

NUEVOS MÉTODOS y SISTEMAS DE MEDICIÓN DE NIVELES DE UNA SEÑAL DE AUDIO

Gabriel Data¹ – Guillermo Jardón²

PALBRAS CLAVES

AUDIO DIGITAL, MEDICION DE NIVELES DE SEÑAL, SONORIDAD

RESUMEN

Desde mediados de la década de 1990, las producciones de CD's de Audio sufrieron la tendencia a una sobre-compresión con el objetivo de conseguir altos valores de RMS con la nociva consecuencia de una drástica reducción del rango dinámico de la señal, niveles insólitamente altos y gran introducción de distorsión por saturación. Este proceso progresivo, conocido como la "Guerra del volumen" llevó a algunos Ingenieros de Sonido (como Bob Katz) a proponer nuevos métodos de evaluación y medición del nivel de señal basados en valores subjetivos óptimos de RMS y Factor de Cresta para diferentes tipologías de géneros musicales.

Por otro lado, la International Telecommunications Union (ITU) y la European Broadcasting Union (EBU) vienen realizando desde hace varios años recomendaciones para la creación de nuevos sistemas de medición de niveles de señales de audio digital con el objetivo de compatibilizar la sonoridad de emisiones radiofónicas y televisivas en lo que atañe a las diferencias de nivel entre géneros musicales, publicidad y otras producciones sonoras, sean éstas estéreo o multicanal. Tales recomendaciones demostraron ser compatibles con

¹ Investigador Externo por la Universidad Nacional de Rosario

² Investigador Externo por la Universidad Nacional de Rosario

nuevos medios de difusión como los formatos de streaming que incluyen, además, métodos de compresión sin pérdida de datos.

En Estados Unidos, se aprobó en septiembre de 2010 ley CALM (Commercial Advertisement Loudness Mitigation Act), la cual establece normas para controlar la sonoridad de los anuncios comerciales en televisión abierta, televisión por cable, satelital y diversas plataformas de distribución de material audiovisual.

En este trabajo relevaremos los diferentes métodos de medición actuales orientados a la evaluación y control del nivel de una señal de audio en tanto medios válidos para su aplicación en todo el proceso de producción musical, y aplicables particularmente a la etapa de masterización.

INTRODUCCIÓN

Con la radiodifusión y la telefonía surgen los primeros medios utilizados para la transmisión de una señal de audio; si bien en el primero de ellos una onda de radiofrecuencia modulada por otra de audiofrecuencia es emitida por una antena en la estación emisora, en ambos se manipula una señal de audio analógica (típicamente en la forma de variaciones de tensión en el tiempo). Posteriormente se unirá a este grupo la televisión. Consecuentemente, no resulta extraño que en todos estos medios haya sido vital contar con un método de medición del nivel de la señal para preservar la integridad de la misma, asegurando una adecuada relación señal-ruido sin que los picos superen el nivel de recorte (máximo nivel instantáneo sin saturación a la entrada o salida de un equipo). Los primeros dispositivos de medición aparecen después de mediados de la década de 1930: el Vúmetro, surgido en 1939 por el trabajo combinado de los laboratorios Bell y los organismos de radiodifusión CBS y NBC; y, por otro lado, originalmente desarrollado para su uso en transmisiones radiofónicas, surgen entre 1936 y 1937 (en forma independiente en Alemania y en el Reino Unido) los primeros Medidores de Pico o PPM (acrónimo de Program Peak Meter).

Dada la multiplicidad de métodos de medición coexistentes en lo que se refiere a la construcción y escalas de medición de Vúmetros y PPM's, en 1986 aparece la primera versión de la recomendación ITU-R BS.645 donde se propuso, por primera vez, un sistema unificado de calibración y estandarización de niveles de forma de compatibilizar el nivel de la señal de audio de las transmisiones radiofónicas a nivel internacional.

LA RECOMENDACIÓN ITU-R BS.645-2. "SEÑALES DE PRUEBA Y MÉTODOS DE MEDIDA PARA LOS ENLACES RADIOFÓNICOS INTERNACIONALES"

Estudiaremos aquí sólo algunos aspectos salientes de la norma con el objetivo de establecer un punto de referencia hacia las soluciones que en 1992 buscaban compatibilizar los sistemas de medición para señales de audio analógico en transmisiones radiofónicas

Entre los considerandos de la recomendación se especifica

que muchas degradaciones que se producen en el intercambio internacional de programas por enlaces radiofónicos son atribuibles a las distintas definiciones nacionales de las señales de prueba y métodos de medida diferentes empleados (ITUR BS.645, 1992)

y por lo tanto recomienda

que en los enlaces radiofónicos internacionales sólo se utilicen las señales de prueba definidas a continuación junto con los métodos de medida mencionados en los Anexos 1 y 2. (op. cit)

Las señales de prueba consisten en una Señal de Alineación, una Señal de Medición y una Señal Máxima Permitida.

La **Señal de Alineación** consiste en una señal sinusoidal con una frecuencia de 1 kHz, utilizada para alinear la conexión radiofónica internacional. El nivel de

la señal corresponde a 0 dBu0s, (0,775 V de tensión eficaz, en un punto de nivel relativo cero)³

La **Señal de Medición** es una señal sinusoidal con un nivel de 12 dB por debajo del nivel de la señal de alineación (debe utilizarse para mediciones de larga duración y mediciones en todas las frecuencias)

La **Señal Máxima Permitida** es una señal sinusoidal a la frecuencia de 1 kHz, con un nivel superior en 9 dB al nivel de la señal de alineación equivalente al máximo nivel permitido de la señal radiofónica. La señal radiofónica debe controlarse por el organismo de radiodifusión transmisor de manera que la amplitud de las crestas sólo rebase rara vez la amplitud de cresta de la señal máxima permitida.

Estas señales se utilizan para alinear tanto medidores PPM como también Vúmetros. En la figura 1 se especifica la relación entre estas señales y las escalas de medición utilizadas en diferentes países como aparece en la recomendación publicada.

Los medidores PPM a los que se refiere la recomendación, en realidad son del tipo “cuasi pico” e indican valores de cresta inferiores a los verdaderos

Resulta de vital importancia tener en cuenta la indicación que aparece en texto de la recomendación, particularmente cuando expresa que

5. El vúmetro indica el nivel medio del programa, generalmente mucho más bajo que la verdadera cresta. Los operadores tienen instrucciones para que la cresta de los programas corresponda generalmente a la lectura de 0

³ En la ITU- R BS.645 se introducen unidades de medida de nivel expresados en dBu, dBr y dBu0, siendo $\text{dBu} = \text{dBr} + \text{dBu0}$. Se sugiere consultar la norma disponible en forma gratuita en https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.645-2-199203-III-PDF-E.pdf para una mejor comprensión de estos términos.

VU. La experiencia ha demostrado que las verdaderas crestas de los programas superan los valores indicados en una cantidad comprendida entre +6 dB y +13 dB, según la naturaleza del programa. Cuando se toman además en consideración los errores de los operadores, las verdaderas crestas de la señal pueden ser hasta 16 dB más altas de lo indicado, correspondiendo a la amplitud de cresta de una señal sinusoidal de +16 dBu0s, o +14 dBu0s, según el caso, cuando la aplicación de la señal de nivel de alineación da como resultado una indicación de +2 vu. (op. cit)

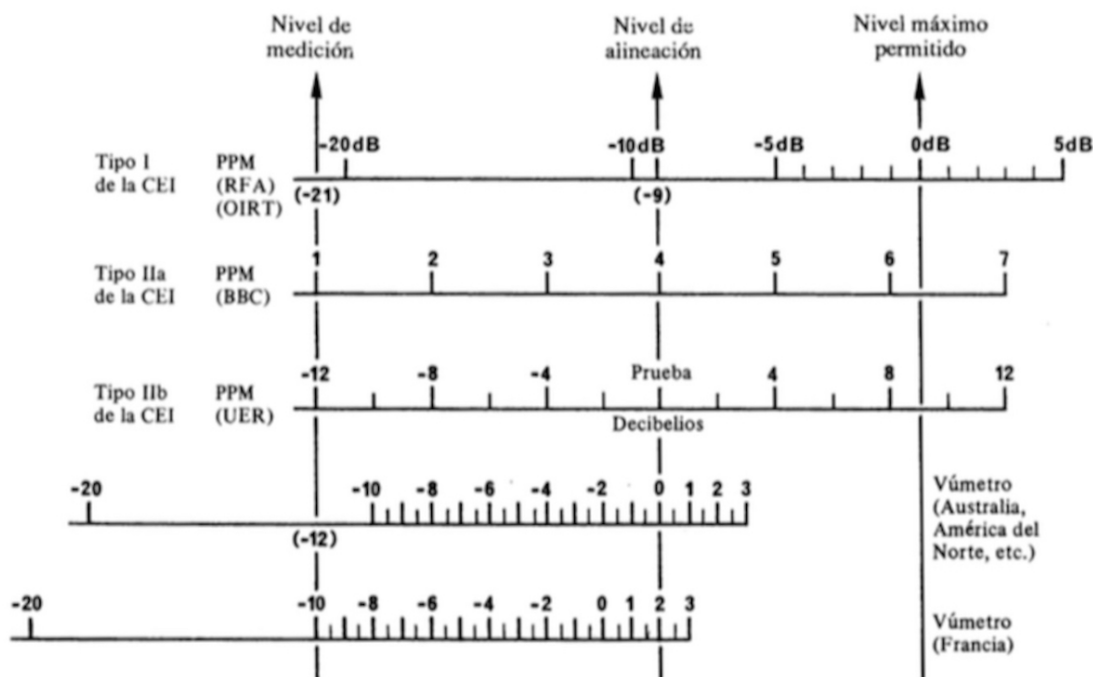


Figura 1. Indicaciones producidas por medidores de varios tipos con las señales de prueba recomendadas

Entonces, las lecturas obtenidas en el Vúmetro sólo expresan la relación del nivel de la señal (cercana al valor de RMS de la misma) respecto al nivel de alineación como una forma (junto a los valores de cresta) de proteger el equipamiento del empobrecimiento de la relación señal ruido y de la introducción de saturaciones lo que está sujeto, además, a la experiencia y

habilidad de los operadores en el procesamiento de la señal.

Por otro lado, un vúmetro es preciso si se trata de señales sinusoidales (con un factor de cresta de +3 dB) y no para evaluar el nivel de una señal compleja; por sus características constructivas (por su balística de 300 ms) la deflexión de la aguja resulta eficaz en el rango de frecuencias de la voz humana y son absolutamente imprecisos para frecuencias graves. En síntesis, nada dicen de la Sonoridad de un programa determinado y esto muestra la total ineficacia de este dispositivo para comparar niveles entre programas diferentes con el fin de lograr un equilibrio dinámico entre ellos.

EBU R68-2000. “NIVEL DE ALINEACIÓN EN EQUIPOS DE PRODUCCIÓN DE AUDIO DIGITAL Y EN GRABADORES DE AUDIO DIGITAL”

En 1992 se publica la recomendación EBU R68, posteriormente revisada en el año 2000, la cual tiene por objetivo buscar la alineación entre los niveles y de señales de audio digital entre sí y de los equipos digitales utilizados en transmisiones radiofónicas, compatibilizándolas con la recomendación ITU-R BS.645-2. Resulta, entonces, que con esta recomendación, se busca compatibilizar los niveles y sistemas de medición en sistemas de sonido que integran tanto equipos digitales como analógicos.

Entonces, teniendo en cuenta la relación de 9 dB entre el nivel máximo permitido y el nivel de alineación para señales y equipos analógicos y que para audio digital el único método de especificación del nivel de señal se referencia al máximo nivel de codificación posible (hoy conocido como 0dBFS⁴), una primera aproximación diría que los equipos digitales deberían calibrarse usando una señal de alineación sinusoidal con un nivel de RMS de -9dBFS.

⁴ En la recomendación no se usa la unidad dBFS. Esta fue propuesta en 1987 por James Colotti en su trabajo “Digital Dynamic Analysis of A/D Conversion Systems through Evaluation Software based on FFT/DFT Analysis”, hoy universalmente aceptada y por tal razón la incluimos aquí.

Pero si se tiene en cuenta que los medidores PPM son “cuasi pico” y que los errores de operación pueden provocar picos con más de 6 dB que el indicado o 15 dB sobre el nivel de alineación, la recomendación EBU R68 propone utilizar una señal digital de alineación con un nivel de -18 dBFS independientemente del número de bits utilizados. Este nivel de alineación, entonces, constituye el primer intento de alcanzar un nivel de alineación estandarizado.

ALINEACION DE NIVELES EN OTROS TIPOS DE PRODUCCIONES SONORAS

Todo lo dicho hasta aquí se refiere a alineación de sistemas de sonido orientados a transmisiones radiofónicas y televisivas; pero para otro tipo producciones de sonido profesional (CD de Audio, Audio para DVD, etc.) existe otra práctica muy aceptada (aunque no hay una norma que la estandarice) de alineación entre equipos analógicos y digitales que sugiere el uso de una señal digital de -20 dBFS de RMS como señal de alineación, la que debe dar una lectura de 0 VU en un vúmetro analógico. La razón de esta elección se basa en el hecho que un equipo profesional de audio analógico tiene un nivel nominal de salida de +4 dBu (1,23 voltios) con un nivel máximo admisible de señal sin saturación de +24 dBu (12,28 voltios), esto es, 20 dB de Headroom. Por tanto, la relación resulta: +4dBu = 0 VU = -20 dBFS.

EQUILIBRIO DE NIVELES DE SEÑAL BASADOS EN LA SONORIDAD

En el proceso de Masterización de audio digital para CD de audio, el equilibrio en la dinámica entre los diferentes tracks es uno de los objetivos más importante a alcanzar. Pero a la hora de utilizar herramientas para la medición de niveles, el contar con vúmetros (aún digitales) para constatar el valor de RMS de la señal y medidores de pico lejos están de lograr este objetivo. Por un lado, los valores de RMS no están correlacionados con lo que percibimos a la hora de realizar una comparación entre diferentes señales. La percepción de la intensidad sonora depende: del rango dinámico de la señal, de la composición espectral, del Factor de Cresta, de los transitorios de ataque, etc. Se podría decir que el valor de RMS es el “centro de gravedad” de la señal de audio. Por

otro lado, como ya dijimos, el máximo valor de señal admisible corresponde al máximo nivel de codificación posible acorde a la resolución en bits utilizados; pero los medidores de pico (con un tiempo de integración de unos 10 milisegundos), dejan de medir los valores de todas las muestras, resultando que los picos “verdaderos” sobrepasen el 0 dBFS en varios dB’s sin que esto se refleje en la medición. Entre mediados de los ‘90 y primera mitad de la década del 2000, en las producciones de la mayoría de CD’s de Audio se priorizó el objetivo de alcanzar altos niveles de RMS de la señal con la consecuente sobre compresión del material sonoro, reducción ostensible del rango dinámico y del headroom de la señal en una corriente de producción conocida mundialmente como la “Guerra del Volumen”. Si a esto le sumamos la popularidad del uso de algoritmos de compresión con pérdidas de datos (como el mp3) la calidad del audio se vio severamente degradada.

Con el progresivo incremento de los medios de compra y reproducción “en línea”, ha disminuido la popularidad del CD para la distribución de producciones musicales en formato de audio digital, siendo reemplazado progresivamente por dispositivos móviles, compra y descarga de archivos de audio por internet o por *streaming* en tiempo real. Este cambio implica un cambio paradigmático en la forma en que se escucha música hoy, lo que incluye también el uso de “listas de reproducción” donde se mezclan audios de diferentes producciones, como ocurre en las transmisiones radiofónicas y televisivas. Las preguntas inmediatas son: ¿cómo equilibrar los niveles ante la diversidad de “programas sonoros” puestos en la lista de reproducción? y ¿qué sistema de medición es el más adecuado para “visualizar” este equilibrio y que resulte consistente con lo que percibimos auditivamente?

El parámetro acústico que se correlaciona con nuestro sistema de percepción de la dinámica es la Sonoridad; pero al ser éste un parámetro subjetivo, la disyuntiva fue encontrar un sistema de medición que lo haga objetivo (medible), al menos para la mayoría de las personas.

Recomendaciones EBU R-128 e ITU-R BS 1770.

En el año 2010 la European Broadcasting Union (EBU) elabora una serie de recomendaciones para la implementación de la normalización por sonoridad, para programas de radio y televisión, denominadas EBU R-128. Estas recomendaciones detallan de qué forma se debe medir la sonoridad y cómo debe realizarse la normalización por sonoridad de los programas.

La EBU R-128 se basa en las recomendaciones de la International Telecommunication Union, denominadas ITU-R BS 1770 y, en cuatro documentos técnicos adicionales, denominados EBU Tech Doc 3341, 3342, 3342 y 3344, donde se detalla la forma de implementación de la normalización por sonoridad y se definen los siguientes parámetros y las pautas para su medición:

- **Sonoridad de programa** (Programme Loudness): es la sonoridad integrada a largo plazo durante todo un programa.
- **Rango de sonoridad** (Loudness Range): variación estadística entre el mínimo y máximo nivel de sonoridad de un programa.
- **Nivel de Cresta Verdadero** (True Peak Level): indica el valor máximo (positivo o negativo) en el dominio temporal continuo de una señal de audio.

ITU-R BS 1770

En julio del año 2006 la International Telecommunication Union publica las recomendaciones ITU-R BS 1770 en las que se describen los algoritmos para medir la sonoridad de los programas de radiofónicos (Program Loudness) y el nivel de cresta de audio real (True Peak)

Algoritmo de medición de sonoridad

El algoritmo de medición de sonoridad descrito en las ITU-R BS 1770 es multicanal, pudiéndose implementar en sistemas de uno a cinco canales excluyendo el parlante LFE de graves.

El algoritmo consta de cinco bloques de procesamiento de señal, cada canal

consta de cuatro bloques de proceso, independientes por canal, y el último bloque realiza la sumatoria de las sonoridades de todos los canales del sistema.

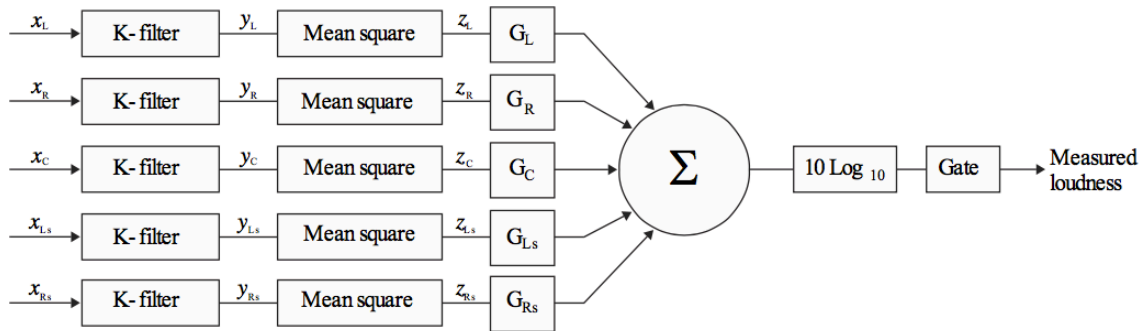


Figura 2. Diagrama de bloques del algoritmo de medición de la sonoridad multicanal

En el primer bloque la señal es procesada mediante un filtro que compensa los efectos acústicos de la cabeza humana, en el segundo bloque se aplica otro filtro denominado RLB cuya curva se denomina B modificada, a la curva resultante de la aplicación de los dos filtro se la denomina curva de ponderación K (K-weighting) que emula la variación de percepción del nivel de presión sonora en función de la frecuencia del oído humano (sonoridad)

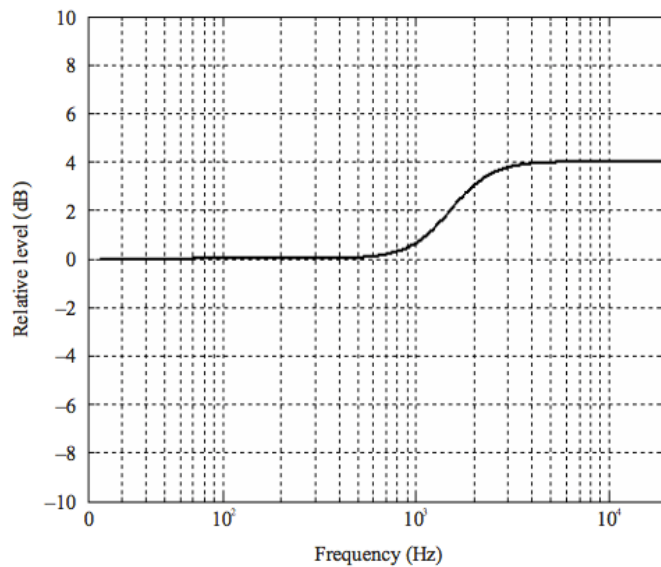


Figura 3. Respuesta del filtro previo utilizado para representar los efectos acústicos de la cabeza

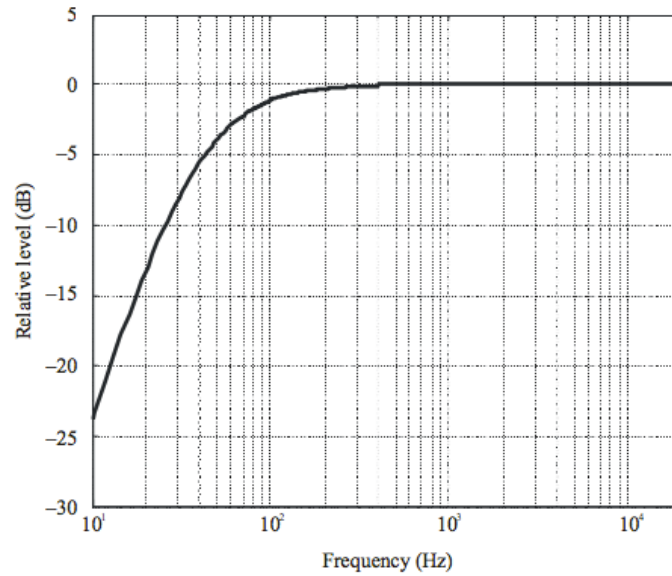


Figura 4. Curva de ponderación RLB

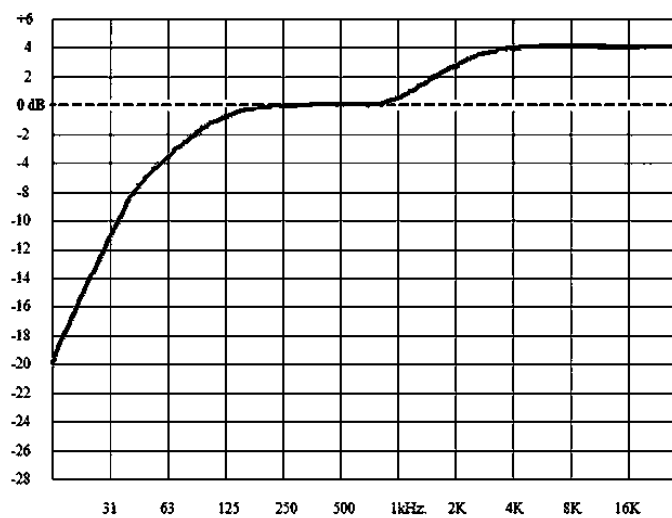


Figura 5. Curva de ponderación K. Imagen tomada de Dave Moulton en <http://www.tvtechnology.com/audio/0098/using-bs---for-fun-and-profit/213234>

El tercer bloque realiza el cálculo de la energía cuadrática media de cada canal en un intervalo de tiempo estipulado.

En el cuarto bloque se pondera la ganancia de cada canal de acuerdo al ángulo de llegada, 0dB para los canales frontales y +1.5 dB para los canales traseros.

El quinto bloque realiza la sumatoria lineal de todos los canales para obtener la sonoridad total del programa. La fórmula resultante es la siguiente:

$$\text{Sonoridad} = -0,691 + 10 \log_{10} \sum_i^N G_i \cdot z_i$$

Donde: -0.691 es la constante que indica el efecto de los bloques uno y dos.

G_i es la ponderación aplicada a los canales de entrada

$$z_i = \frac{1}{T} \int_0^T y_i^2 dt, \quad \text{representa los valores de energía cuadrática media de cada canal.}$$

canal.

En una revisión de marzo del 2011 la ITU agrega importantes recomendaciones: fija el tiempo de medición mediante la conformación de bloques de 400 ms (solapamiento del 75%). Además indica la utilización de una compuerta con dos umbrales, el primero a -70 LKFS (sonoridad, ponderada en K, en relación con la plena escala nominal), que al ser superado, el algoritmo comienza a medir la sonoridad, esto evita las mediciones cuando los niveles de sonoridad son mínimos, en general en el principio y el final del programa. El segundo umbral a -10 LKFS con respecto al nivel medido tras la aplicación del primer umbral. Los sonidos por debajo de este umbral no son medidos, esto evita que se incluyan períodos de silencio o bajo nivel en el cálculo total de la sonoridad del programa.

En las recomendaciones EBU R-128 se estandariza la utilización de LU (Loudness Units) como unidad de medida de sonoridad, siendo $1\text{LU}=1\text{dB}$ y LUFS (Loudness Units relative to Full Scale) como unidad de medida en relación a la escala digital completa, siendo análogo a LKFS.

Poner en funcionamiento este algoritmo es sencillo y los costos muy bajos, lo cual favorece de gran manera la utilización del mismo como estándar internacional.

Algoritmo de medición del Nivel de Pico Real (True Peak)

El algoritmo de medición del nivel de pico real descrito en las ITU-R BS 1770 soluciona el problema que se presenta al medir los niveles de pico de una señal digitalizada cuando el nivel máximo de la señal de audio analógico se encuentra entre dos muestras de digitalización, dando como resultado un valor distinto del nivel real de cresta de la señal. El algoritmo está planteado para señales digitales con frecuencia de muestro de 48KHz y consta de seis bloques de procesamiento.

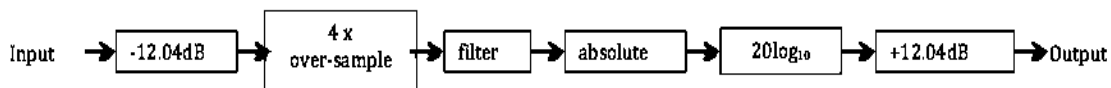


Figura 6. Diagrama de bloques del algoritmo de medición del nivel de pico real

El primer bloque atenúa la señal en -12.04dB (desviación de 2 bits) para dar un margen de seguridad y evitar la saturación de la señal en los procesos siguientes, si la digitalización se realiza con coma flotante no es necesaria la aplicación de la atenuación. El segundo bloque sobremuestra la señal al cuádruple (x4) de la velocidad de muestreo, pasando de 48Khz a 192KHz, aumentando la precisión de la representación digital de la señal analógica. Si la velocidad de muestro es mayor, por ejemplo 96KHz la tasa de sobremuestro será menor, en este caso se duplica (x2). Posteriormente en el tercer bloque se calculan los valores en bits de las muestras, mediante la aplicación de un filtro y en el cuarto bloque se obtienen los valores absolutos de las muestras, invirtiendo los valores negativos. Luego se aplica logaritmo a los resultados obtenidos ($20\log_{10}$) para obtener los valores en dBTP (decibeles True Peak). Al final de estos procesos se debe compensar la atenuación que se aplica en el primer bloque aumentando la señal +12.04dB.

EBU Tech Doc 3341, 3342, 3343 y 3344.

Los documentos EBU Tech detallan de qué forma se deben producir y

transmitir los contenidos normalizados por sonoridad.

EBU Tech Doc 3341.

El documento EBU Tech Doc 3341 describe las funciones y característica de un medidor de sonoridad denominado EBU Mode, basado en las recomendaciones ITU Bs 1770, desarrollado para la implementación de la normalización por sonoridad.

El medidor debe poseer tres formas distintas de medir la sonoridad

1) Momentary Loudness [M] (Sonoridad Momentanea): mide la sonoridad mediante una ventana temporal deslizante de 400ms, sin utilizar la compuerta de umbral -10 LU descrita en la ITU Bs 1770.

2) Short Term Loudness [S] (Sonoridad de Corto Plazo): mide la sonoridad mediante una ventana temporal deslizante de 3s, sin utilizar la compuerta de umbral -10 LU.

3) Integrated Loudness [I] (Sonoridad Integrada) mide la sonoridad promedio del programa desde el comienzo utilizando la compuerta de umbral a -10LU.

Todas las mediciones deben ser realizadas utilizando el algoritmo de medición de niveles de pico real (True Peak) descrito en la ITU Bs 1770.

El monitor (display) del medidor puede indicar solamente números a escala completa (Full Scale) o indicar escalas relativas, en ese caso se proponen dos escalas distintas:

1) rango -18.0 LU a +9.0 LU (-41.0 LUFS a -14.0 LUFS), llamada 'EBU +9 scale'

2) rango -36.0 LU a +18.0 LU (-59.0 LUFS a -5.0 LUFS), llamada 'EBU +18 scale'

La escala EBU+9 deberá ser la escala por defecto, para programas con mayor rango de sonoridad se utilizará la escala EBU+18.

Además el medidor debe indicar el promedio de sonoridad del programa completo, denominada Sonoridad del Programa (Programme Loudness), el Nivel de Pico Máximo Real (Maximum True Peak Level) y el Rango de Sonoridad (Loudness Range). El Rango de Sonoridad es definido y descrito en el documento EBU Tech 3342.

En la siguiente figura se muestra la interfaz del plugin WLM desarrollado por la empresa Waves para medir los niveles de una señal de audio digital donde se muestran los parámetros detallados.



Figura 7. Interfaz del plugin WLM de Waves

EBU Tech Doc 3342.

El documento EBU Tech Doc 3342 describe el parámetro denominado Rango de Sonoridad (Loudness Range- LRA). El Rango de Sonoridad es el cálculo de la variación de la sonoridad a escala temporal macroscópica, no en una

ventana temporal fija. Los cálculos de sonoridad se realizan en base a las recomendaciones ITU-R BS.1770, pero con un umbral de compuerta variable distinto. Se utiliza para determinar si se necesita tratamiento de la dinámica de la señal para hacerla coincidir con los requerimientos de una plataforma o canal de transmisión particular. No debe confundirse al Rango de Sonoridad con Rango Dinámico

El Rango de Sonoridad es un cálculo estadístico de las mediciones de sonoridad. El rango de distribución de los niveles de sonoridad es determinado mediante la estimación de la diferencia entre un bajo percentil y un alto percentil de la distribución de dichos niveles.

Se define como Rango de Sonoridad a la diferencia entre la estimación de los percentiles 10° y 95° de la distribución de sonoridad de un programa. Se determinan estos percentiles para evitar que variaciones pequeñas de sonoridad como un fade out en una canción o un disparo en una película afecten el cálculo estadístico de Rango de Sonoridad. La unidad utilizada para el cálculo es LU. Los percentiles pertenecen a una estadística no paramétrica y son empleados en el cálculo del Rango de Sonoridad porque los niveles de sonoridad no pueden en general ser asumidas como pertenecientes a una distribución estadística particular.

El algoritmo utilizado para la medición consta de dos compuertas, una de umbral fijo a -70 LUFS y otra de umbral variable a -20 LU por debajo del Nivel de Sonoridad medido con compuerta. La compuerta variable asegura que no se midan períodos de baja sonoridad que afecten el resultado final. Se utiliza una ventana temporal de 3s para el cálculo con un mínimo de 2.9s de solapamiento entre ventanas, para evitar problemas de precisión. La aplicación de las compuertas en cascada deja solo los niveles de sonoridad del audio en primer plano y los niveles medios de sonoridad del audio en segundo plano, eliminando los niveles bajos de sonoridad, los sonidos de fondo y los silencios.

El valor máximo individual de Rango de Sonoridad generalmente depende del género, plataforma, entorno de escucha y edad del público al que está

destinado el programa.

EBU Tech Doc 3343.

El documento EBU Tech Doc 3343 determina las directrices prácticas para la producción e implementación de la normalización por sonoridad conforme a la EBU R 128. El texto detalla conceptos generales de la normalización por sonoridad, estrategias para la normalización por sonoridad, qué medir en producción y post-producción, producción y serialización basados en ficheros, utilización de metadatos, calibración de señales, implementación y migración, y cuestiones específicas de los géneros televisivo y radiofónico.

Dentro del texto se definen los valores para los siguientes parámetros:

Sonoridad de Programa (Programme Loudnes): -23 LUFS \pm 1 LU

Nivel de Pico Máximo Permitido (Maximun True Peak Level) -1dBTP

Además se proporcionan diferentes ejemplos de Rango de Sonoridad en función del entorno de reproducción.

EBU Tech Doc 3344

Este documento proporciona directrices prácticas para los sistemas de distribución conforme a la recomendación EBU R-128. En el mismo están explicadas la formas de emisión y re-emisión de servicios de televisión y radio, la normalización por sonoridad en sistemas de distribución digital y de distribución analógicos, la alineación de niveles en sistemas de distribución digitales y analógicos, el trabajo con sintonizadores externos, receptores-decodificadores integrados, televisores, IDTVs, home theater, reproductores multimedia y receptores de radio FM y DAB .

Sobre las Unidades L_k LU, LKFS, LUFS

En las recomendaciones ITU BS.1770 versión 1 publicadas en el año 2007 (la primer versión de las normas se denomina 0 y fueron publicadas en el año

2006) se introduce la denominación LKFS como unidad del valor de Sonoridad de Programa. Se define LKFS como unidad de sonoridad absoluta a la escala completa (Loudness, Kweighted, relative to nominal full scale) que significa, sonoridad, ponderada en K, en relación con la plena escala nominal. Originalmente esta medición se realizaba sin la compuerta variable que evita medir los sonidos de bajo nivel.

Posteriormente en el año 2010 se publican las recomendaciones EBU R128 donde se proponen las denominaciones LUFS (Loudness Units, relative to digital Full Scale) como unidad de sonoridad relativa a la escala completa nominal y LU (Loudness Units) como unidad relativa de sonoridad, utilizando una compuerta variable a -8LU para evitar medir los sonidos de bajo nivel. Al introducir la compuerta de -8LU en la medición se produce una diferencia entre LKFS (propuesto por ITU) y LUFS (propuesto por EBU). La diferencia entre las mediciones en LKFS sin compuerta y LUFS con compuerta es entre 1 y 2 LU mayor en LKFS para un programa de 20 a 30 minutos de duración.

En marzo del 2011 la ITU publica una actualización de las recomendaciones denominadas ITU BS.1770-2, donde incorpora las recomendaciones de las EBU R128 agregando la compuerta variable para evitar medir los sonidos de bajo nivel. La ITU y la EBU acuerdan fijar el umbral de la compuerta variable a -10LU con respecto al Nivel de Sonoridad, a partir de ese momento LKFS y LUFS son iguales. Por lo tanto:

$$\text{Nivel de Sonoridad } L_k = x \text{ LKFS} = x \text{ LUFS}$$

Por el momento no hay acuerdo entre la ITU y la EBU para la utilización de un solo término por lo tanto conviven los dos términos como unidad de sonoridad relativa a la escala completa.

Para realizar la alineación de señales en broadcasting no hay modificaciones en el método. En un sistema stereo utilizando para la calibración una señal sinusoidal de 1KHz, en fase en los dos canales, a -18dBFS se obtiene como Nivel de Sonoridad $L_k = -18\text{LUFS}$ o -18LKFS en el medidor de sonoridad en la

escala completa (+5LU en la escala relativa Modo EBU). Si se utiliza en un solo canal la señal sinusoidal de 1KHz a -18dBFS se obtendrá en el medidor de sonoridad -21LUFS (+2LU en la escala relativa Modo EBU)

En las recomendaciones ITU-R BS.1770 se aclara que dB y LKFS son equivalentes, un aumento de +1dB en la señal incrementará la sonoridad en +1LKFS.

K-SYSTEM

Las recomendaciones anteriores proveen métodos muy eficientes de medición con el fin de lograr una estimación objetiva de la Sonoridad de un programa de audio y de “Pico Verdadero” en transmisiones radiofónicas y televisivas mediante algoritmos relativamente fáciles de implementar; pero la Sonoridad, como parámetro de la psicoacústica, depende directamente del Nivel de Presión Sonora promedio y de los transitorios de la señal acústica. Orientadas al control de los niveles de señal para su transmisión, nada dicen sobre las condiciones de escucha óptimas para la producción y post producción del material sonoro⁵. Específicamente se reconoce

(...) que la sonoridad medida es una estimación de la sonoridad subjetiva e involucra un cierto grado de discrepancia sobre los oyentes, el material de audio y

⁵ En mayo de 1998, se publica la EBU Tech 3276-E donde, en el apartado 2.5 “Niveles de escucha” se sugiere utilizar ruido rosa con un nivel de señal de alineación de -18 dBFS y ajustar las ganancias de cada altavoz con la siguiente fórmula:

$$L_{LISTref} = 85 - 10 \log(n) \text{ dBA}$$

Donde n es el número de canales de reproducción

No se da referencia del origen de tal fórmula, la ponderación del sonómetro muestra el énfasis en la palabra y no en la música y extrañamente no aparece referenciada en las recomendaciones R128 ni en las EBU Tech's

condiciones de escucha (ITU-R BS.1770-2, 2011)

Pero en la producción y post producción de audio para cine y video es fundamental el control de los niveles de señal teniendo en cuenta las “condiciones de escucha”. En el año 2000, los laboratorios Dolby publicaron una guía de producción para sonido en formato 5.1 en la que se detalla el método de calibración del sistema de monitoreo: un generador de ruido rosa con un nivel de -20 dBFS debe producir, en la posición de monitoreo, un nivel de presión sonora de 85 dBC por cada canal frontal (medido con sonómetro con ponderación C y respuesta lenta) y de 82 dBC por cada canal trasero. Esta recomendación, válida para salas de mezcla de sonido para cine no es adecuada en producciones musicales o para video y televisión (por las condiciones de escucha hogareñas); por esto la recomendación es que cada canal produzca un nivel de presión sonora de unos 79 dBC (siempre en la posición de escucha) para todos los canales de audio,

Debido a que el nivel de escucha de referencia utilizado puede dramáticamente afectar el balance e inteligibilidad de la mezcla, es importante considerar tanto el nivel al cual el programa es mezclado como así también el nivel de escucha típico en el hogar para el mismo programa (Dolby, 2000)

Ahora bien, en salas de mezcla pequeñas y para un mejor balance entre canales frontales y traseros, se recomienda atenuar unos 2 dB los canales traseros.

Una lectura detallada revela que el sistema de medición al que hace referencia el escrito sigue siendo el vúmetro, igualando el 0 VU a un nivel de -20 dBFS para el ruido rosa en el dominio digital, lo cual es, en ningún momento se hace referencia a un medidor basado en sonoridad.

Bob Katz propuso en el año 2002 un sistema de medición que, acorde a las recomendaciones dadas por Dolby, refleje la adecuación de niveles de

diferentes programas sonoros basados en su propia experiencia como ingeniero de sonido de masterización según a la función del material sonoro y al género musical: el K.System. La propuesta contempla tanto los niveles promedios de la señal como así también su rango dinámico y el headroom necesario para los diferentes tipos de grabaciones, siempre estableciendo un único nivel de presión sonora de monitoreo de referencia, 83 dBC para un ruido rosa emitido por cada canal. Según el tipo de programa o género musical, cada nivel promedio constituye una suerte de “nivel objetivo” a alcanzar en cada caso asociándolo a una parte “forte” de dinámica musical.

Propone tres niveles promedios: -20 dBFS, -14 dBFS y -12 dBFS. El primero para producciones con necesidad de un amplio rango dinámico y de Headroom, el segundo para la media de CD's de música pop, rock y jazz y el tercero para la radio y televisión o música que necesite una mayor compresión dinámica. La particularidad consiste en que el 0 dB del vúmetro digital es adaptativo, ya que refleja el nivel promedio que se quiere alcanzar y no el fondo de escala como en los medidores habituales, por esto hay tres escalas diferentes: K-20 (20 dB a fondo de escala), K-14 (14 dB a fondo de escala) y K-12 (12 dB a fondo de escala). Pero además es sumamente importante comprender que el K-System propone que **siempre** el nivel de presión sonora de monitoreo debe ser de 83 dB por canal. Para evitar tener que recalibrar el sistema si se cambia de escala, Katz sugiere el uso de un control maestro de ganancia con una precisión de ± 1 dB por paso.

Entonces, para calibrar el sistema se emite ruido rosa con un nivel de -20 dBFS por cada canal el cual debe dar una lectura de 0 dB en la escala K-20 y se ajusta la ganancia de cada altavoz de forma que el nivel de presión sonora en la posición de monitoreo sea de 83 dBC por cada canal. Si se desea cambiar la escala a K-14, se debe atenuar la ganancia de todos los altavoces 6 dB con el control maestro; así, cuando el medidor muestre 0 dB en esta escala, el nivel de presión sonora será nuevamente de 83 dB por canal. Para la escala K-12 se debe atenuar la ganancia en 8 dB. En la siguiente figura se muestra la equivalencia entre niveles de presión sonora y escalas.

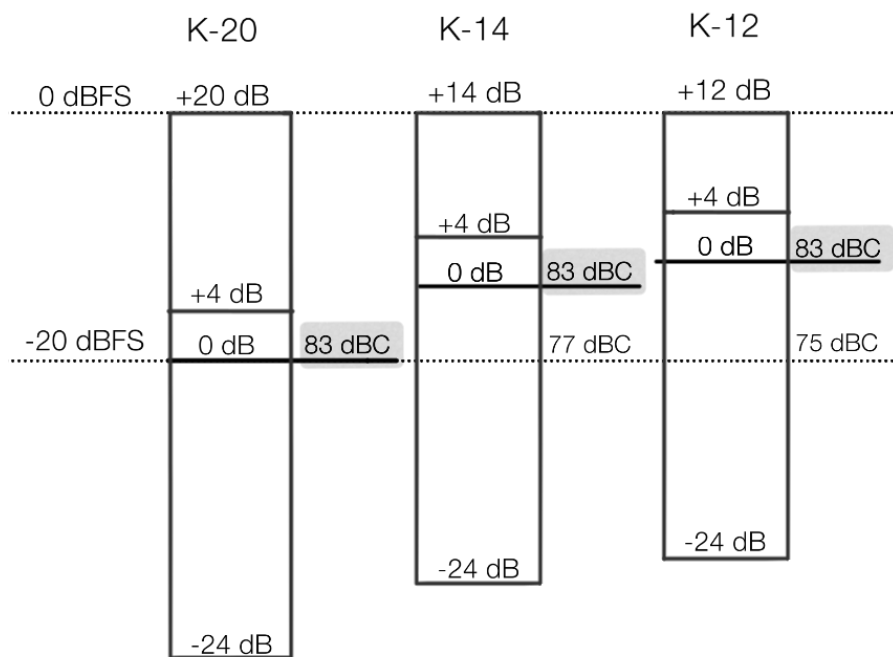


Figura 8. Escalas y niveles en el K-System

En la figura anterior, el margen de +4 dB está destinado a partes del programa sonoro con sonoridad esporádicamente más elevada (típicamente fortissimo); por encima de este nivel, sólo picos de la señal son deseables.

Un medidor basado en este sistema debe constar esencialmente de dos secciones, una que muestre el pico más alto de la señal en los últimos 10 segundos con una lenta caída de 2 segundos cada 24 dB y otra para medir el nivel promedio de los últimos 10 segundos y por esto la medición guarda una relación más estrecha con la sonoridad que el vúmetro tradicional. Según Katz,

El K-System no sólo una escala de medición, es un sistema integrado ligado a la ganancia de los monitores (Katz, 2000)

OTRAS PROPUESTAS DE MEDICION DE LA SONORIDAD

En este trabajo hemos abordado la problemática de la medición de la sonoridad de una señal de audio centrándonos en aquellas soluciones que hayan dado lugar al desarrollo de dispositivos de medición. Aún así, no podemos dejar de lado el hecho que otros trabajos y propuestas han realizado importantes aportes para la medición de la sonoridad.

Una dirección de trabajo fundamental proviene del campo de los psicoacústicos: teniendo en cuenta el sistema de audición humano, el procedimiento de medición divide al espectro sonoro en bandas de frecuencia (típicamente una octava o en tercios de octava) basándose en el concepto de banda crítica, el cual se relaciona con el fenómeno psicoacústico de enmascaramiento. De éstos trabajos se destaca el propuesto por Eberhard Zwicker y que fuera estandarizado en la norma ISO 532-B de 1975. La importancia fundamental del método de Zwicker reside en el hecho de que tiene en cuenta las curvas de igual sonoridad o curvas isofónicas y las características de transmisión del oído externo y medio hacia el oído interno. su principal desventaja radica en que se basa en un procedimiento gráfico (muy tedioso) de estimación de la sonoridad de una señal acústica según el cual, partiendo del trazado de los niveles de presión sonora medidos por tercio de octava, la sonoridad total resulta ser igual al área bajo la curva resultante. Si bien este método (y otras propuestas) dieron lugar al desarrollo de algoritmos y programas de computadora para su aplicación, no han dado lugar (al menos hasta hoy) a implementaciones prácticas en forma de herramientas útiles en el campo de la producción y postproducción de audio digital.

CONCLUSIONES

El uso de herramientas de medición basadas en la percepción de la sonoridad resultan mas adecuadas para la estimación del nivel de una señal de audio antes que el uso de vúmetros y medidores de pico tradicionales.

Todas las metodologías propuestas dan por resultado señales de audio con niveles promedios más adecuados y con un mayor rango dinámico al que se obtenía en la época de la llamada “guerra del volumen”.

Señales con niveles exagerados de RMS y sobrecomprimidas suenan muy degradadas cuando se aplican procedimientos de normalización por sonoridad frente a señales con niveles promedios mas bajos y un mayor Headroom

La normalización por sonoridad utilizada para transmisiones radiofónicas,

televisivas y por streaming tiene un impacto directo en la forma de producción y postproducción de audio.

Para la producción y postproducción de audio se deberá indefectiblemente tener en cuenta, y compatibilizar, los diferentes métodos de medición de sonoridad, tanto por los nuevos paradigmas de escucha y transmisión como para la calibración óptima del sistema de monitoreo.

El desarrollo de las herramientas de medición en formato de programas de computadora o plugins permite una fácil implementación de estas técnicas sin una inversión económica importante.

BIBLIOGRAFÍA

ITU-R BS-1770-2. (2011). Algoritmo para medir niveles de sonoridad de programas de audio y de picos verdaderos. International Communication Union. Génova: International Communication Union.

Katz, B. (2000). An Integrated Approach to Metering, Monitoring, and Levelling Practices. *Journal of the Audio Society of America*, 48 (9), 800 - 809.

Katz, B. (2002). *Mastering Audio: The Art and the Science*. Burlington, MA, USA: Focal Press.

EBU R128. (2014). Loudness Normalisation And Permitted Maximum Level Of Audio Signals. Génova: European Broadcasting Union.

Skovenborg, E. y. (2009). Loudness Descriptors to Characterize Wide Loudness-Range Material. New York: Audio Engineering Society .

Fridrich, R. (s.f.). Notes on SQ Topics: Fridrich Acoustical Consulting. Recuperado el Febrero de 2016, de Fridrich Acoustical Consulting: <http://fridrichacoustical.com/media/5aab44bbc747b14ffff963cffffe417.pdf>

Music Loudness Alliance (2012) Loudness Normalization: The Future of File-Based Playback. Recuperado en Febrero de 2016, de Music Loudness

Alliance: <http://music-loudness.com/index.php/white-papers>

EBU Tech 3341. (2016). Loudness Metering: 'Ebu Mode' Metering To Supplement Ebu R 128 Loudness Normalization. Génova: European Broadcasting Union.

EBU Tech 3342. (2016). Loudness Range: A Measure To Supplement Ebu R 128 Loudness Normalization. Génova: European Broadcasting Union.

EBU Tech 3343. (2016). Guidelines For Production Of Programmes In Accordance With Ebu R 128. Génova: European Broadcasting Union.

EBU Tech 3344. (2011). Practical guidelines for distribution systems in accordance with EBU R 128. Génova: European Broadcasting Union.