

5. Edición de sonido por ordenador

5.1. Introducción

Hemos visto cómo digitalizar un sonido procedente de diversas fuentes analógicas y almacenarlo en la memoria o el disco duro del ordenador, pero el enorme potencial del audio digital por ordenador no comienza a intuirse hasta que abrimos un potente editor gráfico de audio. El paradigma de este tipo de aplicaciones es muy sencillo y varía muy poco de un programa a otro. Básicamente consiste en aplicar procesos matemáticos a un sonido (o a una porción temporal), con la ayuda de un entorno gráfico que permita seleccionar cómodamente los fragmentos a tratar y que muestre los resultados tanto sonora como visualmente. Mediante estos procesos se consigue corregir defectos en los sonidos originales, ensalzarlos, modificarlos ligeramente o generar otros nuevos totalmente irreconocibles.

Todas estas posibilidades se fundamentan sobre una disciplina de investigación que combina la ingeniería, la física y las matemáticas, y que cuenta con varias décadas de existencia: el procesado digital de señal.

5.2. Cómo elegir un programa

Las tarjetas incorporan siempre en su software un editor gráfico de sonido. Lamentable, en muchos casos estos programas son excesivamente sencillos. Un paquete de software que se salva de la quema es el *Sound Impressions* que acompaña a muchas tarjetas de fabricantes diversos. En cuanto a otros programas muy extendidos como el *Creative Wave Studio* (distribuido con las tarjetas Sound Blaster), funcionan como ventajosos sustitutos de la grabadora de sonidos de Windows, pero se quedan muy cortos a la hora de ofrecer posibilidades de edición avanzadas.

El problema se resuelve rápidamente, pues existen dos excelentes aplicaciones *shareware* fácilmente accesibles en Internet, que pueden competir en prestaciones con editores comerciales. Nos referimos a *Cool Edit* de David Johnston y *Gold Wave* de Chris S.Craig, ambas con versiones actualizadas de 32 bits. Existen otros programas de dominio público, pero difícilmente pueden competir con estos dos gigantes.

En el terreno comercial, son tres los programas que se llevan la palma: *Wave for Windows*, de Turtle Beach Systems (fabricante de las famosas tarjetas de sonido homónimas), *Sound Forge* de Sonic Foundry, y el recién llegado *WaveLab* de Steinberg (fabricante del mítico secuenciador MIDI, *Cubase*).

Wave for Windows fue el primer editor de audio profesional para PC compatibles. Uno de sus puntos fuertes es la librería de efectos preestablecidos que incorpora. Su uso es sin embargo algo más lento y engorroso que el de sus competidores. *Sound Forge* goza de un diseño más cómodo y eficaz e incorpora casi cualquier efecto imaginable, lo que lo convierte en el más potente de los tres. Por su parte *WaveLab* promete ser un programa muy completo, pero la versión 1.0 acaba de aparecer y, como es frecuente, no incluye todavía todo lo que uno podría esperar de un programa de su categoría.

De momento, estos programas operan en diferido; no son capaces de realizar los efectos en tiempo real. Sin embargo, con la inminente llegada de los sustitutos del Pentium que incorporarán DSP en la placa, y la arquitectura abierta de algunos de estos programas (*Sound Forge* y *Wave Lab*), posiblemente cambie en breve el panorama.

5.3. Utilización y posibilidades de un editor gráfico de audio digital

Simplificando, la forma de trabajar con estos programas es la siguiente: se carga un fichero de sonido digital (normalmente de tipo .wav), o se digitaliza desde el mismo programa, teniendo en cuenta todas las consideraciones del capítulo anterior. Inmediatamente, la onda queda representada gráficamente en una ventana. Con el ratón se selecciona un fragmento (igual que seleccionarías unas cuantas palabras contiguas en un procesador de texto) o el fichero en su totalidad. Mediante opciones de menú o iconos se aplican a este fragmento algunos de los procesos de modificación incluidos en el programa.

Uno de los principales raseros con que valorar las prestaciones de estos programas es el número, calidad y versatilidad de los efectos disponibles, pero existen otros factores importantes. Pasemos a enumerar y describir brevemente algunos de ellos, teniendo en cuenta que si no se indica lo contrario, los cinco programas mencionados en el apartado anterior (*Cool Edit*, *Gold Wave*, *Wave*, *Sound Forge* y *Wave Lab*) incorporan la opción.

- **Soporte multiventana.** La posibilidad de tener varios ficheros de onda abiertos y visibles, es fundamental para poder combinar fragmentos de diferentes procedencias. Todos los programas citados incorporan sofisticadas formas de gestión de memoria que permiten presentar en pantalla muchos más sonidos de los que podrían caber en la memoria RAM del ordenador.
- **Deshacer (Undo).** Dado que los ficheros de sonido pueden llegar a ocupar decenas o centenares de Mb, el guardar varios niveles de anulación puede en ocasiones ralentizar excesivamente la ejecución, o resultar incluso inviable. En estos programas la opción de **deshacer** es configurable por el usuario.
- **Edición independiente de canales.** Los ficheros estéreo se muestran siempre en dos ventanas superpuestas. En ocasiones es deseable, tal como se aprecia en la figura 5.1, poder seleccionar un fragmento de un único canal.



Figura 5.1. Selección de una única pista de un fichero estéreo en *Cool Edit*

- **Zoom de visualización.** Para trabajar los fragmentos con mayor detalle es conveniente poder modificar la escala de visualización. El zoom se puede aplicar a dos parámetros: el tiempo (eje horizontal) y la amplitud (eje vertical). Todos los programas soportan varios niveles de zoom temporal, pero no todos permiten zoom de amplitud. Esta última posibilidad es especialmente útil cuando se trabaja sobre fragmentos de muy poco nivel (casi silenciosos), ya que de lo contrario estaremos viendo una línea recta donde en realidad hay sonido¹.
- **Cambio de unidades temporales.** Existen normalmente tres unidades alternativas para medir las coordenadas temporales: tiempo (ms), muestras y compases. Esta última es útil cuando estamos editando un fragmento musical con un tempo preciso (que el usuario deberá indicar).
- **Memoria de puntos clave.** En un fichero largo es fácil perderse mientras se navega hacia delante y hacia atrás utilizando el *scroll* horizontal. Por ello es muy útil poder colocar marcas en algunos puntos importantes (inicio de un sonido, de una palabra, etc.). *Sound Forge* es el programa más completo en este aspecto, ya que no sólo permite colocar tantas marcas como deseemos, incluso de forma automática, sino que también las guarda en el fichero, para posteriores sesiones. En la figura 5.2 se muestra una ventana de este programa, con una lista de marcas.
- **Reproducción de listas.** Este punto está relacionado con el anterior: una vez colocadas varias marcas, es posible editar una lista de reproducción para alterar el orden de ejecución de los fragmentos, y repetir u omitir algunos de ellos. Esta es una forma rápida y cómoda de realizar modificaciones temporales, sin necesidad de reorganizar cada vez

¹ Tenga en cuenta que la máxima resolución vertical de una pantalla completa no supera nunca los 800 pixels, mientras que un sonido de 16 bits puede poseer hasta 65.535 niveles diferentes.

millones de muestras. *Sound Forge* automatiza además ciertos procesos relacionados con las listas como, por ejemplo, eliminar (borrar) todos los fragmentos de silencio de un fichero y crear una lista de reproducción con los mismos tiempos del sonido original.

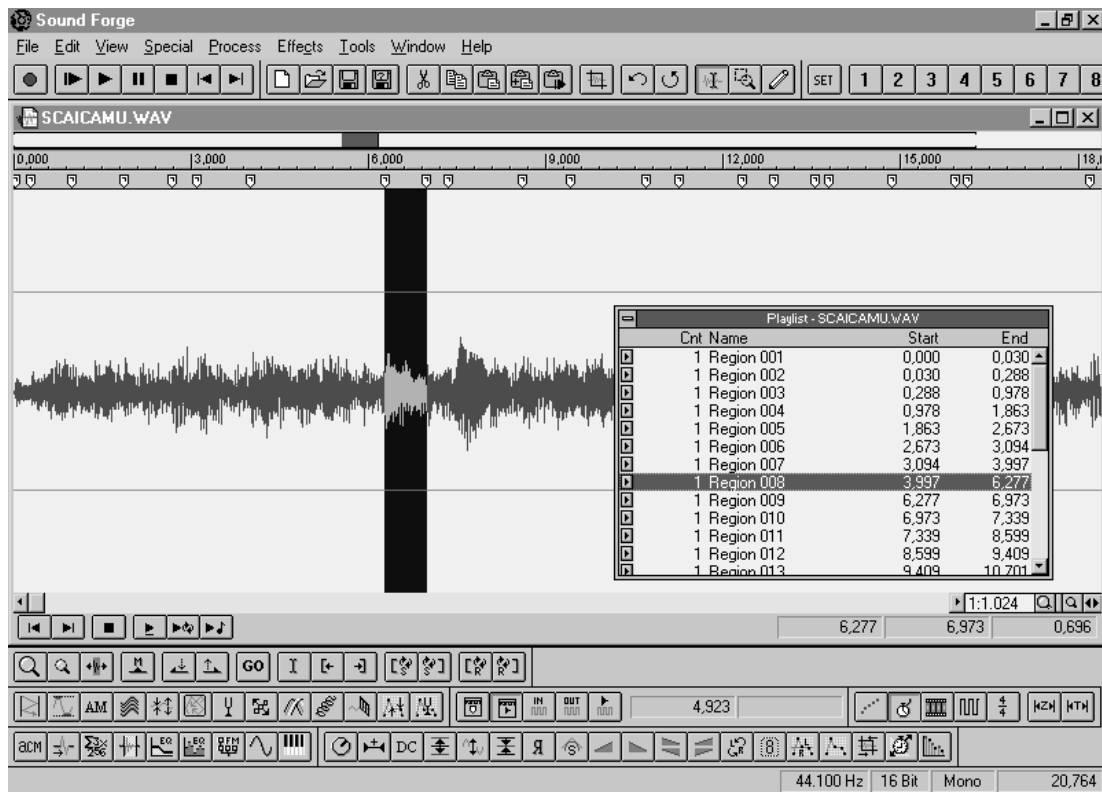


Figura 5.2. Marcas de localización y edición de listas en *Sound Forge*

- **Importación/exportación de otros formatos de onda.** Los programas para Windows trabajan principalmente con el formato .wav, pero en ocasiones es necesario importar y/o exportar formatos diferentes. El número de formatos soportados varía de un programa a otro.
- **Posibilidad de salvar configuraciones de efectos.** Muchos de los efectos aplicables comportan varios parámetros configurables. La posibilidad de salvar en disco cualquier conjunto de parámetros simplifica enormemente el trabajo.
- **Programación de nuevos efectos.** El siguiente paso en la personalización de un programa es la posibilidad de crear nuevos efectos. Esto es bastante más complejo pues no sólo requiere ciertas nociones de programación, sino un profundo conocimiento del procesado digital de señal. *Gold Wave* dispone de un editor de fórmulas que se muestra en la figura 5.3, con el que es relativamente sencillo crear efectos sorprendentes. La arquitectura abierta de *Sound Forge* y *Wave Lab* permite la creación de plug-ins escritos en lenguaje C, aunque ésta es una labor muchísimo más compleja.

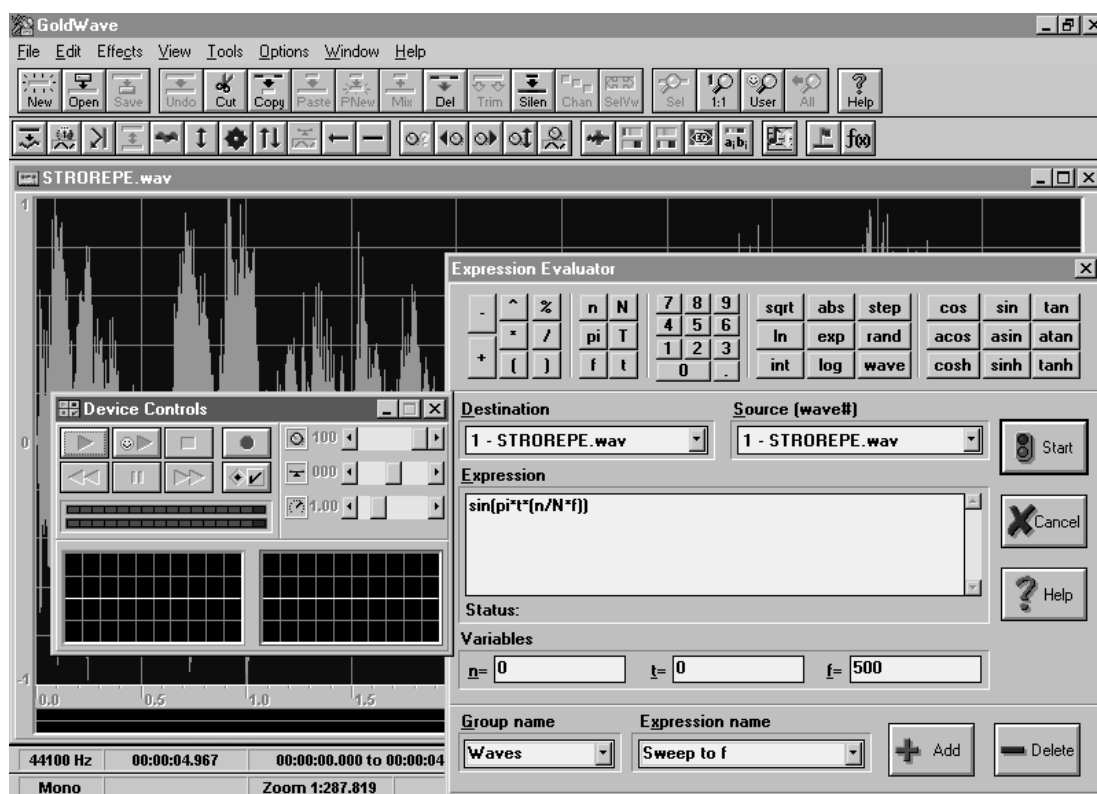


Figura 5.3. Editor de expresiones de *Gold Wave*

- **Proceso por lotes.** Cuando es necesario repetir acciones sobre un cierto número de ficheros, los usuarios de MS-DOS o Unix encuentran a faltar la posibilidad de ejecutar procesos *batch* o en lotes. En este aspecto, *Cool* es el programa mejor preparado ya que permite la escritura de sofisticados *scripts* o macros.
- **Capacidad de síntesis.** Además de todas las posibilidades de procesamiento de archivos de sonido ya existentes, estos programas ofrecen algunas opciones para crear sonidos sintéticos, es decir, partiendo de cero. En el apartado 9.7 estudiaremos varios programas exclusivamente dedicados a la síntesis.
- **Dibujar a mano alzada.** Dibujar ondas de sonido (con cierto control sobre los resultados) no es tarea fácil, pero en algunas ocasiones es la forma más sencilla de resolver algún problema. La modificación de la forma de onda mediante el ratón, es por ejemplo un método eficaz para eliminar clics u otros ruidos muy breves. *Wave* y *Gold Wave* son los únicos programas que permiten esta operación. Utilícelo con precaución.
- **Control de CD Audio - Extracción digital directa.** Los dos programas shareware ofrecen la posibilidad de controlar el reproductor de CD Audio para facilitar la grabación de fragmentos de discos compactos. La última versión de *Cool Edit*, permite además la extracción digital directa de este material (véase apartado 4.10).
- **Envío a samplers (*Generic Sample Dump Standard*).** Esta opción es útil únicamente para los poseedores de un *sampler* externo (véase apartado 9.6), ya que permite la transmisión de ficheros de audio a través de un cable MIDI utilizando el protocolo *Generic Sample Dump Standard*. *Sound Forge* y *WaveLab* son, de momento, los únicos programas que incorporan esta posibilidad, aunque si no dispone de este dispositivo externo, jamás la encontrará a faltar.
- **Activación vía MIDI.** Mediante esta opción se pueden disparar ficheros de onda (o fragmentos) a partir de mensajes MIDI, enviados, por ejemplo, desde un programa

secuenciador, con lo que es posible sincronizar un tema MIDI con fragmentos de audio real (voces, guitarras, etc.). *Sound Forge* es el único que permite esta operación, aunque los nuevos secuenciadores con audio digital incorporado, que se estudian en el apartado 18.7, "Sistemas multipista integrados en el ordenador", hacen que este mecanismo deje de ser necesario.

5.4. Clasificación de las herramientas de procesado digital de sonido

Las posibilidades de modificación del sonido digital almacenado en un fichero de ordenador son interminables. Para estudiarlas de una forma concisa, pero clara y comprensible, en lugar de seguir los menús de opciones de un determinado programa, intentaremos establecer una taxonomía que nos permita organizar estos efectos de una forma coherente, así como explicar someramente algunos de los principios básicos de procesado de señal involucrados. Una clasificación de los efectos más frecuentes sólo puede ser aproximada, ya que muchos de ellos requieren de una combinación de operaciones que los hace partícipes de varias categorías diferentes. Aun así, los dividiremos inicialmente en tres grandes grupos: los efectos que se aplican al tiempo, los que modifican la amplitud, y los que se aplican a la frecuencia y al timbre.

5.5. Efectos simples en el dominio temporal

En la mayoría de programas estos procesos suelen estar en la opción de menú **Edit**.

- **Copiar, cortar y pegar**, ocultan pocos secretos: se selecciona un fragmento con el ratón y se elimina o se inserta en otro lugar. En el estudio analógico (y en el cine), estas operaciones se han realizado durante décadas, cortando y pegando fragmentos de cinta magnética, Pero a pesar de esa aparente sencillez, en el dominio digital estas operaciones no deben realizarse sin ciertos cuidados. Para evitar clics y ruidos, cuando eliminamos o insertamos un fragmento, no deberíamos dejar una discontinuidad excesiva entre las nuevas muestras contiguas. La forma más sencilla de preservar esta continuidad es, tal como se muestra en la figura 5.4, seleccionar fragmentos con inicio y final nulos. Algunos programas se pueden configurar para que cuando seleccionemos un fragmento, la selección se redondee hasta los ceros más próximos. Una opción complementaria, que no suele estar en otros tipos de programas, es la de **Trim** (que podríamos traducir como **podar**), que consiste en eliminar todo excepto la zona seleccionada.

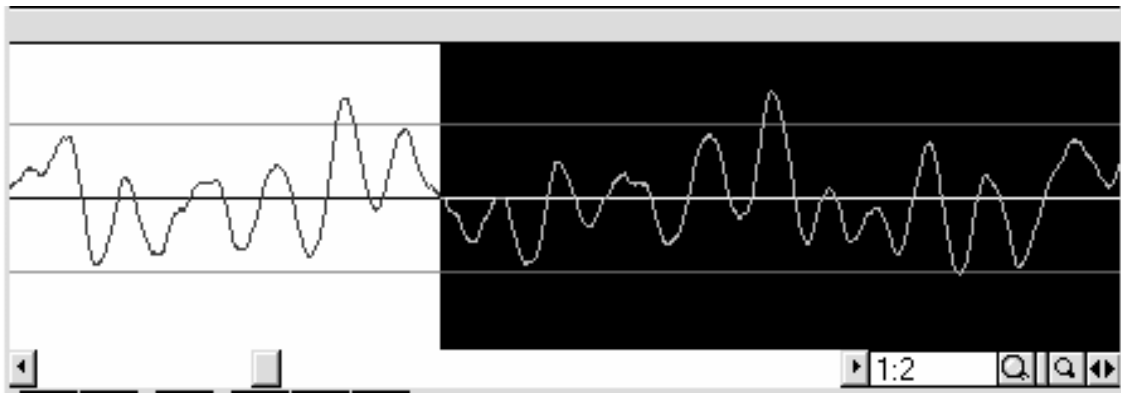


Figura 5.4. Inicio correcto de una selección

- **Reverse**, o darle la vuelta a un sonido, es también un efecto que se viene realizando en cinta magnética desde hace décadas. En la figura 5.5 se muestra un sonido y su inversión temporal.

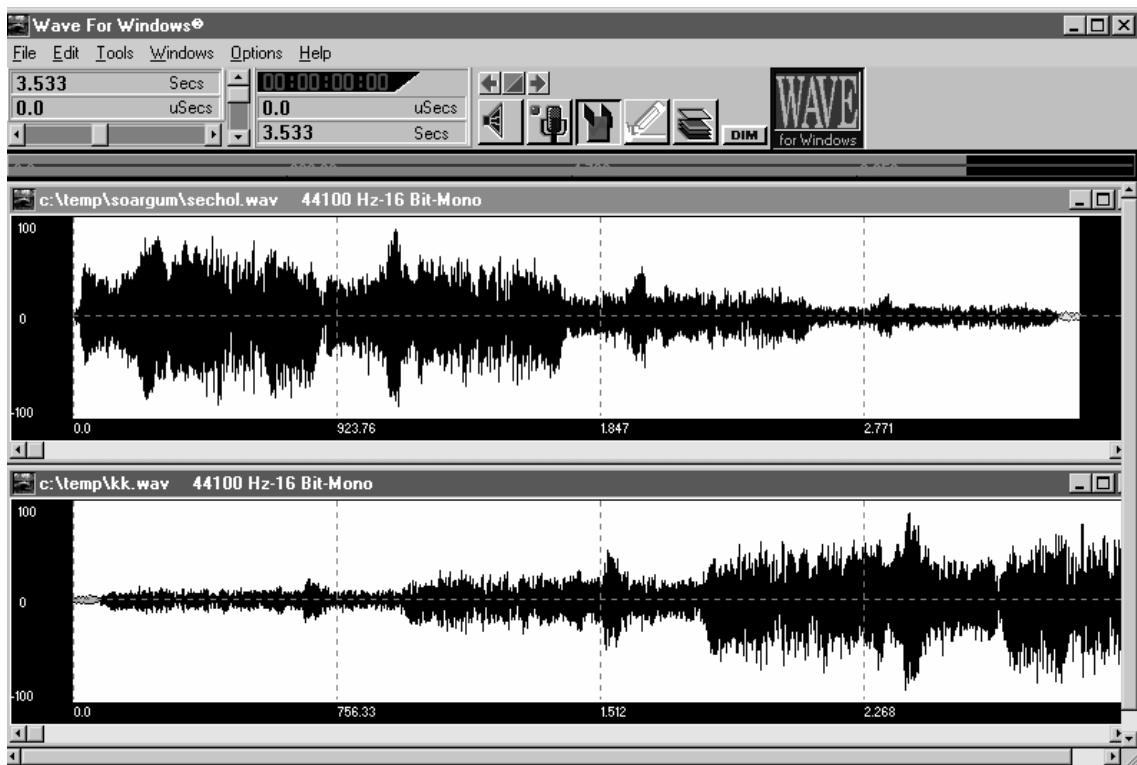


Figura 5.5. Inversión temporal de un sonido

- **Eliminar silencios**, borra los fragmentos silenciosos. Suele utilizarse para eliminar los fragmentos al inicio y final de una grabación. Dado que si el sonido se ha sometido a una conversión A/D sus silencios no serán nunca totalmente nulos, esta opción permite definir el valor de amplitud por debajo del cual el sonido se considera silencio. Tal como se indicaba en el apartado 5.3, *Sound Forge* ofrece la posibilidad de eliminar todos los silencios de un fichero y crear una lista de reproducción automática. De esta forma, el fichero resultante podrá tener un tamaño muy inferior pero sonará igual que el original.

- **Insertar silencio**, pregunta siempre la duración del fragmento a añadir, y lo coloca en la posición del cursor. Es frecuente insertar silencio al final de un fichero antes de aplicar un efecto de reverberación o de eco (véase apartado 6.2).

5.6. Efectos simples sobre la amplitud

Todos los efectos aquí descritos realizan diversas multiplicaciones a cada muestra de un fragmento.

- **Modificar ganancia**, consiste en multiplicar cada una de las muestras por un valor real. Si el valor está comprendido entre 0 y 1 el nivel sonoro disminuye, mientras que a partir de 1 aumenta. Se puede aplicar para potenciar sonidos que se han grabado con un nivel excesivamente bajo.
- **Silenciar**, consiste simplemente en multiplicar por cero la zona seleccionada.
- **Puerta de ruido** (*noise gate*), silencia las muestras por debajo de determinado valor umbral, introducido como parámetro. La figura 5.6 muestra el efecto de aplicar una puerta de ruido del 10% (silencia las muestras con una amplitud inferior al 10% del valor máximo posible). Este efecto permite eliminar el ruido de fondo, aunque como veremos más adelante, existen formas más sofisticadas de reducción de ruido, ya que este sistema sólo puede eliminar el ruido en los fragmentos en los que no hay música. También se ha utilizado mucho en los últimos años para procesar pistas de batería², con lo que se consiguen ataques y decaimientos más bruscos.

² A veces este efecto se conoce como “efecto Phil Collins”, que fue quien lo puso de moda.

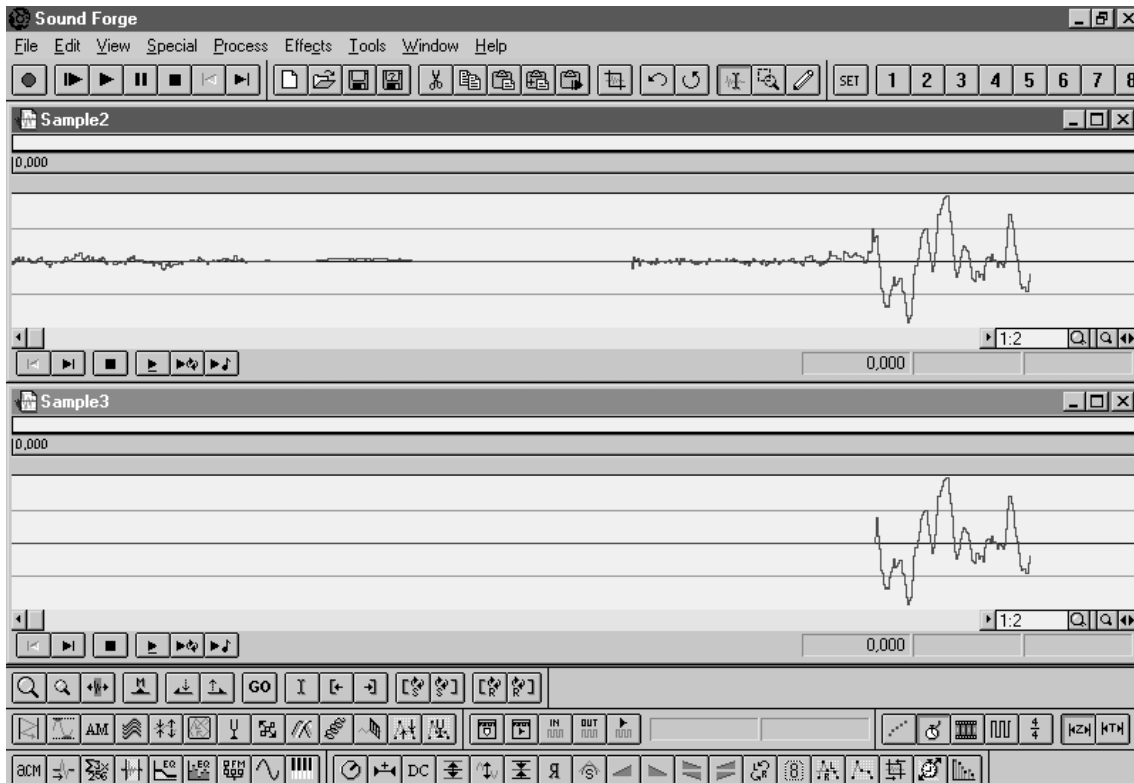


Figura 5.6. Un fragmento sonoro, antes y después de aplicar una puerta de ruido del 10%, en *Sound Forge*

- **Normalizar** es un caso particular de modificación de ganancia, que consiste en obtener la máxima amplitud posible sin que se produzca distorsión. Esto se consigue recorriendo la zona seleccionada y guardando su amplitud máxima. Una vez obtenida ésta, se multiplican todas las muestras del fragmento por el cociente *máxima amplitud posible / máxima amplitud del fragmento*. Si por ejemplo, la aplicación detectara un máximo de 12.345, todas las muestras de la zona seleccionada se multiplicarían por $2,654273 (32.767 / 12.345)^3$.
- **Aplicar envolventes.** Una envolvente es una curva que determina la evolución temporal de la amplitud (véase apartado 9.4.1). Estos programas permiten que el usuario dibuje envolventes con un número variable de puntos, y posteriormente realizan el producto de la envolvente y la señal seleccionada. En la figura 5.7, se muestran las ventanas de definición de envolventes en los programas *Cool Edit* y *Gold Wave*, y en la figura 5.8 el efecto obtenido al aplicar la primera de las dos envolventes a un sonido estacionario.
- **Fade in y fade out** son dos envolventes particulares, que normalmente cuentan como opciones de menú propias, por lo mucho que se utilizan. La primera (fundido de entrada), es una envolvente con valor inicial cero y valor final uno, que se aplica al inicio de un fragmento, mientras que la segunda (fundido de salida), se inicia con uno y termina con cero, y se aplica normalmente al final de un fragmento.

³ 32.767 es el máximo valor positivo en un sonido de 16 bits.

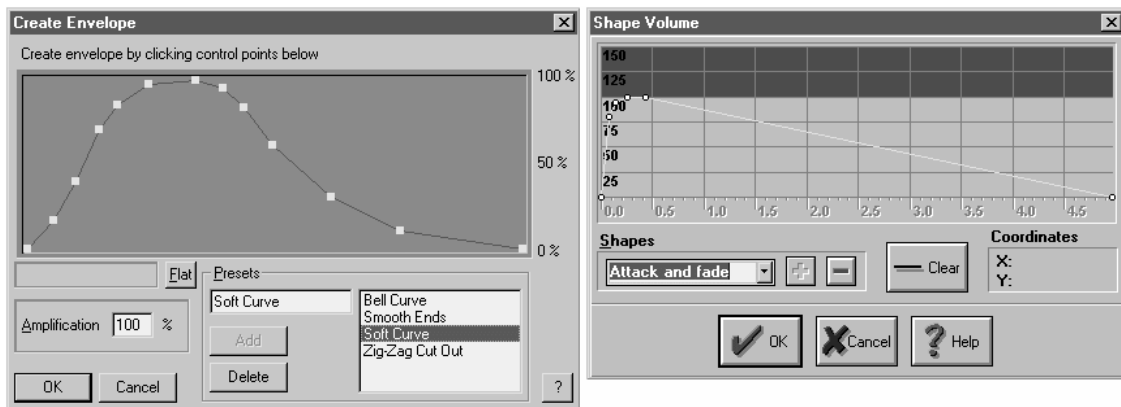


Figura 5.7. Definición de envolventes de amplitud en (a) *Cool Edity* (b) *Gold Wave*

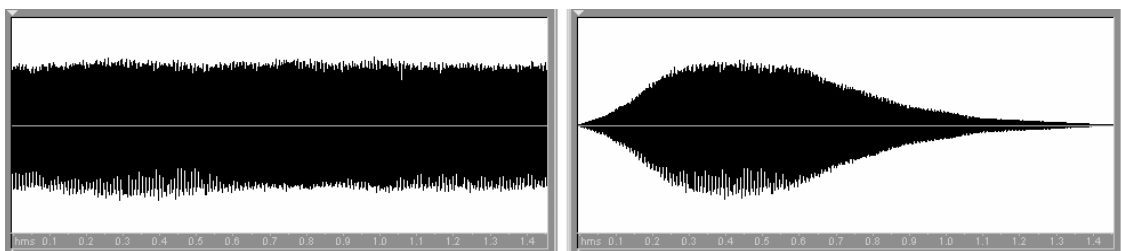


Figura 5.8. Fragmento sonoro antes y después de aplicarle la envolvente de la figura 5.7.a

- **La Modulación de amplitud**, consiste en multiplicar el fragmento por una señal periódica, normalmente sinusoidal, lo que en terminología musical se denomina *trémolo*. Los parámetros son frecuencia y amplitud de la modulación.
- **Invertir**, no debe confundirse con la opción **reverse**, indicada en el apartado 5.5, ya que con esta operación se realiza una reflexión respecto al eje horizontal. De esta forma los valores positivos pasan a ser negativos y viceversa, con lo se consigue un cambio de fase. Es un efecto muy sutil que se percibe mejor cuando la inversión se aplica a un único canal de un sonido estéreo.

A diferencia de los anteriores que son multiplicativos, los dos efectos restantes operan de forma aditiva.

- **Mezclar**, suma dos fragmentos (de un mismo fichero o de ficheros diferentes). Esta operación es normalmente accesible desde los menús como un caso especial de **pegar**, por lo que debe ir precedida por la acción de **copiar** uno de los dos fragmentos a mezclar. A continuación, y antes de ejecutar el comando, posicionaremos el cursor en el inicio de la zona destino. Esta operación se utiliza mucho para combinar varios sonidos.
- **DC Bias Offset**. Algunas tarjetas de sonido introducen un error permanente al digitalizar, que ocasiona un desplazamiento del cero. Puede comprobarlo grabando un fragmento de silencio y observando los resultados en un editor que ofrezca la suficiente resolución vertical. Si la línea recta correspondiente al fragmento de silencio grabado no se corresponde exactamente con la línea del cero, la tarjeta presenta un error de desplazamiento (que podrá ser positivo o negativo). Estos programas incorporan un

comando que calcula automáticamente el desplazamiento y corrige la señal, restándole este mismo valor.

5.7. Efectos simples aplicables a la frecuencia

Existen varias operaciones sencillas que comportan modificaciones frecuenciales, y la comprensión de todas ellas es fundamental para poder trabajar con soltura con el sonido digital.

5.7.1. Modificación de la frecuencia de muestreo

Supongamos que tenemos un sonido digitalizado a 44.100 Hz. Si le cambiamos la frecuencia de muestreo a 22.050 sin modificar el sonido en sí, el resultado sonará una octava por debajo. Esta operación equivale a reproducir una cinta o un disco de vinilo a velocidades diferentes de la original. Un LP de 33 r.p.m. reproducido a 45 r.p.m. sube la frecuencia en un 35 %⁴. Frecuencias superiores hacen que el sonido resultante sea más agudo, mientras que valores inferiores lo tornan más grave. En el ordenador, este efecto se consigue modificando tan sólo el valor de frecuencia en la cabecera del fichero, sin modificar para nada los datos. Esta operación conlleva además la modificación de la duración del sonido (a una frecuencia doble, el sonido durará la mitad). En los programas esta suele denominarse como **Change Playback Rate**.

5.7.2. Resample

Con este término, que podría traducirse por *remuestreo*, se designa la operación consistente en modificar la frecuencia de muestreo de un fichero, sin alterar la frecuencia del sonido. Esto se consigue normalmente eliminando o repitiendo algunas muestras. Para pasar por ejemplo, de 44.100 Hz a 22.050, se elimina directamente una muestra de cada dos, mientras que para el realizar el cambio inverso, cada muestra es duplicada. En realidad, para obtener una mayor calidad se realiza una interpolación, de forma que si una muestra vale 1000 y la siguiente 1020, la que se añade tomará el valor 1010 (cuando el cociente de las dos frecuencias no es un valor entero, las matemáticas involucradas se complican un poco más, pero el principio sigue siendo el mismo). Utilizaremos esta opción cuando queramos reducir el tamaño (y la calidad) de un fichero.

Cuando la nueva frecuencia sea inferior a la antigua, es conveniente filtrar el sonido *antes* de remuestrear, para evitar el *aliasing* (véase apartado 2.4). Algunos programas ya realizan por defecto este filtrado, pero en otros las dos operaciones deben realizarse por separado.

Cuando la nueva frecuencia sea superior a la anterior, no será necesario filtrar, pero tenga en cuenta que en este caso, tampoco se obtendrá ninguna mejora en la calidad del sonido, ya que no hay forma de reinventar la información que ya no existe (es como salvar con miles de colores una imagen de 16 colores). Aun así, esta operación puede ser necesaria por razones de compatibilidad entre programas o ficheros de sonido.

⁴ Si no le cuadran los cálculos es porque los discos a 33 están en realidad a 33 y 1/3, y por consiguiente $45/33,33 = 1,3500001$.

5.7.3. Transposición

El término musical *transponer* se utiliza para subir o bajar una nota o una melodía. En el tratamiento digital de sonido, *transponer* significa modificar la altura, pero manteniendo la frecuencia de muestreo. El resultado sonoro es similar al de la modificación de la frecuencia de muestreo del apartado 5.7.1, pero el efecto es en realidad una combinación de los dos anteriores. El valor de la transposición se puede indicar como un porcentaje, o bien como un intervalo musical (subir dos semitonos, etc.). Estas operaciones comportan también la variación de la duración del sonido (más breve cuanto más agudo, más largo cuanto más grave).

5.7.4. Pitch Bend

Este efecto es una generalización de la transposición, ya que permite transponer el sonido de forma continua, a partir de la definición de una envolvente de frecuencia (cuando esta envolvente sea una línea recta, tendremos la transposición del anterior apartado). Musicalmente, el efecto obtenido se conoce como *glissando* y es el que se obtiene por ejemplo, al desplazar rápidamente la mano izquierda sobre las cuerdas de un violín o una guitarra.

5.7.5. Recapitulación

Algunas de estas operaciones pueden llevarse a cabo mediante algoritmos alternativos que ofrecen una mayor calidad, pero que son también más complicados de explicar y de comprender. Hemos preferido describirlas de la forma más sencilla posible ya que son operaciones fundamentales, que ayudan a entender la naturaleza del sonido digital.

En los diversos programas de edición, estas operaciones no siempre se presentan de la misma manera. *Gold Wave* incluye, tal como se muestra en la figura 5.9.a, exactamente las tres opciones de menú, pero *Cool Edit* las presenta todas en una única caja de diálogo bastante más críptica.

Conviene indicar, por último, que mediante otros tipo de recursos matemáticos es también posible modificar la altura de un sonido sin modificar su duración y, al contrario, modificar la duración preservando la altura. En el próximo capítulo, en el que seguiremos estudiando las posibilidades del proceso digital de señal, daremos más información al respecto.

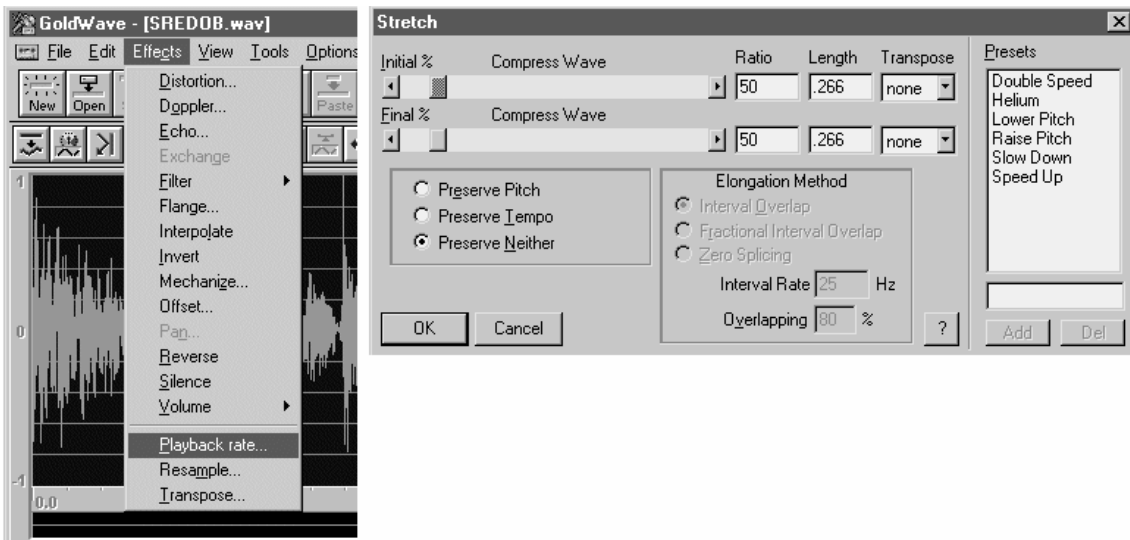


Figura 5.9. Acceso a las modificaciones frecuenciales en *Gold Wave* (a) y en *Cool Edit* (b)