

## **Los principales procesos de Masterización de Audio. Masterización para la Restauración de archivos en soporte Analógico.**

**Por: Daniel Schachter<sup>1</sup>**

### **Introducción**

En un artículo presentado como parte integrante del primer informe de Avance correspondiente a este proyecto, titulado *El Circuito de Restauración de Patrimonio Sonoro. Las técnicas de Masterización en la recuperación de antiguas grabaciones* hemos hecho referencia a las diferentes etapas de lo que denominamos *Circuito de Restauración* considerando las diversas convenciones internacionales sobre preservación del patrimonio sonoro y audiovisual, buscando un ordenamiento de pasos necesario para cumplir esos objetivos. El Tercer Paso del mencionado circuito hace referencia a la Restauración de Audio, involucra el estudio de todas las técnicas habituales de Masterización de Audio Digital que resultan adecuadas para su aplicación a la recuperación de grabaciones históricas en soporte de vinilo, originalmente analógicas, a partir de la Migración de las mismas a la plataforma digital y han sido probadas en las pruebas piloto sobre algunos ejemplares de la colección de registros del sello EDUL por lo que en su conjunto constituyen la metodología de trabajo para la restauración de registros históricos de sonido sobre soportes analógicos de vinilo.

A lo largo del presente artículo se describirán esas técnicas y cuando corresponda se indicarán las diferencias entre el trabajo sobre grabaciones originales que desde un primer momento se encuentran almacenadas en un soporte digital, y la Restauración de antiguas grabaciones en soporte analógico a partir de su digitalización.

---

<sup>1</sup> Compositor, Artista Sonoro, Profesor e Investigador, Director del Centro de Estudios en Producción Sonora y Audiovisual (CEPSA) del Departamento de Humanidades y Artes, Universidad Nacional de Lanús, Buenos Aires, Argentina (UNLa) [daniel@schachter.com.ar](mailto:daniel@schachter.com.ar) <http://www.schachter.com.ar>

Con respecto a las grabaciones en soporte analógico y a su etapa de Migración, corresponde hacer referencia al artículo de Pablo Miechi *Registro y reproducción de discos de Vinilo. La etapa analógica previa a su migración al dominio digital*, que es parte integrante del primer informe de avance de este proyecto donde describe minuciosamente las características de las grabaciones analógicas en soporte de vinilo y el procedimiento para su digitalización.

### **La Masterización de audio en el dominio digital**

Cuando hablamos de Masterización de Audio Digital, hablamos de todas aquellas tareas posteriores a la producción de la toma original de sonido, con el objetivo de optimizar las mismas para su posterior inclusión en una producción discográfica, es decir aptas para la duplicación y distribución en el mercado. En líneas generales, todas las tareas posteriores a la primera etapa de producción se conocen como Postproducción, pero específicamente en el caso de las tomas de sonido, se ha reservado esta denominación para todas aquellas rutinas o tareas cuyo destino final sea la creación de bandas de sonido para la industria cinematográfica o el video en general, en cuyo caso se habla de Postproducción de Sonido para la Imagen. Por lo tanto, Masterización será la disciplina destinada a la edición y transformación de tomas de sonido con destino a la reproducción de audio, generalmente a las grabaciones de música. Dice al respecto Katz:

“La masterización es el último paso creativo en el proceso de producción del audio... la última oportunidad para realzar el sonido o arreglar un problema en una habitación acústicamente diseñada, un microscopio de audio... en ocasiones podremos ayudar a trabajar en esa canción problemática... o añadir el toque final que hace que un sonido esté acabado y sea reproducible en diferentes sistemas.” (Katz, 2002)

En este párrafo, Katz no solo define las características generales de la Masterización de Audio como disciplina sino que además puntualiza que el trabajo de Masterización debe realizarse en un recinto acústicamente diseñado,

al que conocemos como Estudio de Masterización, y que el objetivo es obtener un producto reproducible en diferentes sistemas, es decir universalmente compatible con los sistemas de reproducción de audio existentes.

### **Los diversos roles en la cadena de producción**

Existen tres instancias o roles o para la obtención de un producto final de audio profesional que esté en condiciones de pasar a la etapa de distribución: La grabación, la mezcla y la masterización. Desde hace ya muchos años se acostumbra denominar como Ingeniero de Grabación, Ingeniero de Mezcla e Ingeniero de Masterización a los principales responsables de cada una de estas instancias, mas allá de la existencia o no en todos los países de carreras universitarias de ingeniería con especialización en cada una de estas disciplinas. De hecho el circuito Grabación – Mezcla – Masterización en algunos casos es llevado a cabo por un equipo de profesionales que trabajan en estrecho contacto. Sin embargo, esto no es lo mas habitual, y esta circunstancia obliga a recomendar límites para cada una de estas especificidades.

La responsabilidad del Ingeniero de Grabación incluirá desde la selección de los micrófonos a utilizar, el recinto donde grabar, el sistema de hardware / software a utilizar, la cantidad de canales que utilizará para las tomas de sonido, la definición de la máxima calidad posible en la toma<sup>2</sup> de manera tal que permita la utilización de un valor adecuado de headroom<sup>3</sup> y por ende la obtención de un material tan claro y tonalmente neutro o equilibrado como sea posible. Esto significa que de ninguna manera la etapa de grabación debería incorporar rutinas de compresión dinámica ni corrección del equilibrio espectral.

---

<sup>2</sup> expresada en KHz de velocidad de muestreo y bits de resolución, un buen sistema profesional debería permitir que las tomas de sonido se lleven a cabo a 48 o si es posible 96 KHz y 24 bits, valores que permiten trabajar con un rango dinámico de 144 dB.

<sup>3</sup> En grabación de audio digital se denomina headroom al margen o diferencia expresada en dB entre el valor máximo registrado en la etapa de grabación y el valor máximo admitido por el sistema, que en el dominio digital es siempre 0 dB.

Luego, en la etapa de Mezcla reside la responsabilidad de un reparto adecuado en el espacio estereofónico de las voces<sup>4</sup> que permita una buena percepción textural<sup>5</sup> de la toma de sonido, de lo que dependerá la profundidad perceptual o claridad de planos y por ende el reconocimiento de las diferentes situaciones sonoras y por supuesto el reconocimiento de los timbres de los diferentes instrumentos participantes. Muchas veces se confunde el reparto de las voces o de los instrumentos con el reparto de las frecuencias mismas. Allí reside la delgada línea de separación entre la Mezcla y la Masterización, dado que en caso de usar ecualizadores o compresores para corregir la ubicación espacial de determinadas frecuencias en la etapa de Mezcla se estaría interviniendo excesivamente la misma antes de la etapa de Masterización.

Es interesante mencionar en este punto de que manera aparece en escena la separación de roles entre Grabación y Masterización. Hasta mediados del Siglo XX los roles del Ingeniero de Grabación incluían la Mezcla y como relata Owsinski no existía el rol del Ingeniero de Masterización como una disciplina en sí misma. Escribe este autor:

*“Hasta 1948, no había distinción entre los ingenieros de audio, porque todo se grababa directamente sobre vinilo (registros de 10 pulgadas en 78 RPM). En 1948... Ampex introdujo su primer grabador comercial. A partir de allí la grabación pasó a hacerse sobre cinta magnética, lo que hacía necesaria una transferencia a un Master de vinilo para entregar a la planta de prensado, ahí nació la primera encarnación del Ingeniero de Masterización. En 1955, Ampex lanzó su sistema Sel-Sync (Sincronía Selectiva) que dio a su grabadora multipista la capacidad de sobregabar. Con esto la industria cambió para siempre, y se instaló la verdadera distinción entre los Ingenieros de Grabación y Masterización pues a partir de allí los trabajos de ambos difieren mucho” (Owsinski, 2008).*

---

<sup>4</sup> Las voces de una polifonía en términos musicales, que en muchos casos corresponde a diferentes instrumentos.

<sup>5</sup> En música en general y particularmente en Sonido, hablamos de textura para referirnos a la percepción de diversas capas o planos que suenan simultáneamente.

Este estado de cosas se consolidó aún más a partir de 1957 cuando comenzó a comercializarse el disco de vinilo de larga duración. Señala Owsinski que desde ese momento el Ingeniero de Masterización se hizo más influyente debido al uso juicioso y creativo de la ecualización y la compresión para hacer que los discos suenen mejor que cuando fueron grabados. Con la introducción del CD en 1982 la Masterización ingresó en la era digital si bien en un principio las herramientas fueron las mismas, poco a poco se fueron adaptando al dominio digital. (Owsinski, 2008)

### **El Estudio de Masterización. Condiciones necesarias para el monitoreo.**

Las características del lugar donde se lleva a cabo la Masterización están en todos los casos relacionadas con la calidad del producto final. El Estudio de Masterización debe estar en lo posible completamente insonorizado y bien aislado de todo ruido que provenga del exterior. Además, debe estar acondicionado acústicamente para no favorecer ni enmascarar ningún registro de frecuencias. Cuando el Estudio sea al mismo tiempo de Grabación y Masterización, la Sala de Grabación debe estar acústicamente aislada de la Sala de Control y esta última será asimismo la de Masterización. El monitoreo de audio llevado a cabo en la Sala de Masterización debe realizarse con monitores de campo cercano de características neutras, es decir de respuesta plana<sup>6</sup> a todas las frecuencias. En esta sala se encuentran asimismo todos los dispositivos que se utilizarán durante un trabajo de Masterización: la consola o mesa de mezcla, la computadora y su placa de audio, incluyendo todo el software necesario, si hubiese racks de efectos en hardware también estarán aquí. Finalmente, auriculares para la escucha detallada del audio. Resulta generalmente útil escuchar al menos una vez con auriculares el resultado final de la masterización.

### **Análisis y optimización de la dinámica a niveles de pico de amplitud**

El estudio de los valores de dinámica de los archivos de audio y su evaluación

---

<sup>6</sup> Hablamos de respuesta plana cuando es igual en todas las frecuencias de acuerdo a nuestro sistema auditivo.

y tratamiento constituye una de las tareas más importantes de un proceso de Masterización digital. Es necesario poner de relieve que no existen valores uniformes o generalmente aceptados de amplitud, razón por la cual es necesario tomar en consideración diversos criterios a un mismo tiempo, como ser por ejemplo valores utilizados para diversos tipos de soportes analógicos y digitales. La optimización dinámica es uno de los aspectos fundamentales en un trabajo de Masterización, imprescindible antes de terminar el trabajo para recuperar el rango dinámico disponible y alejarnos así del piso de ruido del sistema, pero para poderlo abordar es necesario primero comprender algunas cuestiones referidas al análisis del rango dinámico de una toma de sonido y a la percepción de nuestro aparato auditivo.

### **Terminología, principales variables, conceptos básicos y rutinas.**

La relación entre valores máximos y mínimos de amplitud se conoce como *Rango Dinámico* y la optimización de estos valores en una toma de sonido se denomina *Normalización*. Para comprender como llevar a cabo el análisis de las relaciones de dinámica en un sistema de audio, debemos en primer lugar conocer como percibimos las variaciones de dinámica. Nuestro sistema auditivo no está capacitado para escuchar sonidos de cualquier intensidad. Por debajo de un nivel dado no escuchamos sonido, y por encima de un cierto valor, éste nos produce sensación de dolor. Así podemos definir un rango entre un valor inferior en el cual el sonido empieza a ser audible llamado Umbral de Audición y otro superior que (si es superado por cierto tiempo sostenido) nos provocará daños irreversibles en nuestro sistema auditivo: el Umbral de dolor. Como la intensidad de un sonido está directamente relacionado con las variaciones de la presión respecto a un valor medio de 1000 hectopascuales (la presión atmosférica), fluctuaciones de unos  $20 \times 10^{-6}$  Pa corresponden al umbral de audición, mientras que las de unos 20 Pa corresponden al umbral de dolor. Dada la amplitud del rango, se definió una escala de medición más acotada llamada Nivel de Presión Sonora (NPS o SPL en inglés) según la siguiente relación:

$$NPS = 10 \log_{10} \frac{P}{P_0}$$

donde P es la Presión Sonora y  $P_0$  un valor de referencia que corresponde al valor correspondiente del Umbral de Audición ( $20 \times 10^{-6}$  Pa). La unidad del NPS es el decibel (dB) y el rango audible resulta un valor a cuyo mínimo denominamos 0 dB y cuyo máximo está alrededor de los 120 dB. Así, si nos expresamos en términos de Niveles de Presión Sonora o NPS llamamos 0 dB al valor de umbral perceptivo de amplitud, que es por lo tanto el valor mínimo mensurable o valor de referencia, y cima o umbral del dolor al valor máximo que puede soportar nuestro oído, el que a efectos de este trabajo ubicaremos en los 120 dB. Esto significa que en teoría nuestro rango dinámico perceptivo máximo sería de 120 dB. Sin embargo, el valor que expresa el NPS es un valor promedio y no un valor máximo instantáneo (o valor de Pico). Un sonido con una amplitud muy importante durante un corto período de tiempo puede sonar más débil que otro con menor amplitud pero sostenida. El NPS no es independiente de la evolución temporal de la dinámica: el *Perfil Dinámico*.

Además solo podemos escuchar esas variaciones de dinámica en las frecuencias medias alrededor de los 1000 Hz. En frecuencias graves y agudas resulta notablemente reducido por las características propias de nuestro sistema de audición, tema al que nos referiremos mas adelante. El NPS no es independiente de la composición espectral del sonido<sup>7</sup>. Evolución en el tiempo y composición espectral son condicionantes fundamentales de la forma en que percibimos la intensidad de un sonido. Lo propio sucederá con un señal de audio resultante del registro sonoro y, como veremos, el procesamiento dinámico dependerá y afectará ambos aspectos: tiempo, espectro y dinámica no son fenómenos independientes. Como sabemos, una señal de audio analógica refleja las fluctuaciones de una variable (típicamente la tensión eléctrica) que guarda con el sonido que representa una relación de semejanza pero no de igualdad; presión y tensión eléctrica son variables que pertenecen a

---

<sup>7</sup> Para compensar este fenómeno, los sonómetros incorporan curvas de ponderación según el tipo de sonido del que se quiere medir su NPS

universos diferentes y si una evoluciona de una forma semejante a la otra, la relación no es de identidad sino de analogía, y de allí su denominación.

Si tomamos en cuenta que ese valor de 120 dB es el Rango Dinámico de nuestro sistema auditivo y consideramos que una medición razonable del nivel de ruido presente en un recinto no insonorizado está alrededor de los 20 a 25 dB, entonces el rango dinámico teórico máximo de nuestro sistema se ubicará en el orden de los 95 a 100 dB y generalmente se considera que se encuentra en los 96 dB. Pero la merma ya registrada en tales condiciones usualmente se reduce otros 25 dB por efecto del ruido presente en la actividad cotidiana. No vivimos en una habitación insonorizada sino que por el contrario estamos rodeados por ruidos habituales como el rumor del viento, el sonido que proviene del sistema de iluminación de nuestro lugar de trabajo y muchos otros de ese tipo, sin olvidar que nuestro propio sistema circulatorio aporta su propio rumor de fondo.

Así podemos considerar que ese piso de ruido de la actividad cotidiana constituye una merma de cuanto menos otros 25 dB, lo que reduce el Rango Dinámico perceptible de aproximadamente 96 dB a aproximadamente 70 o 71 dB. Precisamente el Rango Dinámico máximo de un sistema analógico se ubica alrededor de los 70 a 72 dB y se considera que un disco de vinilo tiene un Rango Dinámico máximo cercano a esos valores. La normativa estándar para la producción de discos compactos de audio<sup>8</sup> define un valor de 44,1 HKz de velocidad de muestreo y una resolución de 16 bits, lo que arroja un Rango Dinámico máximo de 96 dB. Estos valores que en principio parecen óptimos dado que consideran el rango teóricamente perceptible por nuestro oído, no lo son tanto a la hora de editar audio en el dominio digital, como veremos mas adelante.

---

<sup>8</sup> Conocidas como Libro Rojo e incluidas en las Normas ISO 9660



En un sistema electroacústico<sup>9</sup> basado en la amplificación de la señal es imposible tomar como valor de referencia el valor mínimo o de umbral dado que la característica misma del sistema hace que los valores mínimos queden enmascarados por el ruido del sistema. Un muy buen sistema analógico siempre contará con un valor mínimo de amplitud o piso de ruido, y soportará una cierta sobrecarga por encima de los valores considerados óptimos. Así, el mejor sistema de amplificación será el que provea la mejor relación entre el nivel de la señal y el nivel del ruido, conocida como relación señal-ruido y el valor de referencia en lugar de ser el umbral de percepción de nuestro sistema auditivo, pasará a ser un nivel óptimo donde la señal que entra al sistema sea igual a la que sale. Por lo tanto en un sistema electroacústico de medición nuestra referencia o valor de 0 dB será igual a 1 o 100% de la señal de entrada, y a partir de este valor podremos medir fundamentalmente la atenuación en dB hasta el punto en el que el ruido de fondo nos impida seguir percibiendo el sonido.

Durante la grabación de las tomas de sonido que luego serán objeto de Masterización es fundamental asegurar que no se producirá distorsión por saturación de amplitud, no solo en los niveles correspondientes a cada entrada en el sistema sino fundamentalmente cuando se lleva a cabo la sumatoria de amplitudes en la salida de la mezcla de audio. Para ello se suele grabar a un nivel no excesivamente alto y se aplica en la salida una reducción en decibeles que será proporcionalmente mayor cuanto mayor sea la resolución en bits del sistema. Ese margen existente entre el valor máximo de amplitud de un archivo de audio y el valor máximo antes de distorsión admitido por el sistema es conocido como *Headroom* y en la etapa de Masterización se recupera mediante la Normalización del archivo en cuestión a un valor máximo de amplitud cuyo valor será tan cercano al óptimo máximo admitido por el audio digital<sup>10</sup> como sea posible y estará relacionado con el valor medio de audición del cual hablaremos más adelante. Si bien el valor máximo posible es igual a 0 dB, nuestro sistema auditivo no percibirá merma alguna si el mismo no llega a

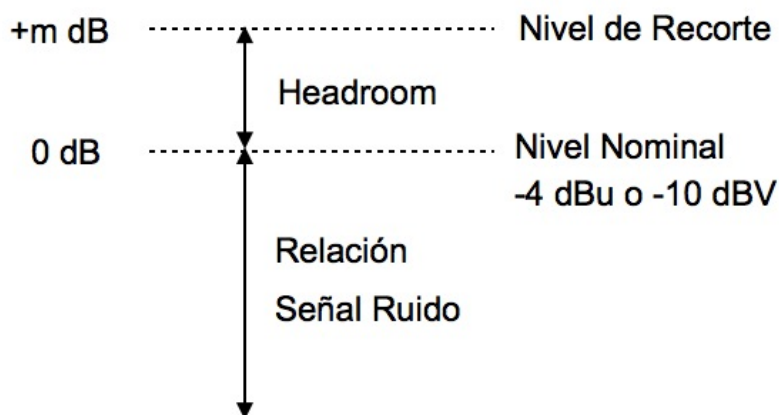
---

<sup>9</sup> Tanto se trate de un sistema analógico como digital, en todos los casos tanto la lectura de los archivos de sonido como su edición y masterización se llevan a cabo a través de la cadena electroacústica.

<sup>10</sup> El valor máximo admitido por el audio digital es igual a 0 dB.

ese valor sino a uno cercano. Podemos afirmar que a partir de 0 dB podemos atenuar ese valor hasta aproximadamente 2 dB sin que deje de ser percibido como óptimo por nuestro oído.

Todo equipo de audio tiene un rango máximo de niveles para el que fue diseñado para trabajar: un valor máximo de tensión de entrada sin que la señal se recorte (o sature): el *Nivel de Recorte*, y un *Nivel de Ruido* generado por el propio dispositivo. Para controlar que el nivel de señal sea el óptimo, los equipos se diseñan con un *Valor de Tensión Nominal* lo suficientemente lejos del *Nivel de Recorte* para evitar que los picos de señal lo superen, y lo suficientemente alto para alejarlo del *Nivel de Ruido*. La diferencia existente entre el Nivel de Recorte y el Nivel Nominal determina el *Headroom* del equipo de audio, que es una zona de seguridad del sistema que indica en cuanto un pico de señal puede superar el *Valor Nominal* sin saturación y la diferencia entre este valor y el *Nivel de Ruido* determinará la *Relación Señal - Ruido* del sistema, como se aprecia en la Fig.1. Históricamente, los equipos se diseñan con dos valores nominales típicos: +4 dBu para equipos profesionales y -10 dBV para semi-profesionales o domésticos.



**Fig. 1**

Así podemos comprender al *Headroom* como una zona de seguridad para asegurar la integridad de la señal y del equipo. En una sesión de grabación multipista habitualmente se utiliza el *Headroom* para evitar el riesgo de saturación por suma de canales en la salida. El *Nivel Nominal* puede ser superado pero se debe tener presente que esto reducirá el rango disponible

para los picos: el control del *Factor de Cresta* se impone en estos casos. A efectos de este trabajo denominamos *Headroom* a la diferencia entre el Nivel Nominal y el Nivel de Recorte o saturación<sup>11</sup>.

Por último, un párrafo aparte merece la noción de Rango Dinámico. Por definición es la diferencia entre el valor máximo y el mínimo de una señal. Aplicado al sonido, esta es la diferencia entre el nivel de NPS más intenso y el más débil, pero aplicado a un sistema electroacústico hay dos variantes posibles: a) diferencia entre el *Nivel Nominal* y el *Nivel de Ruido* (dos valores de RMS) o b) diferencia entre el *Nivel de Recorte* y el *Nivel de Ruido* (un valor de pico y otro de RMS). El nivel de la señal se expresa en términos de valor de amplitud de Pico<sup>12</sup> y valor medio eficaz o RMS<sup>13</sup>

El Valor de Pico marca un límite que de ser sobrepasado producirá distorsión, y el valor medio RMS está relacionado con el nivel que percibe nuestro oído. En el dominio analógico, la amplitud surge de relacionar valores de tensión, y el Decibel es especialmente útil para estas mediciones dado que nuestro sistema de percepción auditiva responde en modo logarítmico<sup>14</sup>. Para definir una escala en dB debemos ante todo adoptar un valor de referencia o valor 0 dB que indicará que el valor de tensión coincide con el de referencia, y a partir de allí si ese valor es atenuado los decibeles serán negativos y si es mayor serán positivos. Para la medición del nivel de línea de una señal de audio<sup>15</sup> fueron definidos dos valores de referencia:  $0,775 V_{RMS}$  y  $1 V_{RMS}$ , el primero da lugar a la escala de dBu y el segundo a la de dBV. Así tenemos que

---

<sup>11</sup> En algunos trabajos se denomina *Headroom* al *Factor de Cresta* cuando solo se considera la señal y no se toma en cuenta la respuesta del hardware, no consideramos adecuada esa denominación porque confunde ambos conceptos.

<sup>12</sup> valor instantáneo máximo de la tensión.

<sup>13</sup> cuyo cálculo veremos un poco mas adelante.

<sup>14</sup> Al igual que el NPS (en decibeles) como unidad de medida del nivel dinámico de un sonido está mejor relacionado con nuestra percepción que la Presión Sonora (en Pascales)

<sup>15</sup> Que es el nivel en que nos centraremos en este trabajo

$$dBu = 20 \log_{10} \frac{V_{RMS}}{0,775V}$$

$$dBV = 20 \log_{10} \frac{V_{RMS}}{1V}$$

Por lo tanto, 0 dBu será el nivel de una señal con  $V_{RMS}=0,775 V$  mientras que 0 dBV corresponderá a una con  $V_{RMS}=1 V$

Claro que ambas unidades están relacionadas entre sí: una señal con  $V_{RMS}=1V$  tendrá un nivel de 0 dBV pero un nivel de 2,2 dBu

Se define como *Factor de Cresta*, a la la diferencia en dB entre el Valor de Pico y el de media RMS de una misma señal<sup>16</sup>. El *Factor de Cresta* junto con el *Rango Dinámico*, *valor de RMS* de la señal de audio serán las variables fundamentales a tener en cuenta para el procesamiento de la dinámica en Masterización.

### **Optimización de la dinámica – Normalización**

La Normalización a nivel de Pico es una rutina habitual imprescindible antes de terminar un trabajo de Masterización, y consiste en un ajuste proporcional de la amplitud de cada una de las muestras que componen la toma de audio digital. Para llevar a cabo esta rutina se individualiza la muestra con mayor amplitud, se calcula cuantos dB es posible elevar esa amplitud para llegar al valor deseado, y se aplica el mismo ajuste en dB a la totalidad de las muestras. Por lo tanto, este ajuste es estrictamente proporcional y por lo tanto no modifica las relaciones internas de dinámica de esa toma de audio. Si el nivel promedio de amplitud se encuentra por ejemplo 20 dB por debajo del pico máximo, podremos comprobar que una vez llevada a cabo la Normalización a Pico esa diferencia se mantiene.

---

<sup>16</sup> En algunos trabajos se lo menciona como relación Pico-RMS implícita.

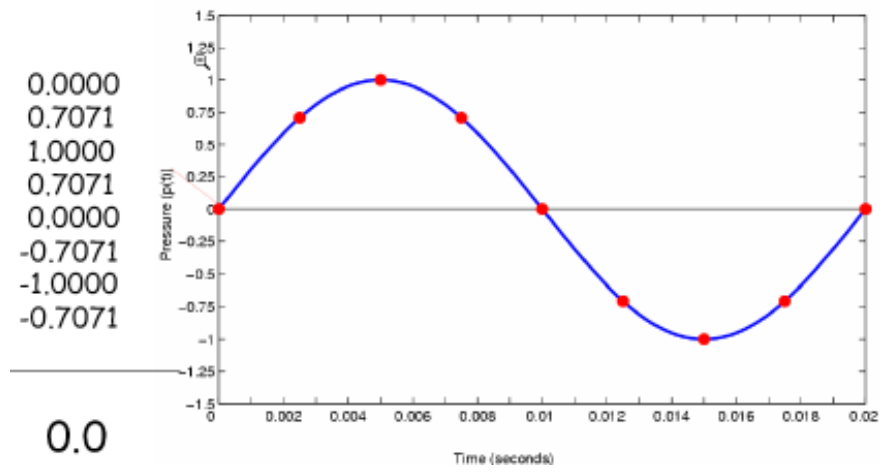
Cuanto mayor sea el rango dinámico que separe al valor de amplitud máximo del valor de amplitud del piso de ruido, diremos que la relación señal-ruido del sistema es mayor. Si consideramos que luego de masterizar las tomas de sonido, las mismas estarán disponibles para ser escuchadas tanto en sistemas de calidad profesional como en otros de audio hogareño, portátiles o semi-profesionales, quedará claro que cuanto mas alto sea su valor de Pico y por tanto más cercano al valor máximo posible, tanto mas alto será también el valor medio o de escucha, por lo tanto será posible reproducir el audio en cuestión en todo tipo de reproductores.

### **Análisis y optimización de la dinámica a niveles de media RMS**

Tal como quedó expresado, nuestro oído el valor medio y no el valor máximo de amplitud. Si decidimos alejar del piso de ruido aquella información que no queremos perder, aún cuando no modifiquemos significativamente el valor máximo o de pico, para hacer esto deberemos modificar la dinámica interna de la toma de sonido. Para comprender esto es necesario ante todo comprender como se calcula ese valor de amplitud promedio.

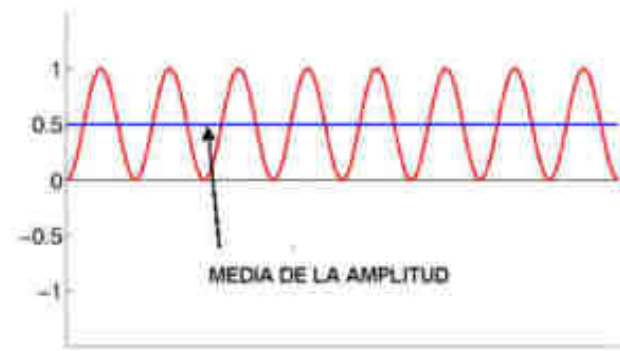
Teniendo en cuenta que el sonido presenta fases positivas y negativas, el valor promedio de la amplitud de todas las muestras arrojará un valor nulo dado que las muestras positivas se anularán con las negativas. En realidad, nuestro oído no percibe el sonido de esa forma. Si invertimos las fases comprobamos que el sonido no se modifica en lo perceptivo, sino que a nuestros oídos suena igual la fase positiva que la negativa de una forma de onda. Por eso, para calcular el valor medio de la amplitud de una toma de sonido debemos recurrir a un valor que surge del cálculo conocido como Root Mean Square o Raíz Media Cuadrática, que se abrevia como RMS.

Veamos el siguiente gráfico de la Fig.1 podemos apreciar como en una onda sinusoidal los valores positivos y negativos se anulan.



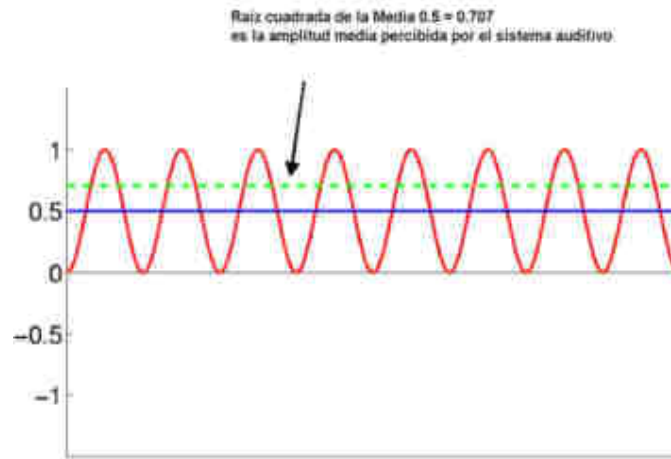
**Fig.1**

Para evitar este inconveniente podemos eliminar el signo elevando al cuadrado todas las lecturas de amplitud y convirtiendo así todas las amplitudes en valores absolutos, de forma que podamos entonces promediarlas según muestra la Fig.2



**Fig.2**

Dado que hemos elevado todos los valores al cuadrado, aplicándoles raíz cuadrada obtendremos un valor denominado Raíz Media Cuadrática o Root Mean Square en idioma inglés, cuyas iniciales son RMS que representa la amplitud media percibida por el oído humano y puede verse en la Fig.3



**Fig.3**

En un proceso de Normalización a un valor de Pico hemos visto que el ajuste será proporcional, por lo tanto la amplitud de todas las muestras será ajustada en la misma cantidad de dB que la muestra de mayor amplitud. Así estaremos seguros de no vulnerar el valor máximo de amplitud dado que tomamos como patrón precisamente la muestra de mayor amplitud para determinar cuantos dB serán aplicados a todas las muestras de la toma de audio.

En cambio si atendemos el valor de media RMS y decidimos incrementarlo por ejemplo 10 dB en principio no tendríamos control alguno sobre el pico de amplitud y podríamos saturar y por lo tanto distorsionar ese archivo de audio. Para evitar esto, tanto en el dominio analógico como en el digital existen los compresores de dinámica<sup>17</sup> que literalmente *suben el piso y bajan el pico máximo* del sonido. Una vez modificados los valores máximo y mínimo, los compresores de dinámica se encargan de optimizar la amplitud de la toma respecto de la relación señal-ruido (como vimos antes), es decir de llevar el pico al máximo posible y ajustar las demás muestras en forma proporcional. Así, en definitiva, el piso o valor mínimo es el que se eleva y las tomas de sonido se perciben con mayor amplitud.

<sup>17</sup> En este caso nos referimos a un Compresor incluido dentro de una herramienta de Normalización. La Compresión dinámica en sí misma y como recurso independiente, es una de las principales herramientas de Masterización y mas adelante será abordada como tal.

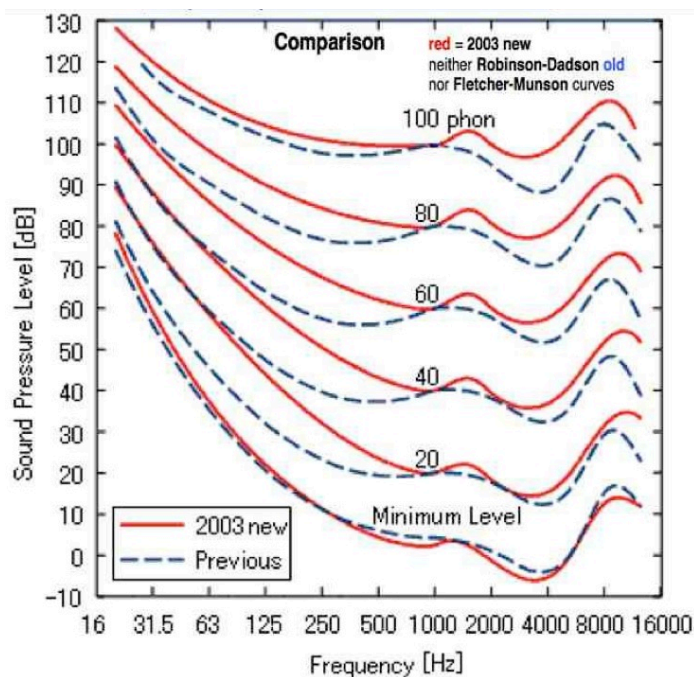
No abundaremos en detalles acerca de la Compresión Dinámica en este punto dado que mas adelante tendrá el tratamiento especial que merece, solo diremos que toda vez que se comprime la dinámica de una toma de sonido, el resultado obtenido es la elevación del valor de media RMS, y obviamente cuando se eleva la media de RMS, se comprime la dinámica de las tomas de sonido. Compresión dinámica y ajuste del valor de media RMS son simultáneos pero si nuestra intención es comprimir, podremos precisar los parámetros de compresión pero no tendremos control absoluto sobre el valor de RMS resultante, precisamente porque se trata en ese caso de el resultado del proceso. En cambio si aplicamos una Normalización a un valor dado de RMS, obtendremos un archivo de audio comprimido pero no tendremos control sobre los parámetros del compresor.

Los valores de RMS generalmente usados varían por cuestiones de estilo y cada empresa productora o sello discográfico. Generalmente rock, pop, jazz, tango y otros estilos de música popular suelen usar valores mas altos de RMS Para la voz suele usarse valores un poco inferiores. La música orquestal, para instrumentos solistas o pequeños grupos de instrumentos acústicos y la música electroacústica utilizan mucho mas los contrastes dinámicos y por lo tanto presentan habitualmente valores mas bajos de RMS.

El programa *Sound Forge* de Sony ha sido el primer software en incluir en su herramienta de Normalización una opción para ajustar los valores de RMS a su valor compensado por las curvas de sonoridad por lo que directamente coloca a la media RMS en valores perceptivos. La herramienta de este programa es especialmente eficiente para medir la media RMS compensada por sonoridad, especialmente porque incluye un compresor de dinámica automático muy eficiente. La herramienta de Normalización de este programa permite trabajar sobre la dinámica dejando por cuenta del software la aplicación automática de la Compresión Dinámica necesaria para alcanzar como resultado un valor buscado de RMS. Sin embargo, esos valores ajustados por sonoridad que proporciona *Sound Forge* son un caso especial y no resultan comparables con los que entregan otros programas que siempre trabajan sobre cálculo estadístico de la media RMS. En la Fig.4 se aprecian las curvas de sonoridad



igual o isofónicas que muestran de que forma responde nuestro sistema auditivo a la percepción de diferentes frecuencias.



**Fig.4**

Las *Curvas de Sonoridad Igual*, también llamadas *Curvas Isofónicas* no son en sí mismas una herramienta de Masterización. Para mayor información sobre este tema remito a la lectura del artículo de Gabriel Data y Pablo Miechi *Procesamiento Espectral y Dinámico en la Masterización de Audio* y al artículo de Daniel Schachter *El Circuito de Restauración de Patrimonio Sonoro. Las técnicas de Masterización en la recuperación de antiguas grabaciones* ambos incluidos también en el Informe Final de este mismo Proyecto.

### **Conversión a estéreo de grabaciones originalmente monoaurales**

En toda grabación estereofónica nos encontramos ante el registro de un discurso idéntico por dos medios diferentes, de forma que tanto las ondas directas como sus reflexiones llegan cada a los diferentes receptores (sean éstos nuestros oídos, un par de micrófonos o un micrófono estereofónico) por caminos levemente diferentes. Eso produce en nuestra percepción la sensación de espacio a la que llamamos imagen estéreo.

Con la sola excepción de aquellas grabaciones originalmente monoaurales por expresa decisión de sus autores, podemos convertir en estereofónica toda grabación originalmente mono y acercarnos a un teórico estéreo original mediante la creación de un falso estéreo por contrafase invirtiendo la polaridad de un canal y luego suprimiendo en ese canal una muy pequeña cantidad de muestras hasta que las fases de ambos canales coincidan nuevamente. Este desfasaje temporal de un canal será percibido como una grabación originalmente estereofónica que no se anulará por suma de canales y contrafase. Debemos tener siempre presente que un archivo de audio estereofónico tiene siempre la misma cantidad de muestras en cada canal, por lo tanto deberemos insertar la cantidad necesaria de muestras silenciosas para lograr este objetivo, la que no debe superar el Umbral de Haas<sup>18</sup> para que la señal se perciba como una sola y no se reconozca como proveniente de fuentes diferentes.

De esta forma, la inversión de fase de un canal se compensará con el brevísimo desfasaje temporal que recompone las relaciones de fase entre canales, y obtendremos un archivo dotado de un aceptable campo estereofónico.

### **Analizadores espectrales y filtrado de Ruidos simples y complejos**

Los *Analizadores Espectrales* son una de las herramientas más importantes para la edición y masterización de audio digital. Nos permiten conocer las componentes de una onda compleja, así como individualizar las frecuencias presentes en cualquier fragmento de la misma. De esta manera podemos literalmente ver un fragmento ruidoso y analizar su comportamiento en el tiempo. El *Análisis Espectral* en el dominio digital se realiza habitualmente mediante la Transformada Rápida de Fourier o FFT que divide el archivo de audio o la porción del mismo que estemos analizando en un número dado de

---

<sup>18</sup> Se conoce como Efecto Haas el que se produce cuando aplicamos una señal idéntica en dos altoparlantes y retardamos en uno de ellos la señal. Cuando superamos los 35 milisegundos de retraso el sonido retrasado se percibe en forma separada del primero y ambos provenientes de fuentes diferentes. Así denominamos Umbral de Haas a esos 35 milisegundos máximos de retraso.

ventanas, comenzando usualmente por una ventana de 128 muestras y llegando en algunos casos hasta más de sesenta mil muestras donde siempre la relación entre las ventanas anterior – posterior es 1:2<sup>19</sup>. A partir de allí, cada ventana es representada por una curva de función. La regla general es que cuanto más pequeña sea la ventana de análisis, tanto más detallado será ese análisis pero cuanto más grande sea esa ventana más exacto será el resultado.

Existe otro método denominado *Transformada Digital de Ondículas* o *Wavelets* conocido por la sigla DWT. Este método tiene la particularidad de reconocer y graficar con precisión los transitorios del ataque de la forma de onda y es particularmente débil para graficar la porción determinada del sonido o sea lo que es la fortaleza principal del método FFT. Por esa razón este tipo de filtros se utiliza para el análisis pero no se lo suele aplicar al filtrado de ruido.

Toda vez que percibamos ruido de fondo presente en un registro de audio digital podremos analizarlo y como resultado del análisis reconocer si se trata de un *Ruido Simple* o de un *Ruido Complejo*.

Denominamos *Ruidos Simples* a aquellos constituidos por una frecuencia claramente perceptible que podrá o no estar acompañada por alguna/s otra/s cuyo nivel de amplitud no alcanza el umbral de audibilidad o bien lo alcanzan solo por el hecho de encontrarse moduladas por aquella frecuencia preponderante. Habitualmente esto sucede con los llamados Ruidos de Línea producidos por una descarga a tierra poco eficiente al momento de llevar a cabo la grabación. Estos ruidos tienen una frecuencia de 50 Hz o 60 Hz según sea el voltaje de la instalación eléctrica<sup>20</sup>. Un *Ruido Simple*, aparte de estar generalmente constituido por una única frecuencia, es muy estable en valores de amplitud. Por lo tanto también lo reconoceremos porque los medidores de Amplitud de Pico quedan prácticamente inmóviles indicando la amplitud máxima de este ruido. En un *Ruido Complejo* por el contrario las componentes son siempre muchas y su valor de amplitud es muy variable.

---

<sup>19</sup> Es decir 128, 256, 512, 1024, 2048, 4096, 8192, 16384, 32768, 65536, etc.

<sup>20</sup> Son 60 Hz para instalaciones de 110 voltios a 120 voltios y 50 Hz para instalaciones de 220 voltios a 240 voltios.

Para filtrar un *Ruido Simple* recurrimos a un filtro simple, habitualmente usamos para tal fin un Filtro Paramétrico pasa-rechaza banda con la mayor profundidad de filtrado y el menor ancho de banda posibles<sup>21</sup>. Cuanto menor sea el ancho de banda y mayor sea la profundidad, tanto mas preciso será el filtrado. Existen muchas herramientas de software que proveen este tipo de filtros. En algunos casos podemos encontrar Filtros Paramétricos con representación gráfica, denominados Paragráficos, que en esencia son iguales a los Paramétricos solo que provistos de una interfaz mas amigable para el usuario.

Para filtrar un *Ruido Complejo* utilizaremos los Filtros de Ruido Complejo por FFT. Este tipo de filtros utiliza la técnica de la Transformada Rápida de Fourier descrita mas arriba. Generalmente cuando un software provee por un lado el Analizador y por otro lado el Filtro Complejo, utiliza para ambas tareas el mismo algoritmo de cálculo. Como quedó expresado anteriormente, para cada ventana existe una función y para que el filtrado sea eficiente es imprescindible solapar las ventanas haciendo que una parte de cada ventana sea a su vez analizada en otra ventana. Esto se debe precisamente a las funciones que se aplican para el cálculo de cada ventana. El Solapado (conocido en inglés como Overlapping) es de tal importancia que los programas que incluyen filtros de estas características incluyen siempre el solapamiento de ventanas y definen un valor de overlapping mínimo 2/3 y un máximo de 3/4 de ventana<sup>22</sup>.

El filtro actúa en primer término como un analizador del espectro ruidoso y obtiene como resultado la representación del ruido que deseamos eliminar. El sistema tiene así la información acerca de un espectro sonoro que se encuentra incluido dentro de otro (la toma de sonido). Una vez individualizado el espectro a filtrar podremos entonces determinar cuantos dB de reducción le aplicaremos para que no sea percibido.

---

<sup>21</sup> El tema de la estructura de los filtros es en si mismo muy extenso y justificaría por si solo un proyecto de investigación. Bastará decir aquí que buscaremos el filtro del mayor Orden posible donde los filtros del 1º Orden atenúan 6 dB por octava y cada número creciente de orden significa otros 6 dB por octava (2º Orden: 12 dB/octava, 3º orden: 18 dB/octava y así sucesivamente).

<sup>22</sup> O sea que el solapado u overlapping no es opcional sino un parámetro que siempre está presente y por tanto define las características de un analisis por FFT.

A modo de recapitulación diremos entonces que los principales parámetros que determinarán la eficiencia de un Filtro de Ruido Complejo por FFT son: en primer lugar el tamaño de la ventana, en segundo lugar la proporción del Solapado entre ventanas y en tercer lugar la atenuación del espectro ruidoso expresada en dB.

Existen múltiples opciones en el mercado pero muchas de ellas no nos permiten el manejo de todos los parámetros con la excepción del Filtro por FFT de Sound Forge que nos permite efectivamente manejar esas variables aparte de otros parámetros ya relativos por ejemplo a la envolvente espectral de filtrado, velocidad de ataque y reléase, etc. Funciona muy bien habitualmente si determinamos correctamente la ventana de análisis y el valor del solapado. Con respecto al tamaño de ventana, como primera medida resulta adecuado definir lo que podemos denominar una *unidad mínima de filtrado* compuesta por dos ventanas y un solapado. Esto permitirá la aplicación de al menos un solapamiento de ventanas que es esencial en la técnica de FFT. Luego la regla general será que cuanto mas grande sea la muestra de ruido tanto mayor podrá ser la ventana de filtrado a utilizar. La práctica profesional ha demostrado que resulta eficiente relacionar el tamaño de la ventana con la posibilidad o no de abarcar la totalidad de la mencionada *unidad mínima de filtrado* con la muestra de ruido pura que hayamos aislado para filtrar. Por ejemplo si tenemos una muestra de ruido puro de 3/8 de segundo en una toma de audio digital a 44.1 KHz, tendremos:

$$44100 / 8 \times 3 = 16537 \text{ muestras de ruido puro}$$

Si deseáramos utilizar la ventana mas grande disponible y esta fuese de 16384 muestras (un valor habitual en este tipo de filtros) comprobaríamos que de acuerdo a lo expuesto deberíamos tener una muestra de ruido puro igual o mayor a la *unidad mínima de filtrado* para ese tamaño de ventana, es decir una cantidad de ruido que permita aplicar un filtro por FFT, el que para funcionar correctamente requiere que al menos exista un solapamiento entre ventanas de análisis. Así la mínima cantidad de muestras de ruido será la que sigue:

**16384 x 1,33 = 21791 muestras de ruido puro necesarias para 16384**

Como vemos, a pesar de que el número de muestras de ruido puro supera el tamaño de la ventana que deseamos utilizar, no es suficiente para permitir la existencia de al menos un solapado. Entonces deberemos probar con la ventana inferior siguiente, que en todo filtro de ruido por FFT siempre tiene un tamaño igual a la mitad de la anterior, o sea 8192 muestras.

**8192 x 1,33 = 10895 con lo que podremos usar la ventana de 8192**

Nada nos impide utilizar la ventana mas grande con una cantidad tan pequeña de muestras, salvo los resultados obtenidos, dado que pueden aparecer errores tales como cierta distorsión armónica y la percepción de sobrefiltrado por merma de transitorios del ataque.

Luego de esto debemos definir la cantidad de dB a filtrar para lo cual ante todo debemos medir la amplitud de pico del ruido que queremos filtrar. Como quedó expuesto anteriormente, nuestro sistema de audición percibe un máximo teórico de 96 dB que de hecho se ven reducidos en aproximadamente 25 dB por el ruido ambiente. Así podemos decir que nuestro aparato auditivo percibe unos 70 a 72 dB de rango dinámico, por lo cual para hacer desaparecer dicho ruido de nuestra percepción alcanzará con reducir su amplitud hasta que su valor máximo o pico caiga a los aproximadamente -71 dB. Así un ruido complejo que presenta una amplitud máxima de 40 dB debería desaparecer a nuestros oídos si reducimos su amplitud en 31 dB. Si reducimos menos que esto, quedará algo de ruido perceptible. Por lo tanto podemos decir que este será el valor mínimo de filtrado posible.

Por otro lado, la mejor prueba para el procedimiento de determinación del tamaño de ventana y porcentaje de solapado es intentar reducir el ruido mucho mas que eso, llevando su valor de pico al piso del rango dinámico disponible. Siguiendo con el mismo ejemplo, si estamos trabajando en 16 bits tendremos un rango dinámico máximo de 96 dB, por lo tanto podríamos reducir la amplitud

del ruido en cuestión en 56 dB y aún así no dañar la toma de audio. Si superamos este valor máximo de atenuación de la amplitud del ruido comenzaremos a percibir tanto distorsión como sobrefiltrado por merma de transitorios de ataque. Por lo tanto podemos decir que este sería el valor máximo de filtrado posible expresado en dB.

## **Compresión dinámica en Masterización**

Gabriel Data y Pablo Miechi han analizado en detalle las herramientas para el Procesamiento Dinámico en su artículo *Procesamiento Espectral y Dinámico en la Masterización de Audio* a cuya lectura remito para el estudio detallado de estas cuestiones. Por lo tanto solo me referiré a algunas cuestiones generales referidas al tratamiento de archivos digitales estereofónicos de audio obtenidos como resultado de las etapas de Grabación y Mezcla.

El análisis de una metodología para la Masterización de audio en el dominio digital no puede dejar de considerar el caso en el que las tareas de Grabación y Mezcla se lleven a cabo por separado de la Masterización. En efecto, cabe suponer que en mas de un caso recibiremos mezclas ya terminadas. Precisamente la digitalización de archivos de audio en soportes analógicos de vinilo es uno de esos casos en lo cuales es imposible intervenir en la Mezcla ni opinar acerca del emplazamiento de los micrófonos previo a la toma de sonido. En estos casos nos encontramos ante preguntas como:

¿Será útil en esos casos la aplicación de un compresor de dinámica? ¿Con que criterio aplicaríamos un compresor a una mezcla aparentemente ya terminada?

Propondremos a continuación alternativas para mejorar el aspecto dinámico de un archivo de audio correspondiente a una mezcla aparentemente finalizada aplicándole Compresión.

Sostiene Katz que en muchos casos es posible intervenir la textura de una mezcla de audio mediante el procedimiento denominado Compresión Paralela. Dice este autor que la superposición de una toma de sonido dada con una

copia de sí misma pasada por un compresor, le otorga transparencia textural. Para ello ante todo es necesario definir valores aplicables a los parámetros principales del Compresor que utilizaremos. Katz propone utilizar un valor de umbral indefinido aunque por cierto muy bajo, por ejemplo alrededor de -50 dB, no ajustar la amplitud de salida del Compresor<sup>23</sup>, y manejar el nivel de la razón o ratio de compresión hasta obtener un resultado aceptable. (Katz, 2002)

Propondré un método alternativo que a mi juicio arroja mejores resultados. No resulta convincente la indefinición de Katz en el nivel del umbral del Compresor, mas aún, al ubicarlo en un valor intencionalmente muy bajo aunque indefinido, lo único que se asegura es que una gran parte del audio pase por el proceso. Luego, tampoco define el valor de la ratio. En lugar de esto podemos asumir la tarea de definir un valor dado razonable y sustentable para el umbral del Compresor. En este sentido, luego de diversas pruebas piloto sobre una importante cantidad de archivos de audio hemos apreciado que el valor teórico de RMS que surge a partir de considerar el valor del Factor de Cresta de un archivo de audio para un valor de amplitud de Pico optimizado a 0 dB es en la mayoría de los casos un aceptable umbral de compresión.

Para conocer este valor basta con multiplicar por -1 al valor del *Factor de Cresta*. Por ejemplo, un archivo de audio que presenta un pico de -2 dB y una media RMS de -22 dB tendrá un *Factor de Cresta* de 20 dB. Si normalizásemos su amplitud de Pico a 0 dB su RMS pasaría a ser -20 dB, pero este paso no es necesario, directamente podemos asignar al compresor un umbral de -20 dB.

El trabajo de campo ha demostrado de en la gran mayoría de los casos, el valor de RMS implícita en el Factor de Cresta de un archivo ha funcionado bien como umbral y ha permitido mantener fijos el ratio de compresión en una relación de 2:1 , el ataque en el valor mínimo posible (usualmente 1ms) y el

---

<sup>23</sup> El ajuste de la ganancia en la salida del Compresor se denomina habitualmente Make Up o Output Gain compensation y es la función que efectivamente produce la elevación de la RMS en un proceso de Compresión. Al eliminar el ajuste de salida del Compresor, de hecho la RMS resultante será inferior a la original, Katz llama a ese proceso Compresión Descendente aunque de hecho el resultado es muy similar al que surgiría si aplicásemos un Limitador.



reléase en un valor usualmente estándar como 500 ms. Con estos valores, activando el ajuste automático de salida en el compresor, se ha logrado en todos los casos una elevación del nivel de RMS entre 3 y 4 dB. El ajuste fino de este procedimiento se encuentra precisamente en la ratio de compresión, la que variando entre 2:1 y 3:1 permite definir el valor a obtener como resultado.

Este procedimiento funciona eficientemente en el proceso de *Compresión Paralela* siempre que no utilicemos el ajuste automático de ganancia en la salida del Compresor, pero también ha demostrado ser muy útil como política de Compresión dinámica para aplicar a aquellos archivos de audio que llegan a nuestras manos como hechos consumados productos de grabaciones y/o mezclas en las que no hemos participado.

### **Ecuación espectral en Masterización**

En su artículo *Procesamiento Espectral y Dinámico en la Masterización de Audio* a cuya lectura remito, Gabriel Data y Pablo Miechi han analizado en detalle las herramientas para la corrección espectral de las tomas de sonido, es decir los Ecuadores. Por lo tanto solo me referiré a algunas cuestiones generales referidas al tratamiento de archivos digitales estereofónicos de audio obtenidos como resultado de las etapas de Grabación y Mezcla. Ante todo, tal como sucede con la Compresión Dinámica, el tipo de ajuste y corrección a emplear durante la Masterización difiere de las rutinas utilizadas durante la Grabación y la Mezcla de Audio. Con respecto a esto afirma Owsinski:

“Una de las responsabilidades más importantes del ingeniero de masterización es arreglar el equilibrio de frecuencias de un proyecto (si es necesario). Por supuesto, esto se hace con un ecualizador, pero el tipo usado y la forma de manejarlo son generalmente muy diferentes a las utilizadas durante la grabación o mezcla. Mientras que en la grabación es posible que el uso de ajustes grandes de ecualización (por ejemplo de 3 a 15 dB) a una cierta frecuencia, en la Masterización casi siempre se aplica una gran cantidad de

incrementos muy pequeños, por lo general desde décimas dB a 2 o 3 como mucho.” (Owsinski, 2008)

Por otra parte los sellos discográficos buscan definir ciertas características en la compresión y ecualización aplicadas a toda su producción, mas allá de los requerimientos específicos de cada grabación, con el objetivo de definir un perfil dinámico y un cierto valor de ajuste espectral genéricos que resulten aplicables a todas sus ediciones y resulten distintivos de su marca.

Para aquellos casos en los cuales debemos trabajar sobre mezclas aparentemente finalizadas, o en el caso de la digitalización de originales sobre soporte analógico, es posible definir una estrategia general para intervenir espectralmente los archivos de audio. En principio trataremos de evitar la enfatización de graves por debajo de los 200 Hz o los agudos por sobre 5 KHz. Esos ajustes resultan generalmente demasiado notorios o artificiales. Normalmente resulta mas eficiente trabajar sobre las frecuencias medias.

Habitualmente las mezclas estéreo tienen dos regiones espectrales cuya atenuación suele resultar beneficiosa. Esas regiones están, una alrededor de los 300 Hz y la otra alrededor de los 800 Hz. No es sencillo determinar el valor exacto de las frecuencias a atenuar, pero podemos comenzar por las frecuencias indicadas trabajando con un ancho de banda muy acotado de aproximadamente un tercio de octava o a lo sumo media octava. Los ajustes de ecualización deben ser sutiles y no necesariamente debemos aplicar la misma atenuación a ambas bandas considerando una atenuación máxima de aproximadamente 3 dB.

### **Edición de Audio en Puntos de Cruce Cero y Montaje de audio digital**

Llamamos *Cero Cross* o “OX” al punto exacto en el que la forma de onda atraviesa el eje que determina el paso de la fase negativa a positiva o viceversa de una forma de onda. Toda edición que considere los puntos de OX asegurará la ausencia de *clicks* producidos por caídas abruptas de amplitud cuando la energía de una onda se interrumpe repentinamente. Cuando

aplicamos cortes en OX estos pueden ser de *pendiente positiva* cuando el punto de corte se da al final de la fase negativa, o de *pendiente negativa* cuando el punto de corte se da al final de la fase positiva. Para evitar ruidos digitales (o *clicks*) debemos siempre cuidar que las fases contiguas (o sea el antes y el después) sean congruentes, es decir que continúen el ciclo como si no hubiera existido el corte.

La edición en Puntos de Cruce Cero es imprescindible para resolver los Montajes entre diversas tomas, muy habituales cuando se llevan a cabo las sesiones de grabación para la producción de un disco, como así también para la correcta eliminación de clicks de gran tamaño producidos por errores de cuantización digital. También será de gran utilidad utilizar la técnica del Montaje para eliminar pequeños pozos cuya duración sea inferior al Umbral de Haas (ya mencionado anteriormente) y por lo tanto su eliminación puede pasar desapercibida. En estos casos alcanzará con la eliminación del ciclo completo que incluye el *click* que deseamos eliminar o en el caso de los pequeños pozos la eliminación de las muestras inmediatamente anteriores y posteriores al mismo hasta alcanzar el primer punto de cruce cero con pase de fase positiva o negativa. No me extenderé mas sobre este tema dado que en la remasterización de grabaciones analógicas registradas en discos de vinilo, no es habitual la utilización de técnicas de Montaje.

### **Utilización de la síntesis de sonido como herramienta de masterización**

En aquellos casos en los que la toma de sonido sea espectralmente poco relevante, sea esto por la utilización de micrófonos inadecuados o mal ubicados, y sea imposible volver a realizar la grabación corrigiendo esos problemas, los archivos de audio sonarán opacos y débiles en textura como resultado de la pérdida de algunos parciales, y con ataques menos claros y efectivos. Cuando esto sucede es posible intervenir esos archivos reforzando algunas frecuencias, particularmente el sostén de la envolvente dinámica de algunos sonidos largos o los transitorios del ataque de ciertos instrumentos en pasajes donde esos ataques hayan quedado opacados.

Para llevar esto a cabo es necesario ante todo llevar a cabo un buen análisis espectral por FFT de los pasajes donde se hacen evidentes esos problemas para detectar cuales son las frecuencias que requieren refuerzo. En todos los casos el principio aplicado es el mismo: la suma de amplitudes entre los sonidos débiles existentes y otros de idéntica frecuencia que crearemos por síntesis. Concretamente en el caso de notas largas que deseemos reforzar, una vez analizado el fragmento que las contiene y conocida la frecuencia, podemos entonces analizar sus valores de amplitud de Pico y Media RMS. Con esta información podemos ya crear por síntesis una onda sinusoidal de idéntica frecuencia, con una duración igual a la del fragmento que deseamos reforzar y con una amplitud igual a la media RMS de la onda a enfatizar, darle a nuestra onda sintética una forma dinámica similar a la del sonido que necesitamos reforzar y luego mezclar la onda sintética dentro del fragmento musical original. Si no resultase suficiente podremos volver un paso atrás y llevar la amplitud de la onda sintética al valor de amplitud de Pico de la frecuencia a enfatizar y repetir el procedimiento.

En el caso de los ataques que deseemos reforzar, el procedimiento es similar en cuanto a la determinación de la frecuencia a enfatizar mediante análisis espectral, pero el procedimiento para obtener los transitorios del ataque puede ser diferente dado que la enfatización de transitorios de ataque nos permite usar por ejemplo síntesis por modulación de frecuencia para obtener transitorios del ataque a partir de la frecuencia previamente determinada por el analizador. Tanto este procedimiento como el anterior pueden denominarse como Enfatización por Síntesis.

También podríamos optar por otro camino, siempre partiendo desde el análisis por FFT del fragmento sonoro a reforzar. Una vez determinada la frecuencia a enfatizar, podríamos tomar una muestra de la toma de sonido que deseamos intervenir y aplicarle un filtrado rechaza-banda de la mayor profundidad en dB y el menor ancho de banda posibles, y luego pegar esta muestra sobre el original. Así estaremos sumando amplitud a todas las frecuencias de ese fragmento pero a la que deseamos enfatizar muchos mas dB que a las restantes. Si

disponemos de un buen filtro, esta alternativa a la que podemos denominar como Enfatización por Muestreo puede resultar muy eficiente.

### **Desplazamiento en el Eje de Continua o DC Offset**

Cada vez que procesamos las tomas de sonido durante el proceso de Masterización, las descomponemos en sus componentes primarios, es decir ceros y unos. Cada proceso implica una gran cantidad de decisiones (cálculos) sobre cada uno de esos números y la salida del mismo es el archivo de audio ya procesado. Durante esos cálculos pueden deslizarse errores, y es habitual que el resultado de esas fallas sea el desplazamiento del eje que marca el cambio de fase en el oscilograma. Este tema es analizado en detalle por Alejandro Brianza en su artículo *Desplazamiento del eje de cero absoluto. Concepto, origen, soluciones* que se presenta junto con este artículo y por tanto integra el Informe Final de Investigación y me remito a su lectura. Solo agregaré que este problema es abordado por las diversas herramientas informáticas disponibles, que el control de ese desvío debe ser permanente, sea cual fuere el procesamiento aplicado y que su reparación es imprescindible cada vez que este inconveniente es detectado para evitar la pérdida irreparable de una parte del Rango Dinámico total del que disponemos.

### **Dithering**

EL último paso del proceso de Masterización es la obtención del nuevo Master de audio, y es aquí donde nos encontramos con que hemos trabajado sobre archivos digitales grabados en la mayor calidad posible para reducir al mínimo la degradación de los mismos como resultado de todos los procesos que hemos detallado. Así, estamos en condiciones de entregar como producto final una serie de archivos de audio que se encuentran en nuestro disco duro a 96 o en todo caso a 48 KHz de velocidad de muestreo y 24 o 20 bits de resolución. El formato CD-DA de acuerdo a la normativa vigente<sup>24</sup> para la producción de discos compactos de audio dispone que debemos crear un Master en 44.1 KHz y una resolución de 16 bits. Por lo tanto, para poder obtener el Master será

---

<sup>24</sup> Conocidas como Libro Rojo e incluidas en las Normas ISO 9660.

imprescindible re-muestrear la totalidad de los archivos. Para llevar a cabo este recorte sin perjudicar la calidad del audio, recurrimos a la herramienta de Dithering<sup>25</sup> que es en sí mismo un proceso mediante el cual se incorpora ruido imperceptible para el oído humano para corregir este problema.

Al reducirse la resolución, la onda pierde puntos de referencia y como resultado de eso tiende a hacerse mas cuadrada y a generar clicks entramados<sup>26</sup> que enturbian el sonido. El Dithering elimina estos errores devolviendo a la toma de audio su transparencia original aún habiendo reducido la resolución a 16 bits. Los programas para la Masterización de audio incluyen siempre herramientas de Dithering. Estas herramientas permiten seleccionar un valor de resolución en bits al que como regla general le asignaremos el mismo valor en bits que tendrá el archivo de salida. De esta forma, si buscamos obtener un Master de acuerdo a la normativa para la producción de discos compactos, usaremos un Dithering de 16 bits.

### **Conclusiones. Aplicación de estas técnicas a la Restauración de archivos de audio en soporte analógico**

Hasta aquí hemos recorrido las principales técnicas de Masterización Digital de Audio que son aplicables a todo tipo de archivos. Hemos analizado los diferentes roles a partir de la grabación de tomas originales, pasando por la etapa de mezcla y luego masterización hasta obtener un Master listo para su replicación. El presente Proyecto de investigación nos lleva a aplicar estas técnicas para la Restauración de un catálogo discográfico editado hace casi cuatro décadas en soporte de discos de vinilo y cuyos Masters se han perdido. Esta circunstancia hizo que el primer paso fuese el rescate, inventariado y recuperación de los discos sobrevivientes, tarea que se realizó en el primer año de trabajo, de acuerdo al primer informe de avance oportunamente presentado.

---

<sup>25</sup> La traducción al español de Dithering es Tramado, utilizado en España pero término poco usado en Latinoamérica donde es mas habitual la expresión original en inglés

<sup>26</sup> Errores de cuantización digital que aparecen en forma de múltiples y pequeños clicks de baja amplitud, entramados por encontrarse ocultos dentro de la trama sonora.

El hecho de contar con discos de vinilo como grabaciones originales implica como primer paso la necesidad de migrar los mismos al dominio digital, para lo cual resulta imprescindible digitalizar los discos pasando por un pre-amplificador compatible con la curva RIAA que permite reconstituir la forma de onda original que ha sido limitada en frecuencias graves en oportunidad de editar los discos en dicho soporte.

Los registros originales sobre soporte analógico de vinilo deben ser ante todo preservados para evitar que su deterioro continúe, y esa tarea debe realizarse de acuerdo a las recomendaciones *TC-03* de IASA (2005) y a la *Convención para la Salvaguardia del Patrimonio Cultural Inmaterial* UNESCO (2003). Desde el punto de vista patrimonial consideramos adecuado ordenar los pasos a cumplir de acuerdo a lo expuesto en el artículo de Daniel Schachter *El Circuito de Restauración de Patrimonio Sonoro. Las técnicas de Masterización en la recuperación de antiguas grabaciones*, que propone un ordenamiento razonable en el tiempo de los pasos necesarios para llevar a cabo esta tarea garantizando la preservación de los archivos históricos, su migración al dominio digital para obtener copias digitales idénticas a los originales analógicos, como paso previo a la Restauración de los mismos mediante la aplicación de modernas técnicas de masterización para obtener nuevos masters digitales compatibles con los estándares actuales.

Las pruebas piloto desarrolladas en el segundo año de trabajo en el Proyecto permitieron concluir que una vez llevada a cabo la migración, la primera tarea es obtener una copia de esos archivos para trabajar sobre la misma. Una vez hecho esto todo el trabajo posterior debe realizarse sobre las copias de esos archivos a los que consideramos como nuevos originales.

Las mencionadas pruebas piloto pusieron de manifiesto que los registros de EDUL discos presentan condiciones de deterioro muy variables, como así también son variables sus cualidades sonoras y las características de las grabaciones. Por lo tanto, con respecto a las técnicas de Masterización aplicables para llevar a cabo la Restauración de la colección completa de

registros de EDUL, llegamos a la conclusión de que los procesos a aplicar no pueden ser idénticos en todos los ejemplares.

Otra conclusión que se hace evidente al momento de aplicar las técnicas de Masterización para llevar a cabo la Restauración de estos registros es que ante todo debemos separar las grabaciones monoaurales de las estereofónicas. Toda vez que debamos trabajar sobre un original monoaural será imprescindible cerciorarse de que esa condición no responde a la voluntad expresa del autor de la música grabada en ese original. Esto es importante porque de ninguna manera una remasterización puede contrariar la voluntad del autor. Así, solo continuaremos interviniendo aquellos originales monoaurales que de ser por los autores hubiesen sido estereofónicos desde un primer momento. En cambio, dado que en todos los casos estaremos en condiciones de mejorar la relación señal-ruido con la aplicación de un filtrado de ruido complejo por FFT, aquellos registros monoaurales por voluntad expresa de su autor podrán de todas formas beneficiarse con el proceso de remasterización ya que la claridad de la toma de audio y el rango dinámico total se incrementarán.

Las pruebas piloto demostraron que luego de aplicar cada proceso de masterización en el dominio digital, tanto a los discos monoaurales como a los estereofónicos, es necesario controlar y de ser necesario corregir el corrimiento del eje de cero absoluto o *DC offset*. Surge también como resultado de las citadas pruebas que los nuevos originales remasterizados, una vez optimizados sus valores de dinámica presentan un rango dinámico mayor. Además, el filtrado de ruido por FFT permite una mejor apreciación de las texturas registradas en las grabaciones originales. La eliminación de clicks y demás ruidos presentes en los originales sobre soporte analógico de vinilo mejoran sensiblemente la calidad de las antiguas tomas y una vez alcanzado este resultado es aún posible estudiar en cada caso la conveniencia o no de expandir la imagen estereofónica.

Daniel Schachter



## Bibliografía

Gibson, B. (2012). *Mixing & Mastering - Recording Method: Libro 6 (2a. Edición)*. New York, EE.UU.: Hal Leonard.

IASA (2005) TC-03. *La salvaguarda del patrimonio sonoro: ética, principios y estrategias de preservación*. Trad. Osorio Alarcón Fernando. Último acceso 10/03/2014 [www.iasa-web.org/tc03/ethics-principles-preservation-strategy](http://www.iasa-web.org/tc03/ethics-principles-preservation-strategy)

Katz, B. (2002). *Masterización de Audio: el arte y la ciencia (orig: Mastering audio: the art and the science) 1a Edición*. Burlington: Elsevier/Focal Press. Versión castellana: Escuela de cine y video, Guipuzcoa.

Owsinsky, B. (2008). *The mastering engineer's handbook. 2a Ed.: The audio mastering handbook*. Boston: Thomson.

Suzuki, Y., & Takeshima, H. (2004). *Equal-loudness-level contours por pure tones*. Journal of Acoustic Society of America , 918-933.

Truesdell, C. (2007). *Mastering - Digital Audio Production*. Indianapolis, Indiana, EE.UU.: Wiley Publishing Inc.

UNESCO (2003) *Convención para la Salvaguardia del Patrimonio Cultural Inmaterial*". Último acceso 03/03/2014  
<http://unesdoc.unesco.org/images/0013/001325/132540s.pdf>

Watkinson, J. (1994). *An Introduction to Digital Audio*. Oxford, Reino Unido: Focal Press.