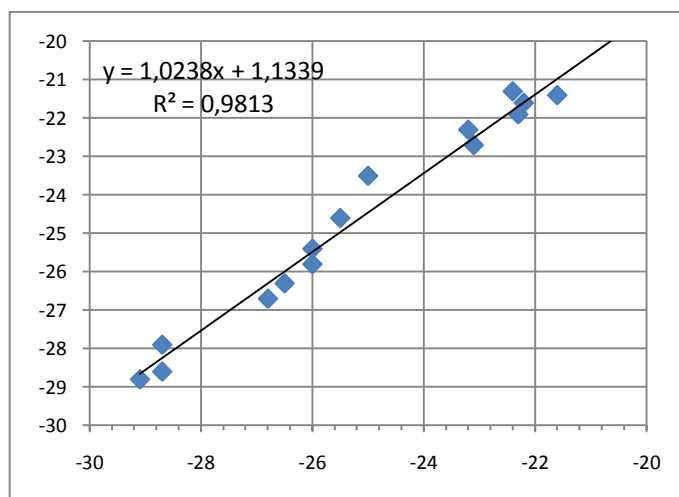
Tabla 9: Resultados de las mediciones

En la tabla 10 se muestran los valores de sonoridad para los distintos umbrales.

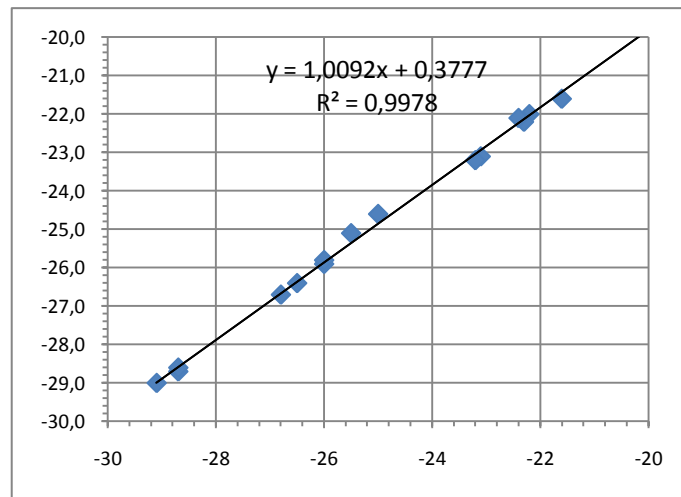
Pista	Sin umbral	Umbral -3LU	Dif	Umbral -8LU	Dif	umbral -10LU	Dif	Umbral -12LU	Dif
Fantasia.wav	-21,6	-21,4	0,2	-21,6	0,0	-21,6	0	-21,6	0
Bataille1.wav	-28,7	-28,6	0,1	-28,7	0,0	-28,7	0	-28,7	0
Música Country mujer.wav	-26	-25,8	0,2	-25,9	0,1	-26	0	-26	0
Multitud_1.wav	-26,5	-26,3	0,2	-26,4	0,1	-26,5	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
Diferencia Máxima			1,5		0,4		0,2		0,1

Tabla 10: Valores de sonoridad para distintos umbrales

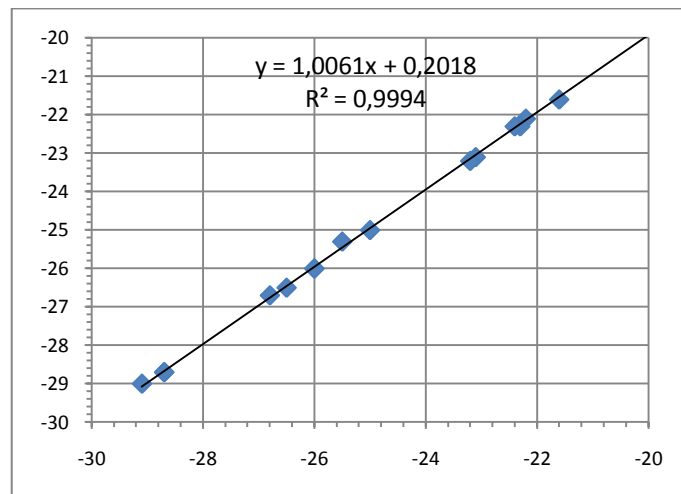
En la Figura 23 (a, b, c) se presentan los gráficos de correlación lineal de las variables, mostrando el grado de dispersión para los tres umbrales elegidos:



a) Umbral -3 LU



b) Umbral -8 LU



c) Umbral -10 LU

Figura 23 (a, b, c): Gráficos de correlación de datos

5: CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

5.1 Conclusiones

La necesidad de ajustar el nivel de las señales de audio con el modelo propuesto consiste en la normalización de los niveles por su sonoridad y no por sus valores pico.

Esto constituye un cambio de paradigma, una nueva visión del entorno de producción, y emisión de material de audio.

Este concepto revolucionario 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.

Cuando se mencionan valores de amplitud de una señal, es necesario indicar si se trata de valores pico ó valores eficaces. Porque de no estar indicado claramente el tipo de medición especificada se tiene una incertidumbre y el error puede ser catastrófico.

Ya que la relación entre el valor eficaz (RMS) y el valor pico, depende de la forma de la señal o sea de su factor de cresta. En las mediciones y las especificaciones normativas es necesario definir explícitamente estos valores.

No es conveniente en audio referirse a valores de sonoridad con ondas senoidales ya que la música tiene mucho más factor de cresta. Es más lógico utilizar ruido rosa ó 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 ó fondo de escala digital, está definido como dBDFS (dB Digital Full Scale) ó dBTP (dB True Peak). Por ejemplo en el uso de archivos WAV RIFF de 16 bits que utilizan little-endian con formato en complemento a 2, el máximo positivo es 0x7FFF y el máximo negativo es 0x8000, correspondiendo a +32767 y -32768.

En la evaluación de la sonoridad del audio digital, la utilización de los LKFS ó 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.

La utilización de tecnología informática permite fácilmente realizar estos cálculos y es posible construir indicadores que cumplan con estos nuevos parámetros.

De ahora en más 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, aún conteniendo pasajes silenciosos.

En las pruebas realizadas se encontró que un umbral relativo de unos -10 LU permite diferenciar entre el sonido de primer plano y el ambiental, determinando el valor de sonoridad integral más adecuado según el tipo de fuente o categoría del material de audio.

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, promocionales y avisos.

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.

Se ha encontrado que cuando se usa directamente el algoritmo de la recomendación ITU BS.1770, se presenta insuficiente para la evaluación sonora de cierto 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).

5.2 Recomendaciones normativas

Se necesita determinar en la Argentina 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, en 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 el 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 ó PRS y de mucho rango ó 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*).

5.3 Futuras acciones y líneas de investigación

Queda pendiente elaborar un medidor que integre un completo sistema multicanal 5.1 incluyendo el canal de efectos de baja frecuencia LFE.

Este canal no tomado en cuenta en el BS.1770, aporta a la sensación de sonoridad un componente que puede tener injerencia en las mediciones de contenidos con bajas frecuencias. Este es un tema que debe ser desarrollado. [50]

La forma de presentación podría integrar en una pantalla la información de valores momentáneos con barras en todos los canales y numéricamente el valor resultante de la sonoridad. También podría indicarse el valor del rango de sonoridad RS, el valor mínimo y el máximo. El dispositivo debe poder integrar la medición y graficar su desarrollo en una línea de tiempo.

Se ha planteado usar una escala tipo radar, donde la medición rota en un círculo dibujando las variaciones momentáneas. Esto permite una visión global de las variaciones así como un sencillo sistema de integración grafica.

Profundizar las investigaciones que permitan obtener una implementación del programa en lenguaje descriptivo del hardware, para su integración en dispositivos modulares.

La integración en un DSP (CPLD ó FPGA) del algoritmo, utilizando VHDL (ANSI/IEEE 1076-1993), permitirá encarar con procesamiento informático, la realización de dispositivos autónomos de múltiples usos. El bajo costo de estos dispositivos y un sencillo diseño industrial permitirán realizar medidores prácticos de uso general, para control y verificación de los parámetros del audio digital, integrables en otros dispositivos.

Considerar una iniciativa para difundir este cambio de paradigma proponiendo a los generadores y distribuidores de contenidos la incorporación de la normalización del audio digital en sus productos.

Se muestra en la Tabla 11 un ejemplo de las correcciones de nivel para la adaptación de los metadatos en el sistema HE-AAC del ISDB-T:

Sonoridad entrada en LUFS	Dif. en sonoridad en LU (PRL)	Cambio en los metadatos	Sonoridad corregida en LUFS	
-5.0	+18.0	+18	-23.0	
-6.0	+17.0	+17	-23.0	
-7.0	+16.0	+16	-23.0	
-8.0	+15.0	+15	-23.0	
-9.0	+14.0	+14	-23.0	
-10.0	+13.0	+13	-23.0	
-11.0	+12.0	+12	-23.0	
-12.0	+11.0	+11	-23.0	
-13.0	+10.0	+10	-23.0	
-14.0	+9.0	+9	-23.0	
-15.0	+8.0	+8	-23.0	
-16.0	+7.0	+7	-23.0	
-17.0	+6.0	+6	-23.0	
-18.0	+5.0	+5	-23.0	
-19.0	+4.0	+4	-23.0	
-20.0	+3.0	+3	-23.0	
-21.0	+2.0	+2	-23.0	
-21.9	+1.1	+1	-23.9	
-22.0	+1.0	0	-22.0	↑
-23.0	0.0	0	-23.0	Ganancia uno
-24.0	-1.0	0	-24.0	↓
-24.1	-1.1	-1	-23.1	
-25.0	-2.0	-2	-23.0	
-26.0	-3.0	-3	-23.0	
-27.0	-4.0	-4	-23.0	
-28.0	-5.0	-5	-23.0	
-29.0	-6.0	-6	-23.0	
-30.0	-7.0	-7	-23.0	
-31.0	-8.0	-8	-23.0	
-32.0	-9.0	-9	-23.0	

Tabla 11: Correcciones sobre los metadatos ISDB-T

Notas:

- La sonoridad es el valor incluyendo el factor de corrección de los metadatos.
- La corrección debe estar aplicada en números enteros, para el rango de ganancia unitaria de ± 1 LU, donde no se necesita offset.
- El rango indicado para el metadato Programme Reference Level (PRL) de +18 a -9 LU es un ejemplo y no significa una limitación.

Queda también proponer la publicación de un instrumento normativo que reúna las características técnicas y operativas, basado en los antecedentes y conclusiones presentados en este documento.

5.4 Actividades relacionadas con este trabajo de tesis

Se detallan las actividades realizadas durante el año 2010, relacionadas a este trabajo y las presentaciones referidas a mi actividad sobre este tema.

- Presenté la problemática de la sonoridad en las reuniones de la Comisión Mixta de Televisión Digital CAI-COPITEC (Centro Argentino de Ingenieros y Consejo Profesional de Ingeniería en Telecomunicaciones, Electrónica y Computación) donde expuse en varias ocasiones el tema, reconociéndome mis colegas como especialista en el mismo.
- Expuse el tema Sonoridad en la TVD en el curso de actualización profesional “Sistema Argentino de TV Digital Terrestre” del AFCEA (Asociación de Comunicaciones y Electrónica de las Fuerzas Armadas) en Abril del 2010.
- Realicé una clase abierta sobre el tema en el ISER (Instituto Superior de Enseñanza Radiofónica), con asistencia de los alumnos y los profesores de la casa.
- Publiqué un artículo en la revista especializada “Radio y Televisión Americana” sobre la sonoridad en televisión digital.
- Presenté un White Paper para el Congreso Mundial de Ingeniería 2010 del CAI.
- Expuse el tema en las conferencias académicas de CAPER 2010 (Cámara Argentina de Proveedores de Equipamiento para Radiodifusión) como representante del CAI.
- Publiqué un artículo en la revista HDMagazine, Washington, USA sobre la implementación de la televisión digital en nuestro país cubriendo las necesidades de audio correspondientes.
- Dicté un taller de 3 días sobre el tema Audio Digital, auspiciado por Patagonia Films y la Universidad de la Patagonia Austral en Río Gallegos.
- Obtuve la membrecía como integrante del grupo P/LOUD de la EBU (European Broadcasting Union), compuesto por especialistas en audio. Donde se han expuesto distintos asuntos sobre la implementación de este nuevo esquema de medición de la sonoridad del audio digital.
- Como así también actividades en la organización AES (Audio Engineers Association) filial Argentina.

6: 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”*, Broadcast 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 23rd 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, Waine, et al. (2003), *“Testing Challenges in PC Audio Devices”*, AES 114th Convention, Paper 5814, Amsterdam.
- [7] Intel Corp., (2010), *Especificación de Audio de Alta Definición*, “AZALIA”, Junio, <http://www.intel.com/standards/hdaudio/>. (ultimo acceso Nov. 2010)
- [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 117th Convention, San Francisco, CA.
- [11] Brixen, E.B., (2010), *“Audio Metering”*, Broadcast 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] INTEL, PC99, *PC 2001, System Design Guide*, © Intel & Microsoft, Intel Corporation and Microsoft Corporation. Intel Press (2001).
- [14] Chapuis R. J., (2003), “100 Years of Telephone 1878 - 1960“, Part 1, IOS Press, Netherlands, ISBN: 1 58603 349 2.
- [15] Dennis Bohn, (2003) “*Audio Specifications*”, RANE Note N° 145 , Paper N° 11628, RANE Corporation Inc. Mukilteo, WA. USA.
- [16] Dennis Bohn, (2004) “*Signal Processing Fundamentals*”, Note N° 134, , Paper N° 08834, RANE Corporation Inc. Mukilteo, WA. USA.
- [17] Millard, Andre, (1995) “*America on Record: A History of Recorded Sound*”. Cambridge University Press, ©1995. USA.
- [18] Juristas Contra el Ruido, (2010), “*FISIOLOGÍA AUDITIVA DEL OÍDO*”, Asociación Profesional N° 169.802 - 1ª. NIF: G 91243626 SEVILLA, España. (<http://www.juristas-ruidos.org/LaAsociacion.html>) accedido Nov. 2010
- [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. SAU.
- [21] John G. McKnight, (2006), “*Some Questions and Answers on the Standard Volume Indicator (vu meter)*”, 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., Users Manual. USA.
- [25] EBU (1979), “*Standard Peak Programme Meter*”, Tech 3205-E, 2º 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] Matti Zemack (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*” 2º ed., Information Series 22, Springer – Verlag, Berlin.
- [31] Zwicker E, (2002), “*Audio Engineering and Psychoacoustics...*”, parte del libro de Jeffay, Kewvin, “*Readings in Multimedia Computing and Networking*”, Chapter 1, Digital Audio, Academic Press, 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 Press; 3 ed., ISBN: 9780521688895
- [34] Thomas C. Savell, (1999) “*The EMU10K1 Digital Audio Processor*”, Joint Emu/Creative Technology Center; IEEE, pp: 49 – 57, ISSN: 0272-1732, USA
- [35] Wikipedia, (2010), “*Graphic Display Resolutions*” (ultimo acceso Nov. 2010)
http://en.wikipedia.org/wiki/Graphic_display_resolutions#HUXGA_.284800p.29

- [36] Bob Katz- *“The Lost Interview”* (2010), Written by Brent Randall, © ProRec.com
<http://www.prorec.com/index.php/home/interviews/item/63-bob-katz-the-lost-interview-part-1.html>
 Ultimo acceso: Nov. 2010
- [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 levels of audio signals”*, EBU Recommendation - P/LOUD Task Group. Switzerland
- [45] ABNT NBR 15602-2 (2007), *“Codificación de video, audio y multiplexación 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, 3º 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 , 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, 127th 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 AES119, 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.

7 ANEXOS

7.1 ANEXO I

Algoritmos EXPERIMENTALES

Implementación en Matlab del código para el filtro RLB

```

function RLBloudness = RLB(a1)
%
%--- Calculo de la sonoridad leq(RLB) ---
%
% Este script calcula la sonoridad leq(RLB), para un
% archivo mono o stereo 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.
% --- Abril 2010 VMSA ---
%
% Lee el archivo de audio
audio=wavread(a1);

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

%Calculos para un archivo estereo
if size(audio,2)==2

    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);

else
    %Calculos para un archivo mono
    %
    can1=filter(Bhead, Ahead, audio(:,1));
    can1filt=filter(BRLB, ARLB, can1(:));
    can1filt2=can1filt.^2;
    z11=mean(can1filt2);

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

RLBloudness=sono;

```

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

```

function ITUloudness = ITUOriginal(a1)
%
%--- Calculo de la Sonoridad de acuerdo a ITU-R BS. 1770-1 ---
%
% Este script calcula el valor de la sonoridad para archivos wav
% mono o estereo 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.
% --- Abril 2010, VMSA ---
%
% Lee el archivo de audio
audiol=wavread(a1);

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

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

%Calculos del filtrado para archivo estereo (2 canales)
if size(audiol,2)==2
    can1=filter(Bhead, Ahead, audiol(:,1));
    can2=filter(Bhead, Ahead, audiol(:,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);

else
    %Calculos del filtrado para archivo mono (1 canal)
    can1=filter(Bhead, Ahead, audiol(:,1));
    can1filt=filter(BRLB, ARLB, can1(:));
    can1filt2=can1filt.^2;
    z11=mean(can1filt2);
    sono=-0.691+10*log10(z11);
end
ITUloudness=sono;

```

Implementación del código Matlab para RLB con compuerta fija

```

function RLBgate = RLBgate(a1)
%
%--- Calculo de la sonoridad RLB con la funcion compuerta ---
%
% Calculo de acuerdo a RLB,
% pero los valores debajo del umbral se ignoran.
% --- Abril 2010, VMSA ---

% Lee el archivo de audio
audio=wavread(a1);

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

%Valor de la compuerta de discriminacion
%(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;

%calculo de acuerdo a RLB para archivo estereo

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 ordenados los indices
    % 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);

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

```

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

%Busca el indice en los vectores donde
%el umbra les mayor
a=find(ordcan1<umbral);

%Calcula el valor medio cuadratico 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;
```

Implementación del código Matlab para ITU BS.1770 con compuerta fija

```

function ITUG = ITUGate(a1)
%
%--- Calculo segun ITU-R BS.1770 con funcion de compuerta fija ---
%
% Calculo segun el estandar ITU, pero los valores debajo
% del umbral son ignorados.
% --- Abril 2010, VMSA ---

% Lee el archivo de audio
audio=wavread(a1);

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

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

%Valor del umbral de discriminacion
% (este valor corresponde a -10 dBFS)
%Cambiar para otros valores de umbral
umbral=0.1;

%Calculos de acuerdo a ITU-R BS.1770-1 (stereo)
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(:));
    can1filt2=af1ch1filt.^2;
    can2filt2=af1ch2filt.^2;

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

    %Encuentra el indice del vector, donde
    %el umbral es mayor.
    a=find(ordcan1<umbral);
    b=find(ordcan2<umbral);

    %Se calcula el valor medio incluyendo solo los
    %valores sobre el umbral en los vectores ordcan1 and ordcan2
    z11=mean(ordcan1((length(a)+1):length(ordcan1)));
    z12=mean(ordcan2((length(b)+1):length(ordcan2)));
    sono=-0.691+10*log10(z11+z12);

else

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

```

```
canlfiltcuad=canlfilt.^2;

%Se ordenan los valores al cuadrado
sortcan1=sort(canlfiltcuad);
%Encuentra el indice del vector donde
%el umbral es mayor
a=find(sortcan1<umbral);

%Calculo del valor medio incluyendo solo los valores
%sobre el umbral en el vector sortcan1
z11=mean(sortcan1((length(a)+1):length(sortcan1)));
sono=-0.691+10*log10(z11);
end

ITUG=sono;
```


Implementación en código Matlab para el algoritmo ITU-R BS. 1770 con compuerta dinámica

```
function ITUdb = ITUdb(a1,db)
%
%--- Calculo de la Sonoridad de acuerdo a ITU-R BS. 1770-1 con
compuerta dinamica ---
%
% Este script calcula el valor de la sonoridad para archivos wav
% de acuerdo a ITU-R BS. 1770-1 con compuerta dinamica
% 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 discriminacion.
% --- Mayo 2010, VMSA ---
%
% Lee el archivo de audio
audiol=wavread(a1);

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

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

%Calculos del filtrado para archivo estéreo 2 canales
if size(audiol,2)==2
    can1=filter(Bhead, Ahead, audiol(:,1));
    can2=filter(Bhead, Ahead, audiol(:,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);
else
    %Filtrado para archivo mono (1 canal)
    can1=filter(Bhead, Ahead, audiol(:,1));
    can1filt=filter(BRLB, ARLB, can1(:));
    can1filt2=can1filt.^2;
    z11=mean(can1filt2);
    sono=-0.691+10*log10(z11);
end

%calculo de dB de umbral
x = str2num(db);
```

```

umbral=(sono-x);

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

%calcula valor de sonoridad con compuerta dinámica estereo
if size(audio1,2)==2

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

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

    %Se calcula el valor medio incluyendo solo los valores
    %por encima del umbral dinamico
    z11=mean(ordcan1((length(a)+1):length(ordcan1)));
    z12=mean(ordcan2((length(b)+1):length(ordcan2)));
    sonodin=-0.691+10*log10(z11+z12);

else

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

    %Los valores cuadrados se ordenan
    ordcan1=sort(can1filt2);

    %Encuentra el indice en el vector de
    %los valores mayores al umbral
    a=find(ordcan1<umdin);

    %Calcula el valor medio incluyendo solo los valores
    %sobre el umbral dinamico en el vector ordcan1
    z11=mean(ordcan1((length(a)+1):length(ordcan1)));
    sonodin=-0.691+10*log10(z11);

end

ITUdb=sonodin;

```

7.2 ANEXO II

DEFINICIONES

Definiciones

Las siguientes definiciones describen las conexiones presentes en la mayoría de los dispositivos de audio en una PC.

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 transmitan 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 lazos 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. Esta mas 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. Además 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 muestrea a 48 Khz con 20 o 24 bits. En audio doméstico es de 44.1Khz, 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 frecuencias donde la salida casi no 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 frecuencias.

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 frecuencias 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 y 20 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 frecuencias.

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 respuesta 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 decibelios 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 la 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	Union Europea de Radiodifusión - European Broadcasting Union
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	Union 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)