



POSTPRODUCCIÓN DE AUDIO DIGITAL
EDICION MEZCLA MASTERIZACION

Lino García Morales

LA MÚSICA ES LA ARITMÉTICA DE LOS SONIDOS COMO LA
ÓPTICA ES LA GEOMETRÍA DE LA LUZ
CLAUDE DEBUSSY

LA MÚSICA ES SONIDO ORGANIZADO
EDGAR VARÈSE

LA MÚSICA ES EL ARTE QUE NOS PERMITE RECORDAR LO
QUE
AÚN NO HA SUCEDIDO
JOSÉ MARÍA VITIER

A Hugo, Héctor y Viki.

Índice general

Estudio

Estación de trabajo

Edición

Mezcla

Masterización

Apéndices

Bibliografía

Introducción

La postproducción musical es un proceso complejo orientado al «producto», «soporte» o «medio» para el cual se produce (e.g., televisión, radio, cine, vídeo, música, gráfica, educación, etc.)¹. Cada *medio* tiene diferentes exigencias y estos desarrollos se suelen plantear en forma de proyectos con diferentes etapas, más o menos independientes, relacionadas u organizadas según la función audiovisual que desempeñe. Los medios de reproducción más frecuentes en la postproducción de audio digital hoy día son: CD (Compact Disc), SACD (Super Audio CD), mp3 (MPEG 1, Layer 3) y ¡Vinilo!² En el futuro habrá nuevos medios, con seguridad, pero no serán mejores³ porque el audio digital ha superado con creces las limitaciones del sistema perceptual auditivo. Por ahora se dispone del *streaming*⁴. Dicho de otra manera ¡el audio digital se comporta como audio analógico! El medio es el «soporte de almacenamiento» desde donde vamos a disfrutar del programa musical con el «equipo de reproducción» correspondiente. Se postproduce para un medio.

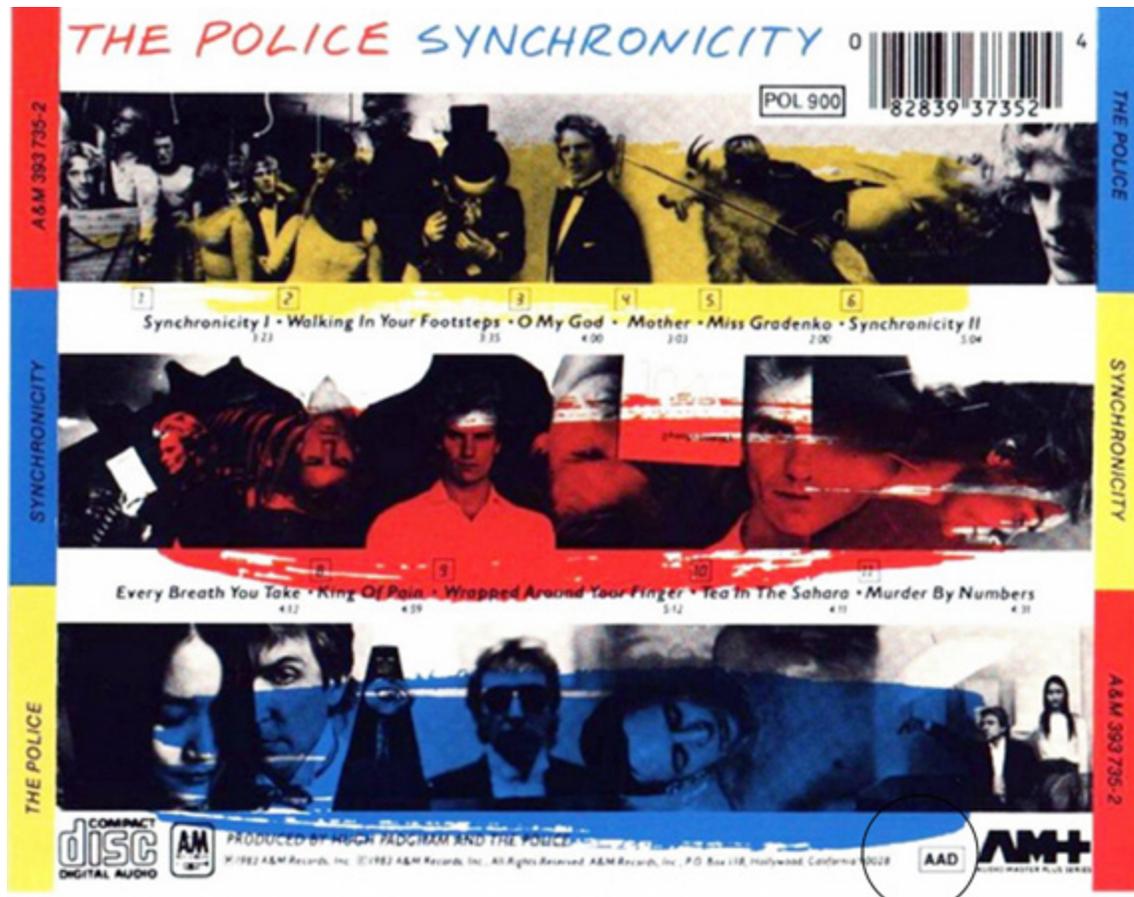


Figura 1: Reverso del CD (Compact Disc) de *The Police, Synchronicity*. Publicado por A&M Records en 1983.

Hasta hace muy poco los programas musicales se marcaban como DDD; tanto es así que desapareció prácticamente de todos los soportes. Cada vez es más común encontrarse con una marca DDA debido a la vuelta del vinilo; a ello volveré más adelante.

Observe que abajo a la derecha de la [Figura 1](#) se indica, encerrada por un círculo, una marca formada por la combinación de tres letras: AAD. Estas letras indican la tecnología (analógicadigital) de grabación, mezcla y masterización (que en éste caso también se denomina *transferencia*), respectivamente. AAD, por lo tanto, indica que tanto la grabación como la mezcla de éste CD se realizaron en el dominio analógico (con tecnología de audio analógica)

mientras que la masterización se hizo para soporte digital; CD, en este caso.

Si el medio es el mensaje, como sentencia McLuhan, la postproducción es la máxima responsable de que el mensaje llegue como debe al receptor, desde el emisor.

Preproducción, Producción, Postproducción

El trabajo de *preproducción* está relacionado con la planificación y organización del proyecto musical. Por ejemplo, el trabajo de producir un disco de música e introducirlo en el mercado requiere de un planteamiento inicial (que puede incluso empezar con un estudio de mercado), la selección y contratación de los músicos, los estudios e ingenieros, los temas, la planificación de todo el proceso, etc. En el mundo del entretenimiento a esta primera etapa se le suele llamar *producción* simplemente. La *preproducción* debe proveer un plan organizado de *dónde*, con *quién* y *cómo* se hará cada parte del proyecto musical: estudios, ingenieros y músicos, calendario, equipamiento, etc. Esta etapa es muy importante si quieres minimizar los gastos del proyecto y terminar en la fecha prevista.

La función principal de la etapa de producción, genéricamente, es la recopilación de todo el material sonoro necesario para la consecución del proyecto. En el caso de la producción musical de un disco, es necesaria la grabación de los temas según la agenda programada en preproducción (por lo cual se le suele llamar sencillamente como etapa de *grabación*). La *producción* debe proveer múltiples tomas o grabaciones de los diferentes instrumentos en cada tema y muchos más temas de los que formarán finalmente el producto musical. Lo más importante de la producción es que el material sonoro esté limpio de distorsión y capture la sensibilidad de los temas.

El término «masterización» proviene de producir un «master» de donde “copiar” o “estampar” las múltiples copias idénticas de consumo. *Master* se puede traducir como *maestro* o *matriz* pero *maestrización* o *matrización* no parecen muy adecuadas por lo que simplemente se castellaniza *mastering* como *masterización*.

La fase de *postproducción*, finalmente, concluye el proyecto musical y está compuesta por tres etapas fundamentales: *edición*, *mezcla* y *masterización*. En una producción musical los temas son las «partes» y el «todo» es el programa final: CD, vinilo, etc. La *edición* tiene por objetivo seleccionar, ajustar y “acomodar” todos los fragmentos grabados para que puedan ser utilizados, de manera óptima, en la *mezcla*. La *mezcla* es el proceso de conformar una pista “correctamente” a partir de todos los fragmentos grabados. Los procesos de *edición* y *mezcla* ajustan y corrigen el material de partida. Estos ajustes pueden ser más o menos creativos que pueden alejar más o menos el resultado final del punto de partida. Por eso se podría afirmar que estos procesos son procesos generadores de ilusión. El producto final, en general, no se atiene a una ética de la verdad, sino más bien todo lo contrario. El producto final es más falso cuanto más se aleje del punto de partida pero de eso se trata, de generar ilusión, no de generar material reproducible⁵. La *masterización* es el proceso final de conformar un producto musical “correctamente” a partir de todas las pistas seleccionadas. Luego se analizará qué se considera “correcto”. El producto final se postproduce para un medio de distribución específico: CD, vinilo, mp3, etc.

Como se puede observar los procesos de edición, mezcla y masterización constituyen un flujo secuencial en el que raras veces es necesario volver a atrás. Las tres etapas comparten algunas herramientas, equipos y procesos; las entradas de una son, en definitiva, las salidas de otra.

Según lo que se quiera, se usan de una manera u otra pero para conseguir un único objetivo: llegar al final con determinado nivel de satisfacción y calidad y también, porque no decirlo, sin salirse del presupuesto. En definitiva la postproducción es un proceso híbrido de ingeniería y arte en un contexto de compromiso calidad-coste. Gran parte de las herramientas, equipos y procesos de «postproducción», en general, son los mismos de «producción»; por lo que, aunque este libro trata solo de la última etapa: la *postproducción*, también puede ser útil en algunos aspectos para la *producción*.

Estudio

El «estudio» es el entorno donde se realizan los procesos de postproducción y también “es” los equipos y herramientas que se utilizan dentro de ese entorno para determinado propósito. Obsere que he preferido utilizar el término *entorno* y no *lugar*. Es evidente que el estudio debe ocupar determinado lugar; al menos donde quepa el técnico encargado de la postproducción y el equipamiento mínimo necesario. Pero entorno no es solo lugar, recinto, sala, un lugar especial desde el punto de vista acústico, sino también equipamiento y herramientas. Las tecnologías digitales reducen la noción de entorno físico a entorno virtual. Muchas de las herramientas digitales no ocupan lugar; si acaso un fragmento de espacio en un disco de almacenamiento, memoria, y tiempo de ejecución de un procesador.

Herramientas

Todos los estudios disponen, al menos, de un conjunto de equipos y/o programas, llamémosle «herramientas» de manera genérica, destinados a: *grabar, reproducir, procesar* y otro, llamémosle conjunto de «auxiliares», que posibilitan y facilitan todo lo demás. Las herramientas de grabación tienen la función de adquirir, del mundo real, el contenido sonoro necesario para el proyecto con la satisfacción de determinados requerimientos de calidad. Por ejemplo, para

un proyecto de *producción* musical, pueden ser necesarios micrófonos, cables, consola de mezclas, grabadores multi-pistas, sistema de monitorización, etc. De todos los dispositivos que intervienen, solo algunos “graban”; pero esto no sería posible sin la asistencia de otros auxiliares. Con las herramientas de *postproducción* ocurre más de lo mismo.

Para tener una realimentación del proceso es necesario *escuchar* el contenido sonoro; en definitiva esa será su función principal, *ser escuchado*, una vez convertido en producto de consumo. Uno o más monitores (o auriculares), con sus correspondientes accesorios (como el amplificador en el caso de los monitores), son imprescindibles en cualquier etapa de realización de un proyecto musical. En el Apéndice *Monitores*, página →, tiene información acerca de monitores.

Usaré el término «monitor» para referirme al conjunto de recinto, altavoz y electrónica interna (ya sea *pasiva* o *activa*; en caso de que la tenga). Es la manera en que se puede «monitorizar» o «monitorear» el sonido.

Las herramientas de procesamiento permiten manipular, transformar, procesar, los contenidos sonoros ya sea *técnicamente* o *estéticamente*; más adelante ahondaré en la diferencia. Las herramientas son las mismas, pero el propósito es diferente. Por ejemplo, la corrección de la afinación en un segmento concreto de la voz cantada (una vez grabada) es una exigencia técnica, ¡la desafinación es inadmisible!, pero la alteración del timbre de un instrumento (con distorsión, por ejemplo) o la incorporación de reverberación artificial para conseguir determinado efecto espacial es algo estético orientado a conseguir determinada “impresión”.

Existe una serie de dispositivos sin los cuales sería imposible equipar adecuadamente un estudio: cables (Apéndice *Conexionado*, página →), soportes, equipos de suministro de energía eléctrica, equipos de comunicaciones,

etc. Tales dispositivos son imprescindibles para realizar cualquiera de las funciones anteriores: grabación, reproducción y procesamiento. Quizá la familia más importante de herramientas auxiliares son los equipos o dispositivos de almacenamiento. Aunque existen aplicaciones musicales que reproducen audio en tiempo real (*streaming*), en lugar de “en diferido”, es raro que no se “guarde” algún contenido: ya sea como memoria histórica, para un procesamiento posterior, una cuidadosa edición o re-edición, etc. Los dispositivos de almacenamiento suelen estar integrados en el equipamiento de grabación y reproducción, e incluso de procesamiento, pero por sus particularidades y por la posibilidad de aislarlo funcionalmente del resto merece un tratamiento exclusivo.

Todos los equipos tienen determinada «calidad», lo que redundará en determinadas prestaciones, que puede ser evaluada mediante la medición de determinadas especificaciones técnicas, y se distribuyen en tres grandes gamas según su orientación al mercado y no sin cierto solapamiento. Existen equipos *domésticos* (de consumo), *semi-profesionales* y *profesionales*. ¿Qué parámetros o especificaciones técnicas determinan estas franjas?



Figura 2: Vúmetro.

Niveles de entrada-salida El nivel de una señal eléctrica suele darse en función de su *valor cuadrático medio* (RMS, *root mean square*) o *valor eficaz*⁶:

$$RMS_{x(t)} = \sqrt{p_{x(t)}}, \quad (1)$$

siendo $p_{x(t)} = \frac{1}{T} \int_T x^2(t) dt$ la potencia de $x(t)$. Observa que $x^2(t)$ es la *energía* instantánea de la señal en t . La integral acumula toda la energía en un período de tiempo T que luego se promedia para obtener la *potencia*; la raíz cuadrada devuelve una estimación del nivel en ese período, es decir, el *valor eficaz*.

El decibelio es un valor logarítmico que expresa la diferencia entre dos niveles o un cambio en el voltaje o la potencia de la señal eléctrica. Los decibelios dependen de la referencia. Así, 0 dBm es una medida de potencia que equivale a 1 mW , 0 dBu es una medida de voltaje que equivale a $0,775$ voltios, 0 dB SPL es una medida de presión sonora que equivale a $0,0002 \text{ dinas/cm}^2$ y 0 dBFS es el nivel más alto que puede tener la señal en el dominio digital donde cada bit aporta, aproximadamente, 6 dB (0 VU está referenciado a dBu y dBFS en dependencia de la referencia que indiquen sus equipos). Un decibelio, como es de suponer, corresponde a la décima parte de un Belio definido como $B = \log \frac{I}{I_{ref}}$, siendo I nivel y I_{ref} nivel de referencia; por lo que $dB = 10 \log \frac{I}{I_{ref}}$. Observe que aquí el dB está relacionado con potencia pero si el dB hace referencia a voltaje o nivel de presión acústica el nivel de referencia cambia según $I = A^2$ por lo que $dB = 20 \log \frac{A}{A_{ref}}$. En entornos profesionales, y en su estudio casero (*home studio*) la referencia es $0 \text{ VU} = +4 \text{ dBu}$. Si en algún vúmetro ves $0 \text{ VU} = +4 \text{ dBm}$, es que la impedancia es 600Ω , que lo hacen equivalente a $+4 \text{ dBu}$. En el mundo digital la escala de dB comienza en el nivel más alto que puede ser digitalizado y va descendiendo (los decibelios son negativos). Por eso se denomina en inglés *Full Scale*.

Esto es, de hecho, lo que muestran los indicadores de nivel de las mesas, por ejemplo, reforzado por un código de colores simple: **zona verde** -ningún problema; **zona amarilla** -cuidado, se invade el margen de guarda (*headroom*); **zona roja** -peligro, saturación = distorsión. En general las medidas se expresan en decibelios, una forma más compacta y relativa de representación; por una parte los números son más pequeños y por otra es relativa a una referencia según lo que se quiere medir. La medida más utilizada en los equipos de audio, por ejemplo, es el *dBu* definido por

$$dBu = 20 \log \frac{RMS_{x(t)}}{0,775}, \quad (2)$$

donde $RMSx(t)$ es el valor eficaz de $x(t)$ medido en t y 0,775 es el valor de voltaje de referencia. Otras unidades de medida que puedes encontrar en equipos de audio analógico son el *dBm* (1 mW de referencia en una carga de 600Ω) y el *dBV* cuya referencia es 1 V. *dBu* y *dBm* son equivalentes cuando la tensión sobre una carga de 600 Ω es 0,775 y la potencia es 1 mW.

En los equipos digitales se utiliza el *dBFS* (*dB* con respecto al fondo de escala, *Full Scale*). Aquí el nivel de referencia es el máximo valor digital (en valor absoluto) que corresponde a la máxima tensión del convertidor analógico-digital o digital-analógico ($\pm V_{max}$). No puede haber valores superiores a los 0 *dBFS*.

$$dBFS = 20 \log \frac{|RMS_{x(n)}|}{2^{N-1}},$$

donde $RMSx(n)$ es el valor eficaz equivalente de la secuencia $x(n)$ y N el número de bits. Si $N = 16$ bits, como el CD, el rango de las muestras en complemento a dos varía

entre +32767 y -32768. La referencia es el máximo valor representable (32768).

Margen dinámico Relación entre el nivel mínimo y el máximo de una señal en un equipo. Los equipos domésticos o de consumo tienen -10 *dBV* (nivel de referencia 1 V) de nivel nominal, mientras que los equipos profesionales tienen +4 *dBu* de nivel nominal. +24 *dBu* corresponde al máximo nivel de salida.

Relación Señal/Ruido (SNR, Signal/Noise Rate) Diferencia en *dB* entre el nivel nominal del equipo y el nivel de ruido que introduce; cuanto más alto mejor. El ruido, 60 *dB* por debajo del nivel de la señal, queda enmascarado totalmente. En un equipo de audio profesional la *SNR* es superior a 80 *dB*. Este valor es el correspondiente al nivel nominal del equipo, un valor elevado de la señal de audio. Lo normal es que el nivel de la señal de audio fluctúe dinámicamente y la *SNR*, en la práctica, sea bastante inferior.

Respuesta en frecuencia Gama de frecuencias a las que responde el dispositivo. Una reproducción de buena calidad debe cubrir todo el margen de frecuencias de audio, normalmente definido en el rango audible: entre 20 Hz y 20 kHz. Un equipo de audio profesional no debe superar 1 *dB* de variación respecto al nivel de la frecuencia de 1 kHz y el margen de frecuencias debe llegar hasta los 20 kHz como mínimo.

Distorsión armónica y/o de intermodulación La *distorsión armónica* mide la "no linealidad" del equipo teniendo en cuenta que lo deseado es que sea lineal. La no linealidad de los equipos se presenta fundamentalmente en los niveles extremos de su margen dinámico (niveles muy bajos, ruido; niveles altos,

saturación=distorsión=ruido). La *distorsión armónica* (THD, *Total Harmonic Distortion*) se produce cuando la señal de salida de un sistema no se corresponde con la señal de entrada.

En un entorno analógico, el audio comienza a ser distorsionado cuando saturamos los niveles a partir de 24 *dBu*. Realmente dependiendo del equipo puede comenzar distorsionar entre 20 *dBu* y 28 *dBu*, pero se ha aceptado 24 *dBu* como un estándar. En audio digital, 0 *dBFS* es el nivel más alto que puede tener el audio y es a partir de aquí cuando la señal se distorsiona. Entonces, se puede decir que 0 *dBFS* corresponde a un nivel analógico de 24 *dBu*. Si se acuerda esta relación, +4 *dBu* es lo mismo que -20 *dBFS* y que 0 VU.

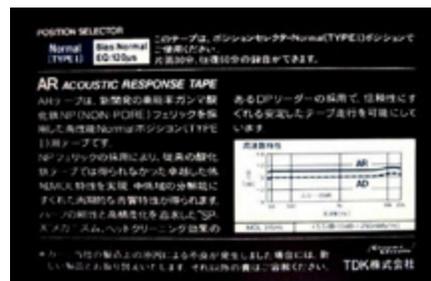


Figura 3: Respuesta acústica de una cinta compacta TDK.

Si entra un tono en el equipo, por ejemplo, salen un conjunto de tonos armónicos de menor nivel múltiplos enteros de la frecuencia de entrada.

El valor

$$THD = \sqrt{\frac{\sum_{k=2}^K p_{x_k(t)}}{p_{x_1(t)}}} 100\%,$$

está dado por la raíz cuadrada de la relación entre las potencias ($p_{x_k(t)}$) de los armónicos “superiores”, $k = 2 \dots K$

de $x(t)$, y la potencia del fundamental $k = 1$; con $K = 5$ armónicos superiores es suficiente para estimar el *THD*. El valor de *THD* para un equipo de audio profesional debe ser inferior a 0,1 %, que corresponde a una diferencia de niveles entre el fundamental y los armónicos de más de 60 *dB*.

La *distorsión de intermodulación* mide la *distorsión armónica* en altas frecuencias. Observa que si intenta medir la *distorsión armónica* en altas frecuencias con una señal sinusoidal, por ejemplo, de 10 kHz, los armónicos resultantes quedan fuera de la banda de audio (>20 kHz) y los armónicos generados por la no linealidad del equipo deben estar dentro del margen de frecuencias de audio (20 Hz a 20 kHz). Para solventar este problema se utiliza una mezcla de señales con dos componentes de frecuencias. Los productos de intermodulación (armónicos de esta mezcla) surgen en las frecuencias $mf_1 \pm n f_2$; donde m y n son números enteros ($m, n = 1, 2, 3, 4 \dots$). El valor de f_1 y f_2 y de la relación de amplitudes de ambas depende de la norma según:

SMPTE $f_1 = 60$ Hz, $f_2 = 7$ kHz, relación 4 : 1

DIN $f_1 = 250$ Hz, $f_2 = 8$ kHz, relación 4 : 1

CCIF $f_1 = 19$ kHz, $f_2 = 20$ kHz, relación 1 : 1, ancho de banda 0 – 20 kHz ó $f_1 = 13$ kHz, $f_2 = 14$ kHz, relación 1 : 1, ancho de banda 0 – 15 kHz

En las normas *SMPTE* y *DIN* los K *productos de intermodulación* $i_k(t)$ aparecen en torno a la componente de frecuencia más alta, 7 ó 8 kHz separados entre sí 60 ó 250 Hz ($f_2 \pm k f_1$), respectivamente: $f_2 \pm f_1$ armónicos de segundo orden (par, $k = 1$), $f_2 \pm 2 f_1$ armónicos de tercer orden (impar, $k = 2$), $f_2 \pm 3 f_1$ armónicos de cuarto orden

(par, $k = 3$), $f_2 \pm 4 f_1$ armónicos de quinto orden (impar, $k = 4$), etc. En el caso de la norma CCIF los armónicos aparecen a partir de 1 kHz, separados entre sí también 1 kHz. Al igual que en el caso de la distorsión armónica, el peso de los armónicos pares o impares depende de la simetría de la curva de transferencia de nivel. Si la curva es simétrica predominan los armónicos impares; en caso contrario, predominan los armónicos pares. El valor de la distorsión de intermodulación

$$IM = \sqrt{\frac{\sum_{k=1}^K p_{i_k(t)}}{p_{x_1(t)} + p_{x_2(t)}}} 100\%,$$

donde $p_{i_k(t)}$ corresponde a la potencia de cada uno de los *productos de intermodulación*, y el denominador, $p_{x_1(t)} + p_{x_2(t)}$ a la potencia de la señal de entrada; correspondiente a la mezcla de las sinusoides de frecuencias f_1 y f_2 .

Separación estéreo o diafonía Describe el acoplamiento (*crosstalk*) entre señales. Si el equipo trabaja con una señal estéreo el término correcto a emplear es “separación estéreo”, a diferencia de “diafonía”. En un sistema multi-pista, por ejemplo, se presupone la no correlación o independencia en las señales de cada pista; no hay nada de una pista en sus vecinas. Un equipo de audio profesional debe tener valores de separación estéreo (o diafonía) superiores a los 60 dB; son valores más que suficientes para enmascarar el acoplamiento entre canales. Si la señal es estéreo el valor de *crosstalk* podría ser inferior; a partir de unos 30 dB de separación el oído es incapaz de percibir acoplamiento entre señales correladas (que se parecen).

He formulado, a manera de ejemplo, cómo se obtienen alguna de las medidas, pero todas estas medidas y sus respectivos protocolos o métodos de medición están perfectamente identificadas y documentadas.

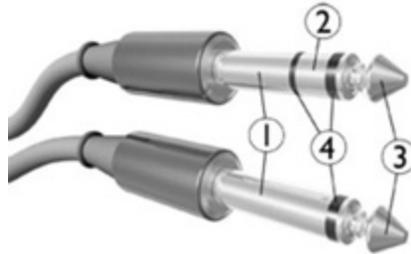


Figura 4: **arriba.** TRS. **abajo.** TS. ① Cuerpo, *Sleeve*: tierra (pantalla, *shield*). ② Anillo, *Ring*: canal derecho para señales estéreo, polaridad negativa para señales mono balanceadas, fuente de alimentación para fuentes de señal mono con fuente de alimentación. ③ Punta, *Tip*: canal izquierdo para señales estéreo, polaridad positiva para señales mono balanceadas, señal de línea (*hot*) para señales mono no balanceadas. ④ Anillos de aislamiento.

El procedimiento de medida en todos los casos es el mismo: se introduce en la entrada del equipo a evaluar una señal (o señales) de referencia y se observa la variación de alguno de los parámetros (amplitud, frecuencia, fase) de la señal en la salida. El equipo (sistema) a medir es como una *caja negra*. Solo se ve lo que entra y lo que sale de él y no se sabe nada más. La señal (o señales) de referencia depende de la medida a realizar⁷.

No solo es importante la calidad de los equipos, sino la calidad de sus interconexiones, de los cables y conectores. Es bueno gastar parte del presupuesto en buenos cables y conectores pero con eso no basta. La coexistencia en una misma instalación de equipos profesionales y de consumo introduce dos tipos de alteraciones. La primera derivada del conexionado diferente (balanceado y no-balanceado); y la segunda debida a los diferentes niveles e impedancias de los equipos involucrados. Las conexiones balanceadas, simétricas, son típicas de los equipos profesionales y las

conexiones no-balanceadas, asimétricas, son típicas de los equipos semi-profesionales y de consumo.

La conexión asimétrica consiste en una señal referenciada con respecto a masa. Ejemplos de estas conexiones son el *jack* de 1/4 pulgada (TS, *Tip-Sleeve*; Punta-Cuerpo) que usa cualquier guitarra, bajo, teclado, etc. o el RCA (*Radio Corporation of America*) que usan los reproductores de CD comerciales. Este es un tipo de sistema de transmisión de señales *single-ended*: la transmisión de señal se efectúa sobre un conductor que retorna por la masa del circuito o potencial de referencia. Normalmente son cables apantallados con una malla (*shield*) que actúa como jaula de Faraday a la vez que conductor de masa y retorno de la señal. En principio, teóricamente, la pantalla del cable debe aislar al conductor de interferencias electromagnéticas (EMI, *ElectroMagnetic Interference*) y protegerlo del ruido. En la realidad la pantalla posee una resistencia finita y las EMI provocan la circulación de pequeñas corrientes dentro de la propia pantalla; corrientes que se suman a las de la señal “deseada”, y son amplificadas de manera tal que, una vez mezcladas con la señal, es prácticamente imposible separarlas⁸.

La conexión simétrica consiste de dos señales referenciadas con respecto a masa. En lugar de uno, tiene dos conductores que “ven” la misma impedancia con respecto a masa (el potencial cero, o de referencia). Pero las amplitudes de la señales en ambos conductores, aunque son exactamente igual, están en contrafase. En el receptor la señal diferencial se obtiene de restar ambas señales en contrafase. El ruido es aditivo, se suma por igual en ambas señales, por lo que se elimina y la resta de las señales, al tener polaridad opuesta, en realidad las suma. Ejemplos de estas conexiones son el *jack* de 1/4 ó 1/8 de pulgada (TRS, *Tip-Ring-Sleeve*; Punta-Anillo-Cuerpo), el Canon o XLR (*eXternal Live Return*), etc.

La diferencia entre conectores no es un problema; se soluciona simplemente con el adaptador adecuado pero la diferencia del «tipo» de conexión: balanceada/no-balanceada, simétrica/asimétrica debe resolverse con el acoplamiento adecuado: una *caja de inyección* (caja DI), por ejemplo. La caja DI es un dispositivo electrónico que se encarga de transformar la señal procedente de una línea no-balanceada (*line*, TS, RCA) en una señal equilibrada de baja impedancia que puede ser transportada por una línea balanceada (*mic*, XLR, TRS). Las cajas DI son muy útiles para adaptar las altas impedancias de los instrumentos o cualquier equipo con salida de línea, a mezcladores de sonido que solo poseen entradas para líneas balanceadas y para realizar tiradas de cable largas. Una caja DI permite conectar una guitarra eléctrica o un piano eléctrico a la entrada *mic* de una mesa de mezclas “correctamente”. Sin una caja DI la conexión de un equipo profesional a un equipo de consumo produce saturación y distorsión y la conexión de un equipo de consumo a un equipo profesional produce atenuación.



Figura 5: ① Cuerpo: tierra (pantalla, *shield*). ② *Hot*: canal izquierdo. ③ *Cold*: canal derecho.



Figura 6: Caja de inyección Behringer DI100 Ultra-DI.

Volviendo a las franjas del mercado: la diferencia del equipamiento en cada uno de estos segmentos radica fundamentalmente en la *calidad y prestaciones* que ofrece; lo que influye en otro factor no menos importante: el *precio*. El desarrollo tecnológico hace posible cada vez mejores prestaciones a menor coste pero la industria mantiene las fronteras. El mercado de consumo o doméstico es el sector más amplio. Los equipos están destinados a las grandes masas. La preocupación prioritaria no es la calidad. Abundan las funciones. En esta franja de mercado: portabilidad, diseño, dimensionamiento, coste, son probablemente los aspectos que exigen mayor atención.

El mercado profesional, en el otro extremo, corresponde al sector élite. Los equipos están destinados a una minoría. Hacen solamente lo que tienen que hacer; para lo que han sido diseñados. El consumidor no es el ciudadano de a pie sino grandes productoras y gestores del medio. El objetivo principal es la *calidad*. Aunque es posible satisfacer determinado nivel de calidad cuantitativamente (traducida a especificaciones técnicas) es más fácil imaginar el significado de parámetros cualitativos como robustez, garantía, tiempo de funcionamiento ininterrumpido, tasa de fallos, etc., para hacerse una idea más amplia. Los equipos

profesionales contienen generalmente el “*state of the art*”⁹ de la tecnología; son, sin duda, el soporte tecnológico de cualquier gran inversión musical.

El mercado semi-profesional, en el medio, se mueve en un sector con grandes exigencias y menos presupuesto. Dado el vertiginoso desarrollo de la tecnología electrónica e informática es posible, cada vez más, encontrar en este sector equipamiento de buen compromiso calidad-precio. La evolución del mercado hace cada vez más difícil la diferenciación de los límites entre estos sectores. Lo que hoy es profesional en uno o dos años estará en el mercado semi-profesional y, probablemente, en cinco o siete años, en el doméstico; aunque estos tiempos se acortan cada vez más.

Existe otro aspecto relacionado con el mercado y las nuevas tecnologías: la *obsolescencia*. El mercado musical, en general, tiene tendencia al empleo de normas “de facto” versus “de lego”. Dicho de otra manera, la industria compete libremente por asegurarse un nicho de mercado y no siempre atiende a la normalización propuesta por grupos de expertos. Esto desde luego, atenta directamente contra el consumidor final: ya sea el ciudadano de a pie, la pequeña empresa o la multinacional gigantescas; pero asegura un dinamismo, dudosamente justificado, que recicla la ordenación del mercado. El equipamiento que no supera la difícil prueba de penetración automáticamente baja de categoría para competir en una liga menor.



Figura 7: La legendaria TG12345, la primera consola de estado sólido de EMI, en Abbey Road Studios. Esta consola resume la evolución del sonido de los años 60 y principios de los años 70 y ayudó a grabar a grupos como *The Beatles*, *Pink Floyd*, y muchos otros.

Piense que toda la tecnología punta que utilizó *The Beatles* en su producción, como muestra la [Figura 7](#), tiene una calidad muy inferior a la tecnología disponible comercialmente para cualquiera que quiera construir un estudio “casero” (*home studio*). Lo que también lleva a otra reflexión interesante. Disponer de la última tecnología, la de mejor calidad, no es ninguna garantía de hacer mejor música; no así lo contrario. Es posible hacer música excelente con equipos “no tan buenos”.

Otra reflexión importante es la siguiente: más caro no significa más bueno (tampoco lo contrario). Bob Clearmountain es un pionero de la postproducción musical. Su trabajo como ingeniero de mezclas es impresionante. Los músicos se lo disputan incluso hoy. A mediados de los 80s Clearmountain hizo una entrevista en su estudio para una revista de divulgación y se dejó retratar (en una pose como la de la [Figura 8](#)) delante de unos flamantes monitores de

estudio Yamaha NS10 (como los de la [Figura 8](#); los monitores del cono blanco).

Una década después los NS10 habían sido sustituidos por unos KRK. El periodista de entonces quedó estupefacto. Después de aquella primera foto, teniendo en cuenta la influencia de Clearmountain, los NS10 se multiplicaron por miles. Cualquiera que tuviera un estudio que se respetara debía tener monitores NS10. ¿Qué había pasado entonces con aquellos NS10?

Estos son mejores, dijo Bob, los NS10 eran lo que me podía permitir en esa época. Los NS10 le ayudaron a pagar la renta, comentó en otra entrevista.



Figura 8: El ingeniero de mezclas Bob Clearmountain en su estudio a mediados de los años 80.

“No soy un tío muy técnico, sólo mezclo música y utilizo las herramientas con las cuales obtengo mejores resultados. La razón porque la que funcionan, no es de gran importancia para mí”. La clave es que, aún no siendo los

“mejores” monitores de estudio: ¡funcionaban! “Sólo los conozco tan bien que realmente puedo estar seguro de lo que está pasando con una mezcla cuando los utilizo”¹⁰. Es mucho más importante explotar perfectamente lo que se tiene que estar a la última y quedarse siempre a medias y otra cosa: no confíe en lo que le vendan, compruébelo usted mismo, mídalo y explore si satisface sus expectativas. Como esta anécdota hay muchas.

Ajeno a las modas y corrientes estéticas es posible realizar un buen producto con bajo presupuesto, gracias a la maestría del técnico pero también gracias al propio desarrollo tecnológico. Hace tan solo veinte años los estudios de producción y postproducción eran asequibles sólo a grandes empresas o monopolios de la industria del entretenimiento. Pero todo ha cambiado, y mucho. El desarrollo y consolidación de la tecnología digital ha hecho posible que, con muy poco dinero, hoy día, cualquier mortal pueda producir material musical desde su estudio, montado en su pequeño apartamento, con igual o mayor calidad que en el pasado. Esta revolución digital ha hecho posible la democratización y masificación de la producción y postproducción musical. A las herramientas volveremos luego. Vayamos primero a la *sala*.

Sala de escucha

Sin profundizar en acústica arquitectónica, la *sala de escucha*¹¹, donde se realiza la postproducción, debe ser silenciosa. La sala de escucha, nuestra sala de operaciones, es tan importante que, en muchos casos, se incluye de hecho como parte del proyecto de equipamiento de estudio y el término *recinto* termina por convertirse en sinónimo de *estudio* cuando el estudio es mucho más que el recinto porque “es” también todos los equipos y las herramientas de trabajo. En un estudio de grabación musical es evidente

que el «acondicionamiento» y el «aislamiento» acústico del espacio reservado a esa función, es muy importante. El *acondicionamiento* debe garantizar determinado *tiempo de reverberación* y determinada *forma de reverberación*, el *aislamiento* debe evitar la propagación del sonido fuera del estudio por encima de cierto nivel. Dicho de otra manera, el *acondicionamiento* es importante para conseguir una buena audición y el *aislamiento* para no molestar a los vecinos y, lo que es más importantes, para que los vecinos no le molesten a usted, ni puedan estropear una sesión de trabajo.

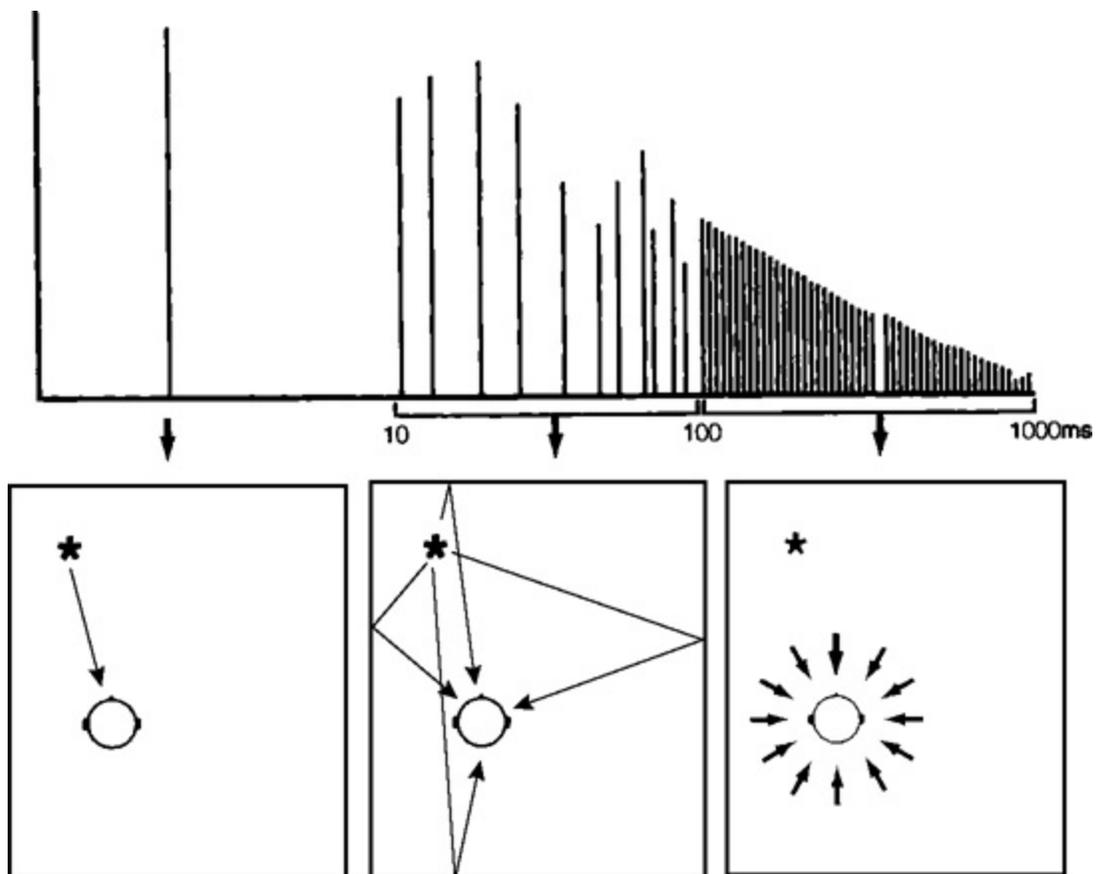


Figura 9: Reflexiones y escucha. En la postproducción solo interesa el *sonido directo (izquierda)* que llega de los monitores. Las *primeras reflexiones (centro)* y la *cola difusa* (reflexiones de alto orden, **derecha**) deben ser evitadas acondicionando debidamente el recinto. Recuerde que el techo y el suelo son también partes del recinto aunque este diagrama sólo represente un plano.

Ambas tareas suelen ser muy caras (requieren de un proyecto arquitectónico especializado). Una decisión errónea, por ejemplo, podría mover parte del presupuesto destinado al acondicionamiento hacia gastos en equipamiento. En ese caso, el rendimiento de los equipos de alta gama, al trabajar en una sala inapropiada, mermará considerablemente. ¿Cuál será el resultado? Comprar un equipo profesional para sacarle un rendimiento semi-profesional o incluso doméstico. Un buen criterio de selección del equipamiento debe adecuarse al recinto de trabajo.

Las especificaciones de la sala están normalizadas por dos estándares: ITU-R BS.1116-1 e ITU-R BS.775-1. El primero se refiere a las *condiciones de escucha*. Observe que, disponer de recintos normalizados, permite escuchar “lo mismo” en cualquier sala. Esta norma especifica parámetros como: forma de la sala, proporciones, tiempo de reverberación, etc. El segundo está orientado a la *disposición de los monitores* y su objetivo es obtener una buena *escucha de referencia*.

El documento EBU Tech. 3276-1998 fusiona las normas anteriores en un solo texto y amplía las recomendaciones para sistemas multi-canal. Según éste último deben tenerse en cuenta las siguientes especificaciones:

Sonido directo Se define como el campo sonoro medido, con los mismos monitores, en condiciones anecoicas; es decir: sin las primeras reflexiones y la reverberación de la sala de escucha. Observe que esta medida está condicionada por la calidad de los monitores.

Primeras reflexiones (early reflections) Se definen como los reflejos de las superficies de contorno que llegan a la zona de escucha dentro de los primeros 15 ms después de la llegada del sonido directo. Los niveles de estas reflexiones deben estar, por lo menos, 10 dB por debajo

del nivel del sonido directo para todas las frecuencias en el rango de 1 a 8 kHz.

Una cámara *anecoica* o *anecoide* es una sala aislada, diseñada para absorber en su totalidad las reflexiones producidas por ondas acústicas o electromagnéticas en cualquiera de las superficies que la conforman (suelo, techo y paredes laterales).

Tanto la amplitud como la frecuencia de las reflexiones individuales se pueden obtener a partir de la *respuesta al impulso* de la sala.

Reverberación Se define como el tiempo que tarda el sonido en caer 60 dB por debajo del nivel inicial. El campo de reverberación es provocado por las reflexiones de las reflexiones desde las superficies de contorno de la sala de escucha. Llega al área de escucha después de las primeras reflexiones (con retardos por encima de 15 ms). El campo de reverberación debe ser lo suficientemente difuso en el área de escucha para evitar efectos acústicos perceptibles como ecos de aleteo.

Respuesta en frecuencia Se define como la respuesta de frecuencia del nivel de presión sonora producido por el monitor (o los monitores) en cualquier punto de escucha. La señal de prueba para esta medición es ruido *rosa* filtrado con un banco de filtros de 1/3 octava. Esta especificación evalúa la influencia mutua de los *monitores* y la *sala de escucha* y por lo tanto la evaluación de las condiciones de escucha. Se corresponde bien con la evaluación subjetiva del sonido reproducido.

Nivel de escucha Caracteriza la sensibilidad de un canal de reproducción. Se utiliza para establecer la ganancia de referencia (0 dB) para ajustar los niveles de las sesiones de escucha.

Ruido de fondo El nivel de presión acústica (RMS) del ruido de fondo continuo de sistemas de aire acondicionado o de otras fuentes externas o internas, medida en el área de escucha a una altura de aproximadamente 1,2 m por encima del nivel del suelo, no debe exceder preferentemente NR 10. En ningún caso, el ruido de fondo debe ser superior a NR 15. Las *curvas de valoración de ruido* (NR, *Noise Rating*) establecen límites aceptables de confortabilidad en diferentes espacios en los que existen unos niveles de ruido de fondo estables.

Estas curvas establecen, para diferentes frecuencias, los niveles de ruido que se consideran aceptables para no causar molestias en distintos tipos de actividad; se selecciona una u otra curva NR en función de parámetros como el tipo de actividad o el local. El ruido de fondo no debe ser sensiblemente impulsivo, cíclico o tonal.

En cuanto a las dimensiones de la sala el área mínima para una sala de escucha referencia es de 40 m² y para una sala de control de sonido de alta-calidad de 30 m². El volumen de la sala no debe exceder los 300 m³.

La mayoría de la postproducción se realiza para dos canales; es decir, estéreo. Los sistemas multi-canal dependen del número de canales. Existen sistemas multi-canal 5.1, 6.1, 7.1, etc., normalizados pero no es posible postproducir para todos y cada uno de ellos; de manera tal que el contenido de un medio se pueda escuchar de una manera o de otra según el escucha. Por lo tanto me centraré en dos fundamentalmente: estéreo y mono.