

121
cop.

ACÚSTICA
Y PSICOACÚSTICA
DE LA MÚSICA

JUAN G. ROEDERER

16/12

- 121 cop -

RICORDI

Título original de la obra
THE PHYSICS AND PSYCHOPHYSICS OF MUSIC
An Introduction

Traducción del inglés de
Guillermo D. Pozzati

5A/0A
- 900 15A -

A Beatriz

compañera de estudios
compañera de trabajo
compañera de mi vida

Este libro se terminó de imprimir el 17 de abril de 1997
en los Talleres Gráficos M. A. Bermejo
11 de Septiembre 539 Haedo - Buenos Aires.

ISBN 950-22-0444-1 RICORDI AMERICANA S.A.E.C.

Originalmente publicado en inglés bajo el título *The Physics and Psychophysics of Music*.

© Copyright 1995, Springer-Verlag New York, Inc.
All rights reserved.

© Copyright 1997 by RICORDI AMERICANA S.A.E.C.

Tte. Gral. Juan D. Perón 1558 - Buenos Aires.

Editor exclusivo de la traducción al español para todos los países de América hispano-
parlante, Brasil, España y Portugal.

Todos los derechos están reservados - All rights reserved.

Queda hecho el depósito que establece la Ley 11.723.

Indice

Prefacio a la edición en español	6
1 Música, física, psicofísica y neuropsicología: un enfoque interdisciplinario	9
1.1 Los sistemas físicos intervinientes	9
1.2 Atributos característicos de los sonidos musicales	11
1.3 El elemento temporal en música	13
1.4 Física y psicofísica	15
1.5 Psicofísica y neuropsicología	19
1.6 ¿Qué es la música?	21
2 Vibraciones sonoras, tonos puros y percepción de la altura	24
2.1 Movimiento y vibración	24
2.2 Movimiento armónico simple	28
2.3 Vibraciones acústicas y sensaciones de tonos puros	30
2.4 Superposición de tonos simples: batidos de primer orden y la banda crítica	37
2.5 Otros efectos de primer orden: sonidos resultantes y armónicos aurales	46
2.6 Efectos de segundo orden: batidos de consonancias desafinadas	50
2.7 Seguimiento de la fundamental [Fundamental Tracking]	53
2.8 Codificación auditiva en el sistema nervioso periférico	60
2.9 Altura subjetiva y el rol del sistema nervioso central	67
3 Ondas sonoras, energía acústica y percepción de la sonoridad	79
3.1 Ondas elásticas, fuerza, energía y potencia	79
3.2 Velocidad de propagación, longitud de onda y potencia acústica	83
3.3 Superposición de ondas; ondas estacionarias	94
3.4 Intensidad, nivel de intensidad sonora y sonoridad	98

Indice

3.5 El mecanismo de percepción de la sonoridad y procesos relacionados	110
3.6 Música desde los oídos: emisiones otoacústicas y mecanismos cocleares	113
4 Generación de sonidos musicales, tonos compuestos y percepción del timbre	120
4.1 Ondas estacionarias en una cuerda	121
4.2 Generación de ondas estacionarias complejas en instrumentos de cuerda	125
4.3 Espectros sonoros y resonancia	135
4.4 Ondas estacionarias longitudinales en una columna de aire ideal	144
4.5 Generación de vibraciones estacionarias complejas en instrumentos de viento	148
4.6 Espectros sonoros de instrumentos de viento	156
4.7 Atrapamiento y absorción de ondas sonoras en ambientes cerrados	159
4.8 Percepción de la altura y el timbre de tonos musicales	163
4.9 Identificación de sonidos musicales	169
4.10 Procesos cognitivos relevantes a la percepción de tonos individuales	171
5 Superposición y sucesiones de tonos compuestos y la percepción musical	180
5.1 Superposición de tonos compuestos	180
5.2 Sensación musical de consonancia y disonancia	183
5.3 Construyendo escalas musicales	190
5.4 La escala standard y el standard de altura	195
5.5 ¿Por qué existen las escalas musicales?	198
5.6 Procesos cerebrales cognitivos y afectivos en la percepción musical: ¿por qué respondemos emocionalmente a la música?	201
5.7 Especialización del habla y del procesamiento musical en los hemisferios cerebrales	206
Apéndice I: Algunos aspectos cuantitativos del mecanismo de frotamiento	212
Apéndice II: Algunos aspectos cuantitativos de los modelos del procesador central de altura tonal	216
Apéndice III: Algunas observaciones acerca de la enseñanza de la física y la psicofísica de la música	226
Referencias	229
Indice alfabético	235
Acerca del autor	240

Prefacio a la edición en español

Muchos músicos temen a los físicos - dicen que no los entienden porque éstos sólo piensan en términos de complicadas fórmulas y desalmadas leyes físicas. Muchos físicos temen a los psicólogos - dicen que no los entienden porque hablan un idioma florido, usando términos de imprecisa definición y difícil representación cuantitativa. Y muchos psicólogos temen a los neurobiólogos, porque estos últimos están encontrando explicaciones «lógicas» a los misterios del comportamiento en base al cableado de las redes neurales y los mecanismos de procesamiento de información en el cerebro.

¡Pero nadie le teme a los músicos! Aprovechando este afortunado hecho, el propósito fundamental de este libro es *mirar* a la física, la psicoacústica y la neuropsicología en una única familia de ciencia interdisciplinaria, usando la música como enfoque y gran reconciliador. Es mi sincero deseo que el resultado de esta reunión de disciplinas, o «*Ciencia de la Música*», satisfaga las inquietudes de los lectores acerca de lo que «en realidad» es la música, y que les provea información actualizada sobre los procesos físicos de generación de tonos en instrumentos musicales, la propagación de ondas sonoras a través del medio ambiente, su detección en el oído y los mecanismos pertinentes de análisis, interpretación y respuesta emotiva en el cerebro. A los músicos en particular, espero que el estudio de este libro también los estimule a apreciar lo que podemos llamar la «*Música de la Ciencia*», o sea, la belleza, armonía, excitación y asombro evocados por el conocimiento científico.

Hice un gran esfuerzo tratando de usar el lenguaje preciso de los físicos, sus métodos y maneras de pensar, sin por ello complicar la presentación con mucha matemática. Esto, por supuesto, impuso ciertos límites a los temas a tratar, obligándome ocasionalmente a sobresimplificar peligrosamente una que otra explicación científica. En general, los conocimientos adquiridos en la escuela secundaria son suficientes para entender la mayor parte de este texto. Lo que sí debo suponer es que el lector esté familiarizado con los conceptos básicos de la música, tales como escalas, acordes, la serie armónica, los varios tipos de instrumentos musicales, y las principales sensaciones tonales.

La primera edición de este libro, escrita en inglés, está basada en apuntes para un curso del ciclo básico, dictado en la Universidad de Denver (Colorado, EE.UU.) a principios de los setenta, titulado «Física de la Música». Muy pronto, este curso se convirtió en interdisciplinario, incorporando los conceptos psicoacústicos y neuropsicológicos relacionados con la percepción de la música. La tercera edición en inglés, publicada en 1995, responde a una revisión fundamental de todas las secciones neurobiológicas, poniéndolas al día. Efectivamente, en los últimos veinte años hubo un progreso inusitado en las investigaciones científicas sobre percepción acústica y musical.

El presente libro es una traducción fiel de la tercera edición inglesa, efectuada con mucha habilidad por un joven compositor argentino, especialista en música electrónica, quien ya estaba familiarizado con el texto inglés por haberlo usado en su dictado de clases universitarias y de conservatorio. El inglés es hoy día, quiérase o no, el idioma internacional de la ciencia. Es un idioma muy conciso, a veces hasta el punto de parecer comparativamente primitivo. En mi opinión como persona trilingüe, la traducción del español (o del alemán) al inglés es, en general, mucho más fácil que una traducción del inglés a los otros idiomas. La tarea se complica considerablemente para textos de carácter interdisciplinario como el presente, en que convergen vocabularios de las ciencias naturales, ciencias del comportamiento y bellas artes.

Una particular dificultad proviene de la existencia de muchos términos que no tienen traducción *exacta* y *unívoca* al español. El más irritante de éstos es la palabra «pattern», que en este libro aparece por doquier. Hemos decidido usar el término *patrón* cuando «pattern» aparece en un contexto «científico»—¡bien conscientes de que, en realidad, «patrón» tiene significados muy diferentes para físicos, neuropsicólogos y peones, respectivamente! Otro problema lo presenta el uso de ciertos términos que son sinónimos en el lenguaje diario, pero que en ciencia poseen significados diferentes. Un ejemplo es «sonido» y «tono» (¡todos los tonos son sonidos, pero no todos los sonidos son tonos! Además, también se usa «tono» para designar un intervalo musical como, por ejemplo, en «semitono»).

Las citas bibliográficas dadas en esta edición en español son aquellas que figuran en el texto original, y por lo tanto corresponden a la literatura publicada en inglés. Sin duda, se me habrán pasado por alto artículos o libros escritos en español que no han llegado a mi conocimiento en estos «pagos lejanos» - desde ya, pido disculpas a sus autores por no haberlos incluido. También debo aclarar que, en vista del carácter multidisciplinario del texto, he dado prioridad a citas en revistas de mayor alcance en el mundo de habla inglesa (tales como *Scientific American*, o el *Journal of the Acoustical Society of America*), sobre aquellas publicadas en revistas muy especializadas. En inglés, existen excelentes libros, más o menos

recientes, sobre acústica musical (p. ej., Benade, 1976; Pierce, 1983; Sundberg, 1991), psicoacústica (p. ej., Zwicker y Fastl, 1990) y psicología musical (p. ej., Deutsch, 1981). El presente volumen no pretende duplicar, sino sintetizar y complementar la literatura existente; su propósito es brindar una presentación uniforme, integral e interdisciplinaria, a un público lo más amplio posible.

Uno de los actos más penosos al escribir un libro es decidir qué tópicos necesariamente deberán ser dejados de lado. Por más concienzuda que sea esta decisión, siempre habrá alguien que proteste por la exclusión de un tema tal o cual. Los siguientes son tópicos omitidos o tratados sólo superficialmente (sin intentar aquí justificación alguna). En la discusión de la generación de tonos en instrumentos musicales, sólo tratamos los mecanismos más básicos, sin entrar en detalles concretos de instrumentos reales. La voz humana - ¡el primero, y quizás más sublime instrumento musical! - se ha omitido por completo, así como toda discusión de tonos inarmónicos de campanas e instrumentos de percusión. En psicoacústica, nos limitamos a la discusión de la percepción de tonos, o de superposiciones de tonos *sinusoidales* (tonos simples), dejando de lado completamente los muchos experimentos realizados con trenes de pulsos y bandas de ruido.

Finalmente, prácticamente no hay referencias al rol fundamental del ritmo en la música, así como a cuestiones de desarrollo histórico de las disciplinas intervinientes.

No se puede enseñar a nadar estando parado, con tiza en mano, frente a una pizarra. Por la misma razón, no se puede enseñar acústica y psicoacústica realmente bien, sin «sumergir» al estudiante en un laboratorio adecuadamente equipado. Lamentablemente, esto requiere instrumentos electrónicos que no son de fácil disponibilidad, ni siquiera en la mayoría de los laboratorios de física universitarios. Siempre que fuera posible, he tratado de describir «experimentos» realizables con medios usuales (instrumentos como el piano, órgano, flauta); cuando hago referencia a experimentos reales, ¡sólo puedo esperar que el lector *me crea* que los efectos descriptos son los que efectivamente ocurren en la realidad!

Ya una vez, en el pasado remoto, procedí a convertir unos apuntes de clase en un librito de texto (*Mecánica Elemental*, EUDEBA, Buenos Aires, 1962). Ese volumen se sigue usando hoy día en Latinoamérica, habiendo sido reimpreso ya varias veces. Estoy curioso: el presente libro, ¿tendrá la misma aceptación?

Juan G. Roederer
Fairbanks, Alaska, febrero de 1997

1

Música, física, psicofísica y neuropsicología: un enfoque interdisciplinario

«Quien no comprende otra cosa que la química, no llega a comprender verdaderamente la química tampoco»

Georg Christoph Lichtenberg (1742-1799)

1.1 Los sistemas físicos intervinientes

Imagínese sentado en una sala de conciertos escuchando a un solista tocar su instrumento. Veamos cuáles son los sistemas que hacen posible que usted oiga la música que se está tocando. En primer lugar, obviamente, tenemos al ejecutante y al *instrumento* que «hace» la música. En segundo lugar tenemos el *aire* de la sala que transmite el sonido en todas las direcciones. En tercer lugar está usted, el *oyente*. En otras palabras, tenemos la siguiente cadena de sistemas: instrumento → aire → oyente. ¿Qué cosa une a estos sistemas mientras la música se está tocando? Vibraciones de un cierto tipo y forma llamadas *sonido* que se propagan de un punto a otro en forma de *ondas* y a las cuales nuestro oído es sensible. (Hay muchos otros tipos y formas de ondas que no podemos detectar en absoluto, o que podemos detectar, pero con otros sentidos, tales como el tacto o la vista).

El físico emplea términos más generales para describir los tres sistemas arriba mencionados. Los llama *fuerza* → *medio* → *receptor*. Esta cadena de sistemas es común al estudio de muchos otros procesos físicos: luz, radiactividad, rayos cósmicos, etc. La *fuerza* *emite*; el *medio* *transmite* y el receptor *detecta*, registra o es en general afectado de alguna manera determinada. Aquello que es emitido, transmitido y detectado es energía en alguna de sus múltiples formas, que dependerá del caso particular considerado. En el caso de las ondas sonoras es energía elástica, porque implica oscilaciones de la presión, es decir, una rápida sucesión de compresiones y expansiones del aire.¹

Examinemos un poco más de cerca a los sistemas involucrados. En la

¹ El sonido, por supuesto, también se propaga a través de líquidos y sólidos.

fuerza, es decir el instrumento musical, identificamos varios componentes diferentes: 1) El *mecanismo primario de excitación* que debe ser activado por el instrumentista², tal como la acción de puntear con el dedo o frotar con el arco la cuerda de un violín, la caña oscilante en un clarinete, los labios del ejecutante en un instrumento de la familia de los metales o la columna de aire soplada contra un borde biselado en la flauta. Este mecanismo de excitación actúa como la fuente primaria de energía. 2) El *elemento vibrante* fundamental que, al ser excitado por el mecanismo primario, es capaz de sostener ciertos modos de vibración bien definidos y de frecuencias prefijadas, tales como las cuerdas de un violín, o la columna de aire en un instrumento de viento o tubo de órgano. Este elemento vibrante es el que realmente determina la altura del tono y, por añadidura, suministra los armónicos superiores necesarios para impartir cierta cualidad característica o timbre al mismo. Además, sirve como almacenamiento de energía vibratoria. En los instrumentos de viento controla en parte al mecanismo primario de excitación a través de un proceso de realimentación (fuerte en las maderas, débil en los cobres). 3) Muchos instrumentos tienen un *resonador* adicional (tabla armónica de un piano, caja de un instrumento de cuerdas), cuya función es convertir más eficientemente las oscilaciones del elemento vibrante primario (cuerda) en vibraciones sonoras del aire circundante y dar al tono su timbre final.

En el medio, también debemos hacer una distinción: tenemos el *medio propiamente dicho* que transmite el sonido, y los *contornos*, es decir, las paredes, el techo, el piso, la gente del auditorio, etc., que afectan sustancialmente la propagación sonora por medio de la *reflexión* y la *absorción* de las ondas sonoras y cuya configuración determina la calidad acústica de la sala (reverberación).

Finalmente, en el oyente distinguimos los siguientes componentes principales: 1) El *tímpano*, que recoge las oscilaciones de presión de la onda sonora que llega al oído y las convierte en vibraciones mecánicas que son transmitidas por vía de tres huesecillos a: 2) El *oído interno*, o cóclea, en la cual las vibraciones son ordenadas según rangos de frecuencia, detectadas por células receptoras y convertidas en impulsos eléctricos nerviosos. 3) El *sistema nervioso auditivo*, que transmite las señales nerviosas al cerebro, donde la información es procesada, desplegada como una imagen de rasgos auditivos sobre cierta área de la corteza (la superficie del cerebro y tejido subyacente), identificada, almacenada en la memoria, y eventualmente transferida a otros centros del cerebro. Estas últimas etapas conducen a la percepción consciente de los sonidos musicales.

²Para hacer una descripción completa deberíamos agregar los «componentes» del ejecutante: la corteza motora de su cerebro, de donde surgen las órdenes a sus músculos, las partes de su cuerpo que activan el instrumento o su tracto vocal, la interacción entre sus oídos y músculos que le ayudan a controlar la ejecución, etc. Pero, lamentablemente, limitaciones de espacio nos obligan a dejar al ejecutante completamente de lado.

Tabla 1.1 Sistemas físicos y biológicos inherentes a la música, y sus funciones correspondientes.

	Sistema	Función
Fuente	Mecanismo de excitación	Suministro de energía
	Elemento vibrante	Determinación de las características fundamentales del tono
	Resonador	Conversión a oscilaciones de la presión del aire (ondas sonoras), determinación final de las características del tono
Medio	Medio propiamente dicho	Propagación del sonido
	Límites	Reflexión, absorción, reverberación
Receptor	Tímpano	Conversión a oscilaciones mecánicas
	Oído interno	Codificación primaria de la frecuencia, conversión a impulsos neurales
	Sistema nervioso	Procesamiento, imaginación, identificación, almacenamiento, y transferencia a otros centros cerebrales

Obsérvese que podemos reemplazar al oyente por un *dispositivo de grabación* tal como una cinta magnética de grabador, un disco digital, o la grabación fotoeléctrica sobre una película, y aun así reconocer por lo menos tres de los subsistemas: la detección mecánica y posterior conversión en señales eléctricas en el micrófono, un limitado procesamiento, accidental o deliberado, en el circuito electrónico, y el almacenamiento en una memoria, sobre cinta, disco o película respectivamente. El primer sistema, es decir el instrumento, puede ser, desde luego, reemplazado por un *dispositivo de reproducción*.

Podemos resumir esta discusión en la tabla 1.1.

El principal objetivo de este libro es analizar comprensivamente lo que ocurre en cada una de las etapas mostradas en la tabla 1.1 y durante cada una de las transiciones de una etapa a la siguiente, cuando se ejecuta y percibe música.

1.2 Atributos característicos de los sonidos musicales

Individuos de todas las culturas concuerdan en reconocer la existencia de tres sensaciones primarias asociadas a un sonido musical dado: *altura*, *sonoridad* y *timbre*³. No intentaremos definir estos atributos subjetivos o

³Las sensaciones a veces citadas de «volumen» y «densidad» (o brillo) son conceptos compuestos que pueden ser «resueltos» en una combinación de efectos de altura y sonoridad (bajar la altura con incremento simultáneo de sonoridad nos lleva a una sensación de incremento de volumen; subir la altura con incremento simultáneo de sonoridad nos lleva a incremento de densidad o brillo).

magnitudes psicológicas, ni involucramos en esta etapa en la discusión respecto de si se trata de cantidades mensurables; solo nos limitaremos a observar que la altura es descripta frecuentemente como la sensación de «altura» o «elevación», y la sonoridad como la «fuerza» o «intensidad» de un tono. El timbre, o cualidad, es aquello que nos permite distinguir tonos de instrumentos diferentes aun cuando tengan igual altura e igual sonoridad. La asociación inequívoca de estas tres cualidades a un sonido dado es lo que diferencia a un «sonido musical» de un «ruido», pues si bien podemos indudablemente asignar una sonoridad a un ruido, es mucho más difícil identificar en él una altura o cualidad tímbrica definida*.

La asignación de altura, sonoridad y timbre a un sonido musical es el resultado de operaciones de procesamiento en el oído y en el cerebro. Esta asignación es subjetiva e inaccesible a la medición física directa (véase Sec. 1.4). Sin embargo, cada una de estas sensaciones primarias puede ser asociada en principio a una cantidad física bien definida del estímulo original, es decir, de la onda sonora, que puede ser medida y expresada numéricamente por métodos físicos. En realidad, la sensación de altura está asociada primariamente con la *frecuencia fundamental* (frecuencia de repetición del patrón vibratorio, descripta por el número de oscilaciones por segundo); la sonoridad con la *intensidad* (flujo de energía o amplitud de oscilación de presión de la onda que llega al oído), y el timbre con el *espectro*, o proporción con la cual otras frecuencias más agudas, llamadas «armónicos superiores», aparecen acompañando a la frecuencia fundamental.

Esta, sin embargo, es una descripción muy simplificada. En primer lugar, la sensación de sonoridad de un tono de intensidad constante parecerá variar si cambiamos la frecuencia. En segundo término, la sonoridad de una superposición de varios sonidos, cada uno con una altura diferente (por ej. un acorde), ya no está relacionada simplemente con el total del flujo de energía sonora; por otra parte, en una sucesión de sonidos de muy corta duración (por ej. sonidos 'staccato'), la sonoridad depende de cuánto dure realmente cada sonido. En tercer lugar, la percepción refinada del timbre, tal como la que se requiere para el reconocimiento de los instrumentos musicales, es un proceso que utiliza mucha más información que la dada por el espectro de un sonido; las características transientes del ataque y de la caída de un tono son igualmente importantes, cómo puede verificarse fácilmente intentando reconocer instrumentos musicales al escuchar una cinta magnética reproducida al revés. Lo que es más, los sonidos de un instrumento dado pueden tener características espectrales que cambian apreciablemente a lo largo de la extensión del

* En lo que sigue reservamos el término «tono» para designar un sonido musical de altura, sonoridad y timbre definidos. «Sonoridad» es el término que corresponde a la palabra inglesa «loudness».

instrumento, y la composición espectral de un sonido determinado puede cambiar considerablemente de un punto a otro en una sala de concierto; sin embargo, los sonidos correspondientes serán reconocidos sin titubeo como pertenecientes al mismo instrumento. Inversamente, un músico experto puede tener grandes dificultades al intentar determinar la altura exacta de un tono puro generado electrónicamente, desprovisto de armónicos superiores y escuchado mediante auriculares, porque su sistema nervioso central carecerá de algunas claves adicionales de información que normalmente vienen con los sonidos «reales» con los cuales está familiarizado.

Otra característica física relevante de un sonido es la dirección espacial en la que llega la onda sonora. Lo que importa aquí es la minúscula diferencia temporal entre las señales acústicas detectadas en cada oído, la cual depende de la dirección de incidencia de la onda. Esta diferencia temporal es medida y codificada por el sistema nervioso para producir la sensación de *direccionalidad sonora* (estereofonía o lateralización).

Cuando dos o más tonos suenan simultáneamente, nuestro cerebro, dentro de ciertos límites, es capaz de diferenciarlos individualmente. Aquí aparecen sensaciones subjetivas nuevas, menos definidas pero de importancia musical, que se relacionan con la superposición de dos o más sonidos y conducen al concepto de *armonía*. Entre éstas se hallan las sensaciones «estáticas» de *consonancia* y *disonancia* que describen el carácter -sea «placentero» o «irritante»- de ciertas superposiciones de sonidos, la sensación dinámica de *urgencia a resolver* un intervalo disonante dado o acorde, el peculiar efecto de los *batidos* y el diferente carácter de los acordes *mayores* y *menores*. Mientras que la correlación de altura, sonoridad y, en alguna medida, timbre con ciertas características físicas de tonos aislados es «universal»- es decir, independiente del condicionamiento cultural de un individuo dado -, no es éste el caso con los atributos subjetivos arriba mencionados de las superposiciones de tonos.

1.3 El elemento temporal en música

Un tono que no variase, sonando prolongadamente con la misma frecuencia, intensidad y espectro, resultaría molesto. Aun más, después de un rato nuestra conciencia no lo registraría más. Solo cuando ese sonido desapareciera nos daríamos cuenta repentinamente de que había estado sonando. La música está hecha de tonos cuyas características físicas cambian de alguna manera a medida que transcurre el tiempo. Es solo esta dependencia temporal lo que hace que un sonido sea «musical» en el verdadero sentido del término.

En general, de aquí en adelante llamaremos a una secuencia de sonidos individuales un *mensaje musical*.

Ese mensaje musical puede ser «significativo» (a veces se habla de una «Gestalt» sonora) si le asignamos cierto «valor» como resultado de una serie de operaciones cerebrales de análisis, almacenamiento en la memoria, comparación con mensajes previamente almacenados y asociaciones. Una *melodía* es el ejemplo más simple de mensaje musical. Algunos atributos de mensajes musicales significativos son elementos clave en la música occidental: *tonalidad* (predominio de un tono sobre los otros en una secuencia), *sentido de retorno* a la tónica, *modulación* y *ritmo*. Una característica fundamental que se presenta en melodías de las más diversas culturas es que los tonos que la forman proceden en saltos discretos, finitos, de altura. Del infinito número de frecuencias disponibles, el sistema nervioso prefiere seleccionar valores discretos que correspondan a las notas de una escala musical, aun cuando somos capaces de distinguir cambios de frecuencia mucho más pequeños que aquellos representados por el intervalo mínimo de cualquier escala musical. El mecanismo nervioso que analiza un mensaje musical presta atención sólo a las *transiciones* de altura. El procesamiento «absoluto» de la altura (oído absoluto) se pierde a edad temprana en la mayoría de los individuos.

Examinemos más de cerca el elemento temporal en música. Existen tres rangos temporales diferentes en los cuales ocurren variaciones de importancia para la psicoacústica. En primer lugar, tenemos una escala temporal «microscópica», en la que ocurren las vibraciones mismas de las ondas sonoras, que abarca un rango de períodos que van desde 0,00007 a 0,05 s. Luego tenemos una escala «intermedia» centrada alrededor de un décimo de segundo, en la cual tienen lugar cambios transientes tales como los representados por el ataque o la caída de un tono, que representan las variaciones de tiempo de los rasgos microscópicos. Finalmente, tenemos una escala de tiempo «macroscópica», que va desde 0,1 s, aproximadamente, hacia arriba, donde ocurren las duraciones de los tonos musicales, las sucesiones y el ritmo. Es importante observar que a cada una de estas tres escalas de tiempo corresponde un particular «centro de procesamiento» con una función específica en el sistema auditivo. 1) Las vibraciones microscópicas son detectadas y codificadas en el *oído interno* y conducen principalmente a las sensaciones primarias del sonido (altura, sonoridad y timbre). 2) Las variaciones intermedias o transientes parecen afectar principalmente mecanismos de procesamiento situados en el *trayecto neural* que va desde el oído hasta el área auditiva del cerebro y proveen claves adicionales utilizadas en la discriminación, identificación y percepción de la cualidad de los sonidos. 3) Los cambios temporales macroscópicos son procesados al nivel más alto del sistema nervioso - la *corteza cerebral*, la superficie con pliegos y el tejido subyacente -; ellos determinan el real mensaje musical y sus atributos. Cuanto más avanzamos a través de estas etapas de procesamiento en la vía neural auditiva, tanto

más difícil se hace definir e identificar los atributos psicológicos a los cuales estos procesamientos conducen, y tanto más el resultado aparece influenciado por el aprendizaje y el condicionamiento cultural, así como también por el estado psíquico momentáneo del individuo.

Por más de cien años los musicólogos se han quejado amargamente porque la física de la música y la psicoacústica habían estado restringidas principalmente al estudio de tonos o conjuntos de tonos constantes, invariables, mientras que la esencia de la música es una secuencia temporal de todo ello. Sus quejas están bien fundadas, pero las razones para tal restricción también lo están. Como explicamos más arriba, el procesamiento de secuencias sonoras ocurre al más alto nivel del sistema nervioso central - abarcando una cadena de mecanismos todavía poco explorada. En este libro intentaremos en lo posible remediar esta situación e ir tan lejos como podamos para disipar algo de estas bien justificadas críticas.

1.4 Física y psicofísica

Podemos describir el objetivo principal de la física de la siguiente manera: proveer métodos por medio de los cuales se pueda predecir cuantitativamente la evolución de un sistema físico dado (o «develar» su historia), basándonos en las condiciones en que el sistema se encuentra en un momento dado⁴. Por ejemplo, dado un automóvil de cierta masa y especificando las fuerzas de los frenos, la física nos permite predecir cuánto tiempo le llevará al auto detenerse y en qué lugar lo hará, siempre y cuando especifiquemos la posición y la velocidad en el instante inicial. Dada la masa, así como el largo y la tensión de una cuerda de violín, la física predice las posibles frecuencias con las cuales la cuerda vibraría si fuera pulsada o frotada con un arco de una manera determinada. Dadas la forma y las dimensiones de un tubo de órgano y la composición y temperatura del gas en su interior, la física predice la frecuencia fundamental del sonido emitido cuando se sopla el tubo.

En la práctica, «predecir» significa suministrar un aparato matemático, una serie de ecuaciones o «recetas» que, basadas en ciertas *leyes físicas* que gobiernan el sistema bajo análisis, establecen relaciones matemáticas entre los valores de las magnitudes físicas que caracterizan al sistema en cualquier instante de tiempo (posición y velocidad en el caso del auto; frecuencia y amplitud de oscilación en los otros dos ejemplos). Estas relaciones son usadas para determinar cómo cambian los valores a medida que transcurre el tiempo.

⁴El objetivo de la física es a veces citado como «El logro de una explicación cuantitativa del universo». Esto, sin embargo, es a lo sumo una jerga filosófica que poco tiene que ver con el real (mucho más preciso, realista y modesto) objetivo de la física.

Con el fin de establecer las leyes físicas que rigen un sistema dado, debemos primero observar el sistema y hacer mediciones cuantitativas de las magnitudes físicas relevantes para descubrir experimentalmente sus interrelaciones causales. Una ley física expresa una cierta relación que es común a muchos sistemas físicos diferentes y que es independiente de circunstancias particulares. Por ejemplo, la ley que rige la gravitación es válida aquí en la Tierra, en la Luna, en el sistema solar y en cualquier otra parte del universo. Las leyes del movimiento de Newton se aplican a todos los cuerpos, sin importar su composición química, color, temperatura, velocidad o posición.

La mayor parte de los sistemas reales que estudia la física, - incluso los ejemplos «simples» y «familiares» dados anteriormente - son tan complejos que se hace imposible la formulación de predicciones exactas y detalladas. Por lo tanto debemos hacer aproximaciones e idear *modelos* que representen al sistema dado en sus rasgos principales. ¡El ubicuo «punto material» al cual frecuentemente un cuerpo es reducido en física - sea éste un planeta, un automóvil o un electrón - es el modelo más simplificado de todos! Muchas veces es necesario descomponer el sistema bajo estudio en una serie de subsistemas más elementales, que interactúan físicamente entre sí, cada uno gobernado por un conjunto de leyes físicas bien definidas.

La física de la vida diaria o *física clásica* presupone que tanto las mediciones como las predicciones deberían ser siempre «exactas» y «únicas»; las limitaciones y errores que pudieran surgir serían debidos únicamente a la imperfección de nuestros instrumentos de medición. En el dominio atómico y subatómico, esta suposición ya no rige. La naturaleza se comporta de modo tal que las mediciones y las predicciones de un sistema atómico *nunca* podrán ser exactas y únicas en el sentido ordinario. Por más que intentemos mejorar nuestras técnicas, las mediciones tendrán siempre una precisión limitada y sólo podrán predecirse *probabilidades* para los valores de las magnitudes físicas en el dominio atómico. En otras palabras, es imposible predecir, digamos, *cuándo* un núcleo radiactivo dado se desintegrará, o exactamente *dónde* se encontrará un electrón dado en un instante dado en su trayectoria desde el cátodo hasta la pantalla de TV - sólo pueden especificarse probabilidades. Una física enteramente nueva tuvo que ser formulada a principios de 1920, apta para describir los sistemas atómicos y subatómicos: la llamada *física cuántica*.

El lector se preguntará a esta altura por qué estamos hablando de física cuántica, cuando ésta parece ser totalmente irrelevante para el estudio del sonido y la música. Sin embargo, la *psicofísica* opera en algunos aspectos de manera sorprendentemente similar a la física cuántica. En términos generales, la psicofísica, como la física en general, trata de formular predicciones sobre el comportamiento de un sistema específico sujeto a determinadas condiciones iniciales. El sistema bajo consideración es el

sistema sensorial (órgano receptor y partes del sistema nervioso relacionadas) de un individuo (o animal); las condiciones están determinadas por los *estímulos físicos que inciden*, en tanto que la respuesta está expresada por las *sensaciones psicológicas* evocadas en el cerebro y reportadas por el sujeto (o manifestadas en la conducta específica del animal). La *psicoacústica*, una rama de la psicofísica, es el estudio que liga estímulos acústicos con sensaciones auditivas. Como la física, la psicofísica requiere que las relaciones causales entre los estímulos físicos entrantes y la respuesta psicológica (o de conducta) sean establecidas a través de la experimentación y la medición. Como la física, la psicofísica debe hacer suposiciones simplificadas e idear *modelos* para poder establecer relaciones cuantitativas y aventurarse en el campo de la predicción. En los primeros tiempos de la psicofísica, las relaciones empíricas entre el estímulo físico y las sensaciones evocadas fueron condensadas en leyes psicofísicas, tratando al «hardware» interviniente como una «caja negra» (equivalente al punto material en física!). Hoy, los modelos psicofísicos toman en cuenta las funciones fisiológicas del órgano sensorial y las partes pertinentes del sistema nervioso.

A diferencia de la física clásica, pero similarmente a lo que ocurre en la física cuántica, no puede esperarse que las predicciones de la psicofísica sean exactas o únicas; solo pueden anticiparse valores que indican probabilidades. A diferencia de la física, pero análogamente a lo que sucede en la física cuántica, la mayor parte de las mediciones de la psicofísica perturbarán sustancialmente al sistema bajo observación (es decir al sujeto que reporta las sensaciones producidas por un determinado estímulo físico) y nada puede hacerse para eliminar completamente la perturbación mencionada. Como consecuencia de esto, el resultado de una medición psicofísica no refleja el estado del «sistema per se», sino más bien el estado más complejo del «sistema bajo observación». También como consecuencia, la psicofísica requiere que se realicen experimentos con muchos sistemas diferentes equivalentes - pero nunca idénticos - (sujetos), y una *interpretación estadística* de los resultados⁵.

Obviamente, hay ciertos límites para estas analogías. En física, el proceso o la «receta» de medición que define una magnitud física determinada, tal como la longitud, la masa o la velocidad de un objeto, puede ser formulada de una manera rigurosa, inequívoca. En la medida en que estemos tratando respuestas fisiológicas, como ser frecuencia de impulsos nerviosos, amplitud de la «piel de gallina» o incremento de la frecuencia cardíaca, las mediciones también pueden ser formuladas de manera rigurosa, inequívoca. Pero en psicoacústica, ¿cómo definir y medir las sensaciones *subjetivas* de altura, de sonoridad, o - para hacerlo aún más

⁵ Debemos enfatizar que éstas son solo analogías. La física cuántica como *tal* no asume un rol explícito en el sistema nervioso, cuya operación involucra procesos eminentemente clásicos.

difícil - la magnitud que representa la urgencia por resolver una melodía en la tónica? ¿Cómo podríamos efectuar mediciones en el terreno de la «audición interna», es decir la acción de provocar imágenes musicales por medio de la voluntad, sin la participación de estímulos externos? ¿Podría esto ser hecho simplemente por interrogación, o deberíamos recurrir a mediciones directas mediante la implantación de electrodos en las células cerebrales?

Muchas sensaciones pueden ser *clasificadas* en diferentes tipos, más o menos bien definidos (cuando son causadas por el mismo órgano, se habla de cualidades sensoriales). El hecho de que se hable de altura, sonoridad, timbre, consonancia, sin mayores desacuerdos sobre el significado de cada uno de estos conceptos, es un ejemplo. Profundizando más, encontramos que dos sensaciones que pertenecen al mismo tipo, al ser experimentadas una a continuación de la otra, en general, pueden ser *ordenadas* por el sujeto que las experimenta de acuerdo a si un atributo dado es percibido como «más grande» (o «más alto», «más fuerte», «más brillante», «más pronunciado», etc.), «igual» o «menor» en una que en la otra. Por ejemplo, cuando a un sujeto se le presentan dos tonos en forma sucesiva, éste puede emitir juicio sobre si la altura del segundo de ellos fue más alta, igual o más baja que la del primero. Otro ejemplo de ordenamiento es el siguiente: se hace escuchar una sucesión de tres tonos compuestos, de igual altura e igual sonoridad; un oyente puede ordenarlos por pares eligiendo los dos sonidos que tengan mayor similitud tímbrica, y aquellos dos que tengan mayor disimilitud. Una de las tareas fundamentales de la psicofísica es la determinación, para cada tipo de sensación, del valor mínimo detectable (valor del umbral) de la magnitud física responsable del estímulo, y del cambio mínimo detectable («diferencia apenas perceptible» o DAP)*.

La habilidad, poseída por todos, de clasificar y ordenar sensaciones subjetivas, les da a éstas un status similar al de las magnitudes físicas y justifica la introducción de la expresión *magnitud psicofísica*. Lo que *no* debemos esperar a priori es que una persona sin entrenamiento previo pueda decir que una sensación es el «doble», «la mitad» o cualquier otro factor *numérico* de una sensación «patrón», adoptada como unidad de medida. Hay situaciones, *sin embargo*, en las cuales es posible *aprender* a hacer estimaciones cuantitativas de magnitudes psicofísicas estadísticamente y, en algunos casos, el cerebro llega a ser muy diestro en esto. El sentido de la vista es un ejemplo. Con suficiente experiencia, la estimación del tamaño de un objeto puede llegar a ser muy precisa, siempre y cuando se disponga de suficiente información acerca del objeto dado. Juicios tales como «el doble de largo» o «la mitad de alto» se hacen sin dudar. Esto muestra claramente que la «unidad» y los correspondientes

* También «umbral diferencial» (N. del T.).

procesos de comparación han sido incorporados en el cerebro sólo *a través de la experiencia y el aprendizaje*, con múltiples contactos con las magnitudes físicas correspondientes. Lo mismo puede lograrse con otras sensaciones psicofísicas como la sonoridad: es necesario adquirir, a través de la experiencia, la habilidad de comparar y emitir juicios *cuantitativos*. El hecho de que músicos de todo el mundo usen la misma notación para la sonoridad es de por sí un ejemplo que sirve de evidencia.

Aquí es donde surgen tal vez las más importantes diferencias entre física y psicofísica: 1) La realización repetida de mediciones del mismo tipo puede *condicionar la respuesta* del sistema bajo observación: el cerebro tiene la habilidad de aprender, cambiando gradualmente las probabilidades de respuesta a un estímulo dado a medida que crece el número de veces que se hace el mismo experimento sobre el mismo sujeto. 2) La *motivación* del sujeto y las consecuencias mentales o físicas derivadas de ésta pueden interferir de una manera altamente imprevisible en las mediciones. Como resultado del primer punto, un estudio estadístico psicofísico realizado con un solo individuo expuesto a repetidas «mediciones» no conducirá de ningún modo a los mismos resultados que un estudio estadístico basado en un solo tipo de medición efectuada en muchos individuos. Esta diferencia no solo se debe a las diferencias subjetivas entre los individuos, sino también al condicionamiento que va apareciendo en el caso de repetir varias veces el mismo experimento. Los complejíssimos procesos del sistema nervioso determinan que las mediciones psicoacústicas sean particularmente difíciles tanto de realizar como de interpretar.

1.5 Psicofísica y neuropsicología

La psicofísica puede considerarse parte de una disciplina más amplia y abarcadora. La psicoacústica, por ejemplo, solo encara la cuestión de «por qué oímos lo que oímos» cuando estamos expuestos a un determinado estímulo acústico, pero no trata la cuestión de *qué es* lo que estamos oyendo! En otras palabras, deja de lado los procesos de más alto nivel que llevan a la cognición, la conducta y la respuesta emocional. La *neuropsicología* es la disciplina que estudia los *procesos* y las *funciones* del sistema nervioso que ligan la recepción de estímulos provenientes del mundo exterior y del organismo con las respuestas en el plano mental y en el de la conducta⁶. En otras palabras, la neuropsicología trata de dar un fundamento biológico, cuantitativo, sistemático de la experiencia mental y de la conducta. Como la física, la neuropsicología también trabaja con

⁶ Obsérvese que esto es muy diferente a la *psicología tradicional*, la cual es básicamente una ciencia descriptiva que se ocupa de la conducta per se, y de conceptos más intangibles como ser sentimientos, motivación, valores morales, conciencia y alma, y de los aspectos clínicos relacionados.

modelos. Estos son principalmente modelos de interrelaciones funcionales; por supuesto que las partes neuroanatómicas y los procesos fisiológicos intervinientes deben ser tomados en cuenta de manera realista. Es importante señalar que la definición de límites precisos entre psicofísica, psicología de la sensación, neuropsicología y psicología es una tarea difícil; ¡incluso la palabra «neuropsicología» aún no figura en muchos diccionarios!

El sistema principal estudiado por la neuropsicología es el *cerebro*⁷. Las más importantes funciones corticales de más alto nivel en el cerebro de un animal son, en breve, *la representación y la predicción de sucesos en el medio ambiente y la planificación de respuestas en el plano de la conducta*, con el objetivo último de acrecentar las posibilidades de supervivencia y perpetuación de la especie. Para lograr esto, el cerebro debe, a largo plazo, obtener la información sensorial necesaria para hacerse un «plano» de los alrededores y descubrir relaciones de causa y efecto que rijan la sucesión de acontecimientos en el tiempo. A corto plazo, tiene que hacer una evaluación del estado momentáneo del organismo y del entorno, identificar los rasgos o cambios más relevantes, hacer predicciones de corto plazo basándose en la experiencia (información adquirida por aprendizaje) y el instinto (información genética), así como planificar y ejecutar respuestas que sean supuestamente beneficiosas para el organismo. Todas estas tareas están guiadas y motivadas por una parte del cerebro, antigua filogenéticamente, llamada *sistema límbico*, el cual origina señales que componen el *estado afectivo* del organismo (placer, miedo, ira, expectativa, ansiedad, alivio, etc.).

El cerebro *humano* puede trabajar con su propio 'output' y planificar respuestas de conducta que sean independientes del estado momentáneo del entorno y del cuerpo, con objetivos que estén desconectados de los requerimientos momentáneos de supervivencia. Por otra parte puede evocar información a voluntad sin que haya existido una estimulación externa o somática, analizarla y realmacenar en la memoria versiones modificadas de la misma para su uso ulterior. Llamamos a esto «*el proceso del pensar humano*». Además, y a causa de esta habilidad de «control interno», el cerebro humano puede desobedecer deliberadamente los dictados del sistema límbico - ¡una dieta es un buen ejemplo! - y dedicarse a tareas de procesamiento de información para las cuales no fue originalmente creado ¡buenos ejemplos de esto son las matemáticas abstractas y el arte!

Todas las funciones cerebrales relacionadas con la percepción y la cognición están basadas en impulsos eléctricos generados, transmitidos y transferidos por neuronas. Hay más de diez mil millones de estas células

⁷ En general, y salvo algunas excepciones, no trataremos la anatomía del cerebro y la neurofisiología; existen muchos libros disponibles en la bibliografía médica (por ej. Brodal, 1969; Sommerhoff, 1974).

en el cerebro; una neurona puede estar conectada a otras cientos o incluso miles de neuronas, y cada operación cerebral, por simple que sea, incluirá millones de aquéllas. Es en la arquitectura y en las interconexiones sinápticas de este conglomerado de células nerviosas donde radica el misterio de la conciencia, la memoria, el pensamiento y los sentimientos. Cada operación cerebral, tal como el reconocimiento de un objeto que vemos, el imaginar un sonido musical o el placer que sentimos en un momento dado, está definida por una *distribución específica espacio-temporal de la actividad neuronal*. La «representación» del entorno, a la cual hicimos alusión más arriba, o la de cualquier imagen mental, no es otra cosa que la evocación de una distribución específica de impulsos nerviosos en determinadas áreas de la corteza; si bien esta distribución es increíblemente compleja, contiene rasgos o patrones que son absolutamente propios de aquello que está siendo representado o imaginado.⁸

El cerebro es el sistema interactivo más complejo del universo tal como se lo conoce hoy en día. Es por lo tanto bien comprensible que cualquier científico - ni qué hablar de los que no tienen esa formación - se enfrente con tremendos problemas al querer comprender en forma rigurosa por qué el funcionamiento de nuestro propio cerebro se nos presenta como algo tan «simple», y como «un Yo unitario» del cual tenemos absoluto control. (Estas características se llaman «la simplicidad natural de la función mental» y «la naturaleza unitaria de la experiencia consciente», respectivamente). Es bien comprensible también que se nos haga difícil aceptar el hecho de que, para alcanzar los objetivos científicos de la neuropsicología, no haya necesidad alguna de invocar por separado conceptos físicamente indefinibles, como «mente» y «alma».

1.6 ¿Qué es la música?

La discusión previa puede haber irritado a algunos lectores. Ellos dirán que la música es «estética pura», una manifestación de la sublime e innata comprensión humana de la belleza, más que el mero efecto de ondas sonoras actuando como estímulos sobre una compleja red de miles de millones de células nerviosas. Sin embargo, como se desprende de lo examinado precedentemente, *incluso los sentimientos estéticos están relacionados con procesamiento neural de información*. La mezcla tan

⁸ Usaremos el término «patrón» para designar la palabra inglesa «pattern», tan importante en biología y en neuropsicología. Reservamos el término «rasgo» para designar «feature», un componente aislado de un «pattern».

⁹ ¡Obsérvese bien que estos rasgos, aunque absolutamente específicos, no guardan ningún parecido «pictórico» con aquello que representan!

característica de patrones regulares, ordenados, alternados con sorpresa e incertidumbre, común a todo 'input' de información sensorial juzgada como estética, puede ser una manifestación del curioso pero fundamental deseo del hombre por ejercitar la inmensa red neuronal con operaciones de procesamiento no esenciales biológicamente, de complejidad variable. En realidad, la creatividad es la más humana de las actividades intelectuales. Mientras podría decirse que inteligencia y capacidad de comunicación son diferentes en el hombre que en los animales solo en una cuestión de grado, la creatividad y la apreciación artística son patrimonio exclusivo de los seres humanos.⁹

¿Entendemos realmente qué es la música? Cuando hablamos, estamos transmitiendo mensajes concretos. El pensamiento transmitido puede ser abstracto, pero el contenido sonoro suministra información. Cuando escuchamos los sonidos que nos rodean - gritos de animales, truenos, corrientes de agua - lo hacemos en respuesta a una tendencia innata a emplear todos nuestros sentidos para tomar conciencia de lo que nos rodea (véase sección previa). El escuchar otorga una ventaja en la carrera por la supervivencia: la interpretación de la información acústica ofrecida por el lenguaje y el medio ambiente es de una importancia biológica fundamental para nosotros. Pero ¿qué información transmite la música? En la mayoría de las culturas la música consiste en sucesiones rítmicas, estructuradas y organizadas de tonos elegidos de un repertorio muy limitado y discreto de alturas de alguna escala. No hay un equivalente directo con los sonidos del medio ambiente¹⁰, y su imitación no ha sido una fuerza impulsora en el desarrollo de culturas musicales. Pero si la música no transmite información que sea biológicamente relevante, ¿por qué nos impresiona? Bellos pasajes musicales pueden ponernos la «piel de gallina», pasajes crudos nos pueden llevar a sentir ira, y bebés que lloran se calman al escuchar los simples sonidos de una canción entonada por su madre ¿Por qué ocurre todo esto? ¿Por qué hay música?

La música bien puede ser un subproducto natural de la evolución del lenguaje humano. En esta evolución, que indudablemente fue un factor esencial para el desarrollo de la raza humana, se fue formando una red nerviosa capaz de realizar las ultracomplejas operaciones de procesamiento, análisis, almacenamiento y recuperación de información sonora necesarias para el reconocimiento fonético, la identificación de la voz y la comprensión del lenguaje. El lenguaje proveyó al ser humano un mecanismo que incrementó la capacidad de su memoria (y la de las operaciones asociadas de almacenamiento, recuperación y comunicación)

⁹ Obviamente, ¡no nos adherimos a la creencia de que plantas, vacas y gallinas, al ser expuestas a tal o cual tipo de música, aumentan su producción a causa de su apreciación artística!

¹⁰ El canto de un pájaro es música para nosotros, ¡pero para los pájaros es información concreta, como ser «este territorio es reclamado» o «este macho está buscando una hembra»!

en miles de millones de veces, al permitir la reducción de las complejíssimas escenas ambientales a breves representaciones simbólicas. En el curso de esta evolución, tuvo lugar una notable división de tareas entre los dos hemisferios cerebrales (Sec. 5.7). El hemisferio izquierdo (en alrededor del noventa y siete por ciento de las personas) ejecuta principalmente operaciones temporales de corto plazo tales como las que se requieren para la inteligibilidad verbal y otras *operaciones de secuenciación* de corto plazo, tal como el pensar. El hemisferio derecho coopera ejecutando operaciones de *integración espacial* y representación temporal de largo plazo. Ejemplos de estas operaciones *holísticas* son la imaginación pictórica y la *percepción musical*. En realidad, como veremos a lo largo de este libro, la percepción musical implica tanto el análisis de patrones espaciales de excitación a lo largo del órgano auditivo receptor, causados por sonidos musicales aislados o por superposiciones sonoras, como el análisis de patrones temporales de más largo plazo de las líneas melódicas. ¿Por qué respondemos emocionalmente a mensajes musicales complejos que no parecen contener ninguna información esencial para la supervivencia? El hecho de que la mayoría de nosotros lo hagamos - con frecuencia sin poseer ninguna preparación especial - indica que *el cerebro humano está instintivamente motivado a entretenerse con operaciones de procesamiento sonoro aun cuando dicha actividad no sea requerida por las circunstancias ambientales del momento*. Esta motivación bien puede ser el resultado de una tendencia innata a entrenarse desde muy corta edad en las altamente sofisticadas operaciones de análisis auditivo necesarias para la percepción del habla, de manera similar al juego del animal, que obedece a una tendencia congénita a mejorar la destreza de los movimientos que le serán indispensables para la caza y la defensa. Bebés que hubieran nacido sin esta propensión a escuchar atentamente, o nacidos de madres no inclinadas a vocalizar simples sonidos musicales, habrían tenido una decisiva desventaja en la carrera por la supervivencia en su entorno *humano*.

Dado que la percepción musical está basada en última instancia en el procesamiento de información acústica, la «causa» final de las sensaciones producidas por un mensaje musical dado podría estar relacionada con la mayor o menor dificultad de identificar el mensaje musical, el grado de éxito en las operaciones de predicción que realiza el cerebro para facilitar este proceso de identificación y el tipo de las asociaciones evocadas por comparación con información almacenada en previas experiencias. Si esto es así, sería obvio que tanto los mecanismos nerviosos innatos (operaciones de procesamiento primario) como el condicionamiento cultural (mensajes almacenados y operaciones de procesamiento aprendidas) deberán controlar nuestra respuesta conductiva y estética a la música.

2

Vibraciones sonoras, tonos puros y percepción de la altura

«Con más de un millón de partes esenciales en movimiento, el órgano receptor de la audición, o cóclea, es el aparato mecánico más complejo en el cuerpo humano.»

A. J. Hudspeth (referencia Hudspeth, 1985)

Escuchamos un sonido cuando el tímpano es puesto en un tipo de movimiento característico llamado *vibración*. Esta vibración es causada por pequeñas oscilaciones en la presión del aire en el canal auditivo por el cual está entrando una onda sonora. En este capítulo analizaremos primero los conceptos básicos referentes al movimiento vibratorio periódico en general para luego concentrarnos en los efectos producidos por las vibraciones del tímpano sobre nuestra sensación auditiva. No nos preocuparemos por ahora en establecer *cómo* el tímpano es puesto en movimiento. Con ese fin, imaginemos simplemente que nos colocamos auriculares y que escuchamos los tonos emitidos por éstos. En el rango de frecuencias más bajas los tímpanos seguirán estrechamente las vibraciones de los diafragmas de los auriculares. Este modo de introducir la materia quizás sea poco ortodoxo, pero nos permitirá lanzarnos directamente al estudio de algunos conceptos clave asociados con la vibración y la percepción del sonido, sin primero tener que gastar tiempo con las ondas sonoras y su generación. Desde el punto de vista práctico, este enfoque tiene un problema: los experimentos que vamos a presentar y analizar en este capítulo requieren necesariamente la generación electrónica de sonido, más que su producción natural con instrumentos musicales. Cuando sea posible, sin embargo, indicaremos cómo un experimento dado podría ser realizado con instrumentos ordinarios.

2.1 Movimiento y vibración

Movimiento significa *cambio de posición* de un cuerpo con respecto a otro cuerpo tomado como referencia. Si el cuerpo que se mueve es muy pequeño con respecto al cuerpo de referencia o al dominio espacial cubierto en su movimiento, de manera que su forma sea prácticamente irrelevante, el problema se reduce a la descripción del movimiento de un *punto* en el

espacio. Esta es la razón por la cual tal pequeño cuerpo es llamado un «punto material» o partícula. Por otra parte, si el cuerpo no es pequeño, pero sabemos de antemano que todos sus puntos están confinados a moverse a lo largo de líneas rectas paralelas («traslación rectilínea»), bastará también con especificar el movimiento de *un* solo punto dado del cuerpo. Este es un caso de movimiento «unidimensional», y la posición de un punto dado del cuerpo (y por ende de todo el cuerpo) queda completamente determinada por sólo *un* número: la distancia a un punto fijo de referencia.

En este libro solo consideraremos movimientos unidimensionales. Supongamos que nuestro punto material se mueve a lo largo de una línea vertical (Fig. 2.1). Indicaremos el punto de referencia sobre esa línea con la letra *O*. Cualquier punto fijo puede servir de referencia, aunque por comodidad, se adopta a veces alguno en particular (tal como la posición de equilibrio para un movimiento oscilatorio dado). Indicaremos la posición de un punto material *P* por la distancia *y* al punto de referencia *O* (Fig. 2.1). La cantidad *y* también se llama el *desplazamiento* de *P* con respecto a *O*, o *coordenada* de *P*. Debemos usar tanto números positivos como negativos, para poder distinguir entre los dos lados de *O*.

El punto material *P* está en movimiento con respecto a *O* cuando su posición y cambia con el tiempo. Indicaremos el tiempo usando la letra *t*. Lo medimos con un reloj; también se requiere que se especifique un instante de referencia en $t = 0$. El movimiento puede representarse matemáticamente de dos maneras: analíticamente, usando las llamadas relaciones funcionales, y geoméricamente mediante una representación gráfica. Nosotros solo usaremos el método geométrico. Para representar gráficamente un movimiento unidimensional, introducimos dos ejes perpendiculares entre sí: uno para representar el tiempo *t* y el otro para

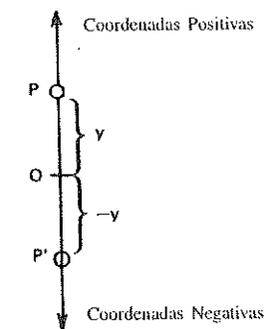


Figura 2.1 Posiciones instantáneas de un punto que se mueve sobre una línea recta. *y*: coordenada; *O*: punto fijo de referencia.

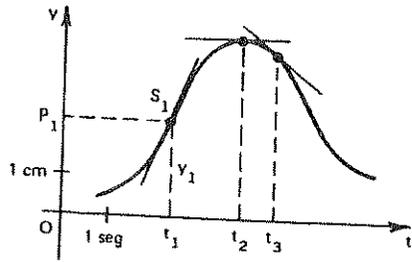


Figura 2.2 Representación gráfica del movimiento de un punto. ¡La curva gruesa *no* es la trayectoria!, determina las posiciones del punto en diferentes instantes t , a medida que se mueve en línea *recta* (coordenadas representadas por el eje y)

representar la coordenada y (Fig. 2.2). Para ambos tenemos que indicar claramente la *escala*, es decir, las unidades que usaremos para representar los intervalos (de tiempo y desplazamiento, respectivamente). Un movimiento puede ser representado marcando por medio de puntos la distancia y a la cual la partícula está momentáneamente ubicada en cada instante de tiempo t . Cada punto de la curva resultante, tal como el S_1 (Fig. 2.2), nos dice que al tiempo $t = t_1$, la partícula P está a una distancia y_1 del punto O , es decir, en la posición P_1 . ¡Hay que dejar bien claro que el punto material *no* se mueve a lo largo de la curva de puntos S ! Esta curva *no* es una trayectoria: es simplemente una «ayuda» que nos facilita la tarea de encontrar la posición y de la partícula en cualquier momento t .

El gráfico de la Fig. 2.2 también nos da información sobre la velocidad del punto material, es decir, sobre la rapidez del cambio de su posición. Esta está determinada por la *pendiente* de la curva en el gráfico: en t_1 la partícula se está moviendo con cierta velocidad hacia arriba; en t_3 lo está haciendo hacia abajo con menos rapidez. En t_2 está momentáneamente en reposo, invirtiendo su dirección de movimiento.

Hay ciertas clases de movimiento en las cuales el punto material sigue un patrón que se repite exactamente a medida que el tiempo transcurre. Este tipo de movimiento se llama *movimiento periódico* o *vibración*. Es el de mayor importancia para la física de la música. Para tener un verdadero movimiento periódico, un cuerpo *no* solo tiene que volver repetidamente a la misma posición, sino que lo debe hacer a intervalos de tiempo iguales y repetir exactamente el mismo tipo de movimiento en el ínterin. El intervalo de tiempo al cabo del cual la característica del movimiento se repite por primera vez se llama *período* del movimiento (Fig. 2.3). Lo indicamos por medio de la letra griega τ . Durante un período, el movimiento puede ser muy simple (Fig. 2.3[a]) o más complicado (Fig. 2.3 [b])¹. La característica o patrón básico del movimiento, que ocurre durante un período y que se repite continuamente, se llama *ciclo*.

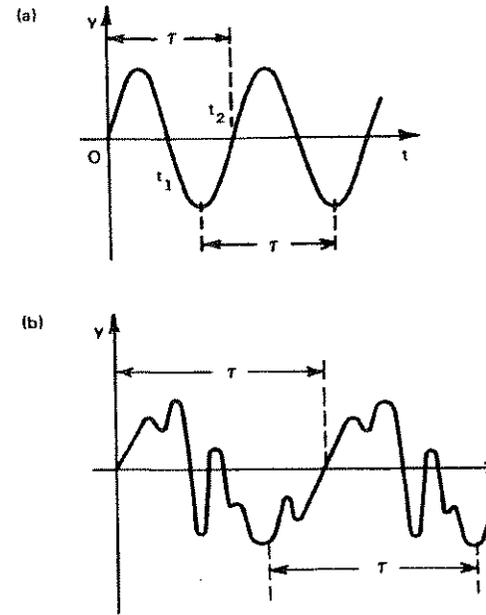


Figura 2.3 Representación gráfica de (a) un movimiento periódico simple; (b) un movimiento periódico complejo (τ : período)

Existen dispositivos mecánicos y electrónicos que pueden trazar automáticamente el gráfico de un movimiento periódico. En un *registrador gráfico*, la pluma reproduce en la dirección del eje y el movimiento periódico que se quiere graficar, al mismo tiempo que escribe sobre una tira de papel que se mueve perpendicularmente al eje y a una velocidad constante. Ya que esta velocidad se conoce, podemos asignar una escala de *tiempo* al eje representado por la tira de papel. La curva obtenida es la representación gráfica del movimiento. Este método, sin embargo, no es práctico para representar vibraciones acústicas. Estas tienen períodos tan cortos que sería imposible desplazar una pluma con suficiente velocidad como para reproducir este tipo de vibración. Un dispositivo electrónico, llamado *osciloscopio*, sirve para este propósito. En esencia, en este instrumento, un haz de electrones (partículas elementales con carga eléctrica negativa) muy estrecho incide en una pantalla de TV, produciendo un punto luminoso. Este haz puede ser desviado tanto en dirección vertical

¹ Un buen ejercicio para entender gráficos como los de las figuras 2.2 o 2.3 es hacer la mímica del movimiento representado moviendo la mano hacia arriba y abajo, representando la distancia (a lo largo del eje y) y su cambio a medida que el tiempo avanza uniformemente (eje t).

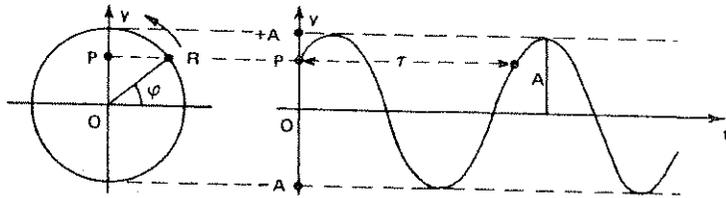


Figura 2.4 Movimiento armónico simple o *sinusoidal* (representado gráficamente a la derecha) obtenido como la proyección sobre un diámetro de un punto en movimiento circular uniforme (φ : fase; A: amplitud; τ : período)

como horizontal. El movimiento vertical está controlado por una señal proporcional a la vibración que queremos graficar (por ejemplo la vibración del diafragma de un micrófono). El movimiento horizontal es un barrido continuo hacia la derecha a velocidad constante, equivalente al movimiento de la tira de papel en el registrador gráfico, representando de este modo una escala de tiempo. De este modo, el punto luminoso sobre la pantalla describe el gráfico del movimiento durante un barrido. Si la imagen del punto luminoso se retiene lo suficiente, aparece como una curva continua sobre la pantalla. Dado que el tamaño de la pantalla es limitado, el movimiento horizontal se restablece instantáneamente al origen una vez que el haz alcanzó el borde derecho de la pantalla, y el barrido empieza de vuelta. Para representar un movimiento periódico, el barrido debe estar sincronizado con el período τ , o con uno de sus múltiplos.

2.2 Movimiento armónico simple

Preguntémonos ahora cuál es el tipo *más simple* de movimiento periódico. Hay muchos ejemplos en la naturaleza: la oscilación del péndulo o de un resorte, las oscilaciones de átomos, etc. Sus movimientos tienen algo importante en común: todos pueden ser representados como la proyección de un movimiento circular uniforme sobre un diámetro (Fig. 2.4)². Cuando el punto R da vueltas de manera uniforme (con período τ , es decir, una vuelta cada τ segundos) la proyección P se mueve hacia arriba y hacia abajo a lo largo del eje y con un movimiento denominado *movimiento armónico simple* (véase el gráfico que aparece en la parte derecha de la Fig. 2.4). A esto también se lo llama *movimiento sinusoidal* (porque y puede representarse analíticamente por una función trigonométrica llamada seno).

² Obsérvese con atención que la parte izquierda de la figura 2.4 es *auxiliar*; el único movimiento real es el movimiento periódico hacia arriba y hacia abajo de la partícula P a lo largo del eje.

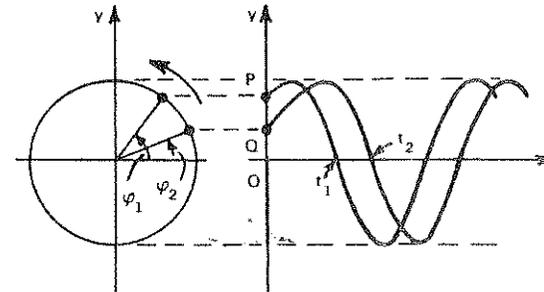


Figura 2.5 Representación gráfica del movimiento armónico de dos puntos de igual amplitud y frecuencia pero fases diferentes φ_1 y φ_2 .

Obsérvese que un movimiento armónico simple representa una vibración que es simétrica con respecto al punto O, el cual se denomina posición de equilibrio. El desplazamiento máximo A (sea para arriba o para abajo) se llama *amplitud*. τ es el *período* del movimiento armónico. Existe un parámetro más que describe un movimiento armónico simple y que es un poco más difícil de comprender. Observemos la Fig. 2.4: en el instante inicial $t = 0$ la partícula (proyección R del punto que rota) está ubicada en la posición P. Podemos ahora visualizar un segundo caso de movimiento armónico con el *mismo* período τ y la *misma* amplitud A, pero en el cual la partícula sale de una *posición diferente* Q (Fig. 2.5). El movimiento resultante será obviamente diferente, pero no en la forma o en el tipo, sino en su relativa «ubicación en el tiempo». En realidad, como se ve en la Fig. 2.5, las dos partículas pasarán por un mismo punto (por ej. por el punto de origen O) en instantes diferentes (t_1 , t_2). Inversamente, ambas partículas estarán en posiciones diferentes en un mismo instante (por ej. P y Q en $t = 0$). Si nuevamente imaginamos el movimiento de la segunda partícula Q como la proyección de un movimiento circular uniforme (Fig. 2.5), nos damos cuenta de que ambos casos pertenecen a diferentes *posiciones angulares*, φ_1 y φ_2 , de los puntos asociados que están rotando sobre la circunferencia. El ángulo φ se llama *fase* del movimiento armónico simple; la diferencia $\varphi_1 - \varphi_2$, que en este ejemplo se mantiene constante, se llama *diferencia de fase* entre dos movimientos armónicos.³

En resumen, una vibración «pura», o armónica, queda determinada por el valor de tres parámetros: el *período* τ , la *amplitud* A y la *fase* φ . Los tres, pero especialmente los dos primeros, desempeñan un rol fundamental en la percepción de sonidos musicales.

³ Use ambas manos para simular los movimientos hacia arriba y hacia abajo representados en la parte derecha de la figura 2.5. Hágalo para distintas diferencias de fase, por ejemplo, 180° (fases opuestas), 0° (concordancia de fase), 90° , etc.

15

Los movimientos armónicos simples ocurren prácticamente en todas partes del universo: las vibraciones de las partículas constituyentes de los átomos, las de los átomos en un cristal, las de los cuerpos elásticos, etc., todas pueden ser descritas en términos de movimientos armónicos simples. Pero existe una razón mucho más importante para considerar al movimiento armónico simple como el más fundamental de todos los movimientos periódicos: se puede demostrar matemáticamente que cualquier tipo de movimiento periódico, por complicado que sea, puede ser descrito como una suma de vibraciones armónicas simples. Trataremos en detalle esta propiedad fundamental más adelante (Cap. 4), por ser de capital importancia para la música.

2.3 Vibraciones acústicas y sensaciones de tonos puros

Cuando el tímpano ejecuta un movimiento periódico, sus vibraciones mecánicas se convierten en el oído interno en impulsos nerviosos eléctricos que se transmiten al cerebro, donde son interpretados como sonido, siempre y cuando el período y la amplitud de las vibraciones ocurran dentro de ciertos límites. En general, el oído es un aparato extremadamente sensible: vibraciones del tímpano tan pequeñas como 10^{-7} cm pueden ser detectadas, como también vibraciones de períodos tan cortos como 7×10^{-5} seg.⁴

Introducimos ahora una cantidad mucho más usada que el período τ , que se llama *frecuencia*:

$$f = 1/\tau \quad (2.1)$$

Físicamente, f representa el número de repeticiones del patrón vibratorio, o ciclos, en el intervalo unidad de tiempo. La razón por la cual se prefiere f a τ es que nuestra sensación de «agudeza del tono» o altura se incrementa a medida que crece la frecuencia. Si τ se expresa en segundos, f se expresa en *ciclos por segundo*. Esta unidad se llama hertz (Hz), en honor al famoso físico alemán Heinrich Hertz. Las vibraciones comprendidas entre los 20 y los 15.000 Hz son percibidas como sonido por una persona normal. Los límites inferior y superior - particularmente éste último - dependerán de la sonoridad del tono y podrán variar considerablemente de una persona a otra, como también con su edad.

Cuando un sonido produce un movimiento armónico simple del tímpano con características constantes (frecuencia, amplitud, fase),

⁴En este libro usaremos la notación exponencial: $10^n = 100\dots(n \text{ ceros})\dots00$. 10^{-n} es simplemente $1/10^n$, es decir, una fracción decimal dada por una unidad en la n -ésima posición decimal.

percibimos aquello que se llama *tono simple*. Un tono simple suena pobre y la música no está hecha de tonos simples aislados. Sin embargo, como dijimos en la introducción de este capítulo, para una mejor comprensión de los sonidos compuestos es aconsejable estudiar primero los tonos simples. Estos tonos tienen que ser generados con osciladores electrónicos; no hay ningún instrumento musical que los produzca (incluso no hay garantía de que tonos simples generados electrónicamente sigan siendo puros en el momento en que llegan al oído). Sea como fuere, dado que la flauta es el instrumento cuyo sonido se aproxima más al de un tono simple, sinusoidal, especialmente en el registro agudo, varios (pero no todos) de los experimentos mencionados en este capítulo pueden realizarse usando una o, en algunos casos, dos flautas - ¡aunque tocadas por expertos!

Cuando escuchamos un tono simple cuya frecuencia y amplitud pueden ser variadas a voluntad, constatamos una correspondencia entre *altura* y *frecuencia* y entre *sonoridad* y *amplitud*. Hoy día se tiene una idea bastante clara de cómo funciona el mecanismo primario de detección de la frecuencia y la amplitud de sonidos simples. En este capítulo nos dedicaremos solo a la altura.

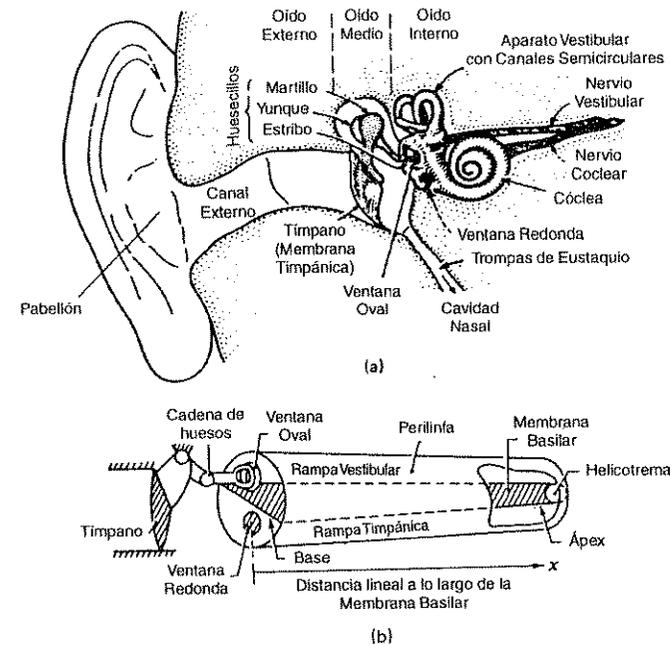


Figura 2.6 (a) Vista esquemática del oído (Flanagan, 1972; Fig. 4) (no en escala); (b) la cóclea 'desenrollada' (versión muy simplificada).

Las oscilaciones armónicas simples del tímpano son transmitidas por medio de una cadena de tres huesecillos en el oído medio, llamados «martillo», «yunque» y «estribo» (o, en términos más eruditos, malleus, incus y stapes, lo cual quiere decir exactamente lo mismo en latín), a la entrada («ventana oval») del oído interno propiamente dicho (Fig. 2.6). La *cóclea*, de tamaño no mayor que una bolita, es un túnel a través del hueso temporal, en forma de espiral, como el caparazón de un caracol. Esta cavidad, mostrada en la Fig. 2.6[b] en una versión muy simplificada y 'desenrollada', está dividida en dos canales, la rampa vestibular y la rampa timpánica, los cuales están llenos de un fluido incompresible, la *perilinf*a (un filtrado directo del fluido cerebroespinal). Ambos canales funcionan como un sistema hidrodinámico, porque están limitados en el extremo final, o ápex, por un pequeño orificio llamado *helicotrema*; la sección inferior está acordonada por una membrana elástica en la «ventana redonda» (Fig. 2.6 [b]) La partición que separa ambas rampas es en sí misma un conducto altamente estructurado de corte transversal triangular, también llamada rampa media, Fig. 2.7[a], la cual contiene otro fluido, la *endolinfa*. Sus límites son la *membrana basilar*, que sostiene el órgano sensorial propiamente dicho (órgano de Corti), la *membrana de Reissner**, que sirve para separar la endolinfa de la perilinfa, y la rígida pared lateral de la cóclea.

La elasticidad de la membrana basilar determina las propiedades hidromecánicas básicas de la cóclea. En la persona adulta, la membrana tiene aproximadamente unos 34 milímetros de largo, desde la base

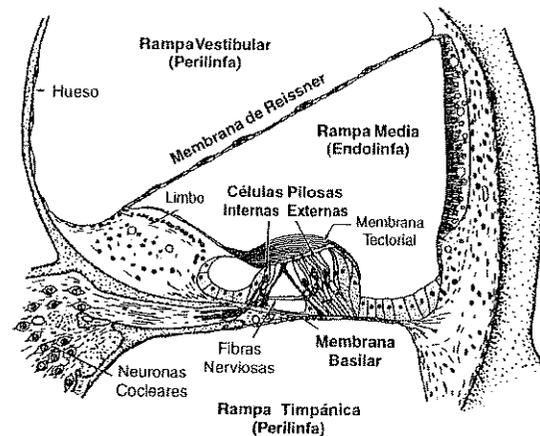


Figura 2.7 (a) Corte transversal del órgano de Corti (según Davis, 1962).

* También llamada membrana vestibular (N. del T.).

(extremo inicial) hasta el ápex; a causa del cambio gradual de su ancho y espesor su rigidez decrece desde la base hasta el ápex unas 10.000 veces, lo cual da a la membrana basilar su función fundamental de analizadora de frecuencias. Las vibraciones transmitidas por los tres huesecillos a la ventana oval, se convierten en oscilaciones de presión en la perilinfa, dentro de la rampa vestibular. Las diferencias de presión resultantes a través de la partición coclear entre las dos rampas flexionan la membrana basilar para arriba y para abajo impartándole un movimiento similar al de una bandera que flamea. A medida que la onda se propaga hacia el ápex, su amplitud alcanza un máximo en un punto determinado de su camino, para luego decaer muy rápidamente hacia el ápex. Alrededor de 16.000 unidades receptoras, llamadas *células pilosas* (o *ciliadas*), ordenadas en filas, una interna y tres externas, a lo largo de la membrana basilar, recogen los movimientos de esta última e imparten señales a las células nerviosas, o *neuronas*, que están en contacto con ellas. El nombre célula pilosa deriva del hecho que en su extremo hay un haz de 20 a 300 ramificaciones llamadas *estereocilias* (Fig. 2.7 [b]) que se proyectan en el fluido endolinfático, y cuyos desvíos disparan una serie de procesos electroquímicos en la célula pilosa y sus alrededores, que culminan con la generación de señales eléctricas en el nervio acústico. La *membrana tectoria* es un tejido gelatinoso suspendido en la endolinfa, sobre el órgano de Corti (Fig. 2.7 [a]), dentro de la cual están insertadas las cilias de las

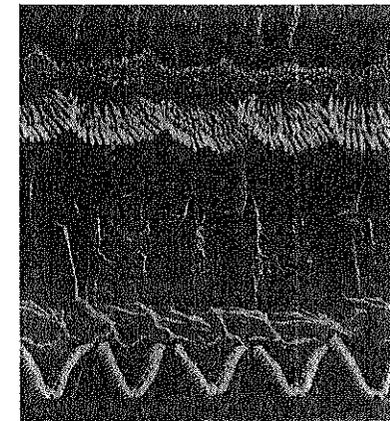


Figura 2.7 (b) Microfotografía electrónica (Bredberg et al., 1970) de la hilera interna (al tope) y externa (parte inferior - sólo se muestran tres hileras) de células pilosas sobre la membrana basilar de un cobayo. (Estos animales, como también las chinchillas y los gatos, tienen sistemas acústicos periféricos muy similares al de los seres humanos y, por esta razón, son los animales más usados en investigaciones sobre la audición).

células pilosas externas; esta membrana asume un rol clave estimulando y recibiendo el movimiento de las cilias. Volveremos a la acción coclear con mucho más detalle en las secciones 2.8 y 3.6.

El hecho notable es que para un *tono simple* de determinada frecuencia, las oscilaciones máximas de la membrana basilar solo ocurrirán en una región limitada de ésta, *cuya posición dependerá de la frecuencia del tono*. En otras palabras, para cada frecuencia hay una zona de máxima estimulación, una «zona de resonancia», sobre la membrana basilar. Cuanto más baja sea la frecuencia del tono, tanto más cerca del ápex (Fig. 2.6 [b]) estará la región de células pilosas activadas (donde la membrana es más flexible). Cuanto más alta sea la frecuencia, tanto más cerca de la entrada (de la «ventana oval») estará localizada esta región (donde la membrana es más rígida). *La posición espacial x a lo largo de la membrana basilar (Fig. 2.6 [b]) de las células pilosas activadas y neuronas asociadas, determina la sensación primaria de altura* (también llamada «altura espectral»). Un cambio en la frecuencia del tono simple que sirve de estímulo, causa un cambio en la posición de la región de células pilosas activadas. Decimos que la información primaria sobre la frecuencia de un tono es «codificada» por el órgano sensorial de la membrana basilar en forma de la *localización espacial* de las neuronas activadas. Dependiendo de cuál grupo de fibras nerviosas se active, la altura se percibirá más aguda o más grave.

La figura 2.8 muestra cómo la posición x (medida desde la base, Fig. 2.6 [b]) de la región de máxima excitación depende de la frecuencia de un tono simple, sinusoidal, en un adulto normal (Von Békésy, 1960). Varias conclusiones pueden extraerse. Ante todo, obsérvese que el rango de frecuencias más importante musicalmente (de los 20 a los 4000 Hz) abarca más o menos las dos terceras partes de la extensión de la membrana basilar (12 a 35 mm desde la base). La gran porción restante de la escala de frecuencias (desde los 4000 hasta los 16.000 Hz - la figura 2.8 no muestra más allá de los 5000 Hz -) está comprimida en el tercio restante. En segundo lugar, cada vez que se dobla la frecuencia de un tono, es decir que la altura salta una octava, la región de resonancia correspondiente se desplaza una cantidad más o menos constante, entre 3,5 y 4 mm, no importa que este salto de frecuencia sea de 220 a 440 Hz, de 1760 a 3520 Hz, o de 5000 a 10.000 Hz. En general, cuando la frecuencia f es multiplicada por un factor determinado, la posición x de la zona de resonancia no es multiplicada, sino simplemente trasladada una cierta distancia. En otras palabras, son los *cocientes* entre frecuencias, y no las diferencias entre éstas, los que determinan el desplazamiento de la zona de resonancia a lo largo de la membrana basilar. Una relación de este tipo se llama «logarítmica» (Sec. 3.4).

Los resultados arriba mencionados derivan de mediciones fisiológicas realizadas en animales muertos (pero bien conservados). Hoy en día tales

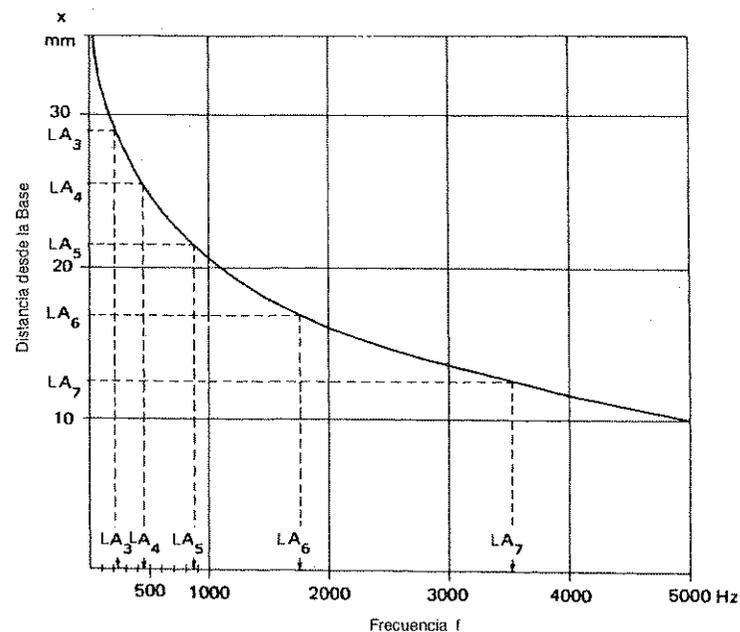


Figura 2.8 Posición del máximo de resonancia sobre la membrana basilar (según von Békésy, 1960) para un tono puro de frecuencia f (escalas lineales).

mediciones pueden ser realizadas en cócleas de animales vivos mediante el efecto Mössbauer. Una pequeña masa de sustancia radiactiva (cobalto 57) es «implantada» sobre la membrana basilar (lo cual, sin embargo, puede afectar en alguna medida la medición). Los pequeños desplazamientos de la membrana basilar pueden ser detectados indirectamente midiendo el cambio de frecuencia (efecto Doppler) de los rayos gamma emitidos por la sustancia (ver por ej. Rhode y Robles, 1974). Más detalles sobre esta cuestión en la sección 3.6.

Consideremos ahora la magnitud psicofísica altura, asociada a un tono simple de frecuencia f . En la Sec. 1.4 mencionamos que una magnitud psicofísica no puede medirse en forma cuantitativa del mismo modo en que se lo hace con magnitudes físicas, tales como la frecuencia. Solo puede establecerse cierto *orden* entre dos sensaciones del mismo tipo presentadas en sucesión inmediata. Para algunas sensaciones es posible realizar estimaciones cuantitativas únicamente después que, de algún modo, el cerebro se haya entrenado en la ejecución de las operaciones necesarias (por ej. un niño que aprende a estimar el tamaño de los objetos que ve), y los resultados tendrían que ser interpretados en forma estadística.

Analícemos la habilidad de un individuo para establecer el orden relativo, en términos de altura, de dos tonos (de la misma intensidad) que le son presentados uno *después* del otro. Hay un límite natural: cuando la diferencia en frecuencia de los dos tonos es demasiado pequeña, debajo de un cierto valor, se juzgará que la altura de ambos tonos es la misma. Esto vale para juicios sobre el orden relativo de cualquier magnitud psicofísica: siempre que la variación de un estímulo físico caiga dentro de cierta «diferencia liminal» o *diferencia apenas perceptible* (DAP), la sensación asociada será juzgada como siendo «la misma» que la precedente; tan pronto como la variación exceda la DAP, será detectado un cambio entre las sensaciones. Obsérvese que la DAP se refiere a una magnitud *física*, que se mide en el sentido ordinario y que se expresa con un número.

El grado de sensibilidad del mecanismo de percepción de altura primaria para los cambios de frecuencia, o capacidad de *resolución de frecuencia*, depende de la frecuencia, la intensidad y la duración del tono en cuestión - y de cuán repentino es el cambio de frecuencia. Varía mucho de persona a persona; es función de la preparación musical y, desafortunadamente, depende considerablemente del método de medición empleado. La Fig. 2.9 muestra la DAP de frecuencia promedio para tonos simples de

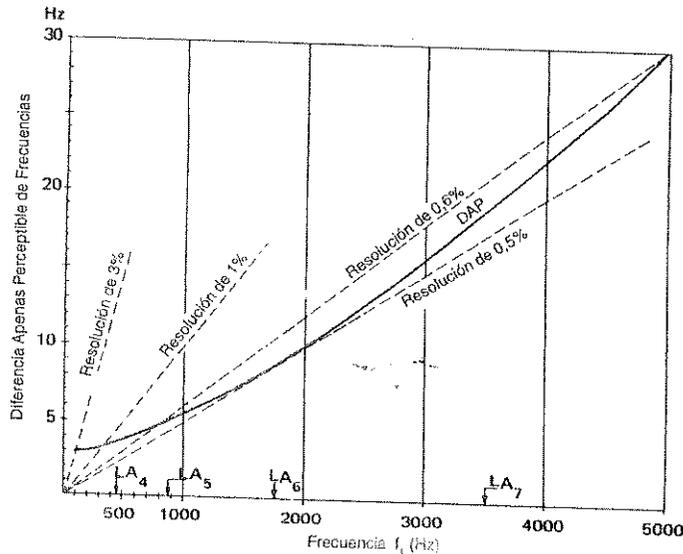


Figura 2.9 Diferencia apenas perceptible (DAP) en frecuencia para un tono puro de frecuencia f_1 (escalas lineales) determinada usando una señal lentamente modulada en frecuencia (según Zwicker, Flottorp y Stevens, 1957).

intensidad constante (80 decibeles, Sec. 3.4); la frecuencia de estos tonos fue lenta y continuamente modulada (Zwicker, Flottorp y Stevens, 1957). Este gráfico muestra, por ejemplo, que para un tono de 2000 Hz, un cambio de 10 Hz - es decir de solo un 0,5% - puede ser ya detectado. ¡Esto es una fracción muy pequeña de semitono! Los cambios *repentinos* en la frecuencia son detectados con una DAP considerablemente menor: hasta 30 veces menos que los valores mostrados en la figura 2.9 (Rakowski, 1971). La resolución de frecuencia empeora hacia las frecuencias bajas (por ejemplo, tres por ciento a los 100 Hz en la Fig. 2.9). También decae a medida que se acorte la duración del tono, por debajo de un décimo de segundo. Por otra parte, la resolución de frecuencia es relativamente independiente de la amplitud (sonoridad). Para una discusión actualizada sobre la DAP véase Zwicker y Fastl (1990).

Desde la aparición de la psicofísica, los psicólogos se sintieron tentados a considerar el mínimo cambio perceptible en la sensación, causado por una DAP del estímulo, como la unidad natural con la cual medir la magnitud psicofísica correspondiente. Por ejemplo, el mínimo cambio perceptible en la altura ha sido usado para construir una «escala subjetiva de alturas» (Stevens, Volkman y Newmann, 1937). Sin embargo, como veremos más adelante, en parte porque la octava ha asumido decisivamente ese rol de intervalo «natural» de la altura y en parte porque todas las escalas musicales se desarrollaron con absoluta independencia de los intentos de establecer una escala subjetiva de alturas, esta última no ha encontrado una aplicación directa en música (véase, sin embargo, Sec. 5.4).

En resumen, de acuerdo con la descripción dada en esta sección, la función primaria del oído interno (cóclea) es convertir un patrón de vibración en el tiempo (el del tímpano) en un patrón de vibración en el espacio (a lo largo de la membrana basilar), y transformar a su vez a este último en un patrón espacial de actividad nerviosa. La descripción teórica de este mecanismo se conoce como *teoría auditiva de localización* [«place theory of hearing»], o teoría de altura espectral. Nos referiremos frecuentemente a ella, y veremos que es una teoría buena, aunque incompleta.

2.4 Superposición de tonos simples: batidos de primer orden y la banda crítica

Antes dijimos que tonos simples aislados suenan muy aburridos. Las cosas cobran un poco más de vida cuando superponemos dos tonos puros, haciéndolos sonar juntos. En esta sección analizaremos las características fundamentales que presenta la superposición de dos tonos puros. Esto nos llevará a conceptos fundamentales de la física de la música y de la psicoacústica.

Hay dos clases de efectos que resultan de una superposición, dependiendo de dónde son procesados en el sistema auditivo del oyente. Si el procesamiento es *mecánico*, ocurriendo en el fluido coclear y a lo largo de la membrana basilar, los llamaremos «efectos de superposición de primer orden», principalmente porque son claramente distinguibles y fácilmente accesibles a la experimentación psicoacústica. Los efectos de superposición de «segundo orden» son el resultado del procesamiento *neural*; son más difíciles de detectar, describir y medir en forma inequívoca. En esta sección trataremos efectos de primer orden solamente.

Analicemos primero el significado físico de «superposición sonora». El tímpano se mueve hacia adentro y afuera comandado por las variaciones de presión del aire en el canal auditivo. Si se le ordena oscilar con un movimiento armónico puro de frecuencia y amplitud determinadas, lo que escuchamos es un tono simple de cierta altura y sonoridad. Si ahora *dos* tonos simples de diferentes características suenan juntos (por ejemplo escuchando dos fuentes sonoras independientes al mismo tiempo) el tímpano reacciona como si estuviera ejecutando dos órdenes independientes al mismo tiempo: una para cada tono puro. El movimiento resultante es la suma de los movimientos individuales que ocurrirían si cada sonido estuviera solo, sin el otro. No solo el tímpano se comporta de esta manera, sino también el medio y todos los componentes vibrantes (esto, sin embargo, deja de ser cierto cuando las amplitudes son muy grandes). Este efecto se llama *superposición lineal* de dos vibraciones, expresión técnica que significa «coexistencia pacífica»: una vibración no perturba los asuntos de la otra, y la superposición resultante simplemente sigue los dictados de cada componente en forma simultánea. En una superposición *no lineal* el dictado de uno de los componentes dependerá de lo que el otro componente tiene que decir y viceversa.

Empecemos nuestra discusión con el análisis de la superposición de dos movimientos armónicos simples de *igual frecuencia e igual fase* (diferencia de fase cero, Sec. 2.2). Puede demostrarse gráficamente (Fig. 2.10) y también de manera analítica que en este caso obtenemos un movimiento armónico simple de igual frecuencia y de igual fase, pero

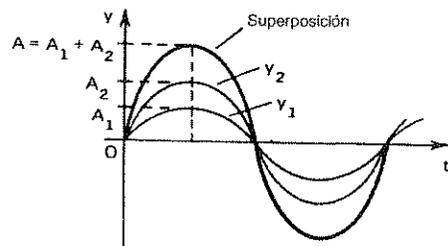


Figura 2.10 Superposición de dos vibraciones sinusoidales de igual fase e igual frecuencia.

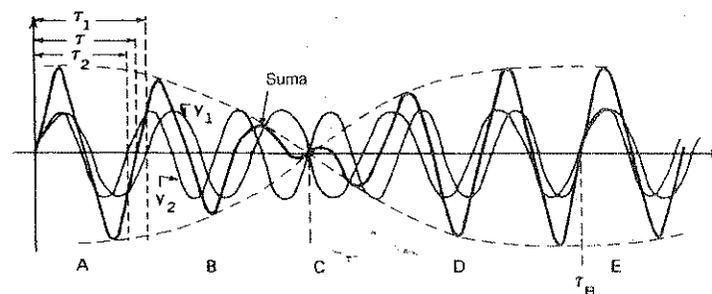


Figura 2.11 Superposición de dos oscilaciones sinusoidales de períodos ligeramente diferentes, τ_1 y τ_2 , correspondientes a las frecuencias f_1 y f_2 .

con una amplitud que es la suma de las amplitudes de las dos vibraciones componentes. Si las dos oscilaciones componentes, con la misma frecuencia, tienen *diferente fase*, su superposición seguirá siendo un movimiento armónico simple con la misma frecuencia, pero su amplitud no será más la suma de las amplitudes componentes. En el caso particular de que las amplitudes de las vibraciones componentes sean iguales y su diferencia de fase ϕ sea 180° , ambas oscilaciones se anularán mutuamente y no se escuchará sonido alguno. Esto se llama *interferencia destructiva* y tiene un rol en la acústica de salas. En suma, cuando dos tonos puros de igual frecuencia alcanzan el tímpano, percibimos solo *un* tono de determinada frecuencia (correspondiente a la frecuencia de los sonidos componentes) y sonoridad (controlada por las amplitudes de los tonos superpuestos y su diferencia de fase).

Consideremos ahora la superposición de dos tonos simples de la *misma amplitud* pero de *frecuencias ligeramente diferentes*, f_1 y $f_2 = f_1 + \Delta f$. La diferencia de frecuencia Δf tiene un valor pequeño. Supongamos que es positivo, de manera que el tono correspondiente a f_2 será un poco más agudo que el correspondiente a f_1 . El patrón vibratorio del tímpano estará dado por la suma de los patrones de cada tono componente (Fig. 2.11). El resultado de la superposición (curva gruesa) es una oscilación de período y frecuencia *intermedia* entre f_1 y f_2 , y de una amplitud lentamente modulada. Obsérvese en esta figura el cambio gradual en la diferencia de fase entre los dos tonos componentes y_1 y y_2 : empiezan en fase (diferencia de fase 0° , como en la Fig. 2.10) en el instante $t = 0$, luego y_2 empieza a llevar la delantera en fase hasta que ambos están completamente fuera de fase (diferencia de fase 180°) en el instante correspondiente a C. La diferencia de fase sigue creciendo hasta que alcanza $360^\circ = 0^\circ$ en el instante t_B . Este lento y continuo cambio de fase es el que da lugar a la amplitud variable de la oscilación resultante: las curvas punteadas en la Fig. 2.11 representan la *envolvente de amplitud* de la vibración resultante (véase también Fig. 2.17 (A)).

¿Cuál es la sensación sonora resultante en este caso? Antes que nada, téngase claro que el tímpano seguirá la oscilación indicada por la curva gruesa en la Fig. 2.11. El tímpano «no sabe» ni «le importa» que este patrón es en realidad el resultado de la suma de otros dos; tiene solo un patrón vibratorio de amplitud cambiante. Algo notable ocurre en el fluido coclear: este patrón vibratorio en la ventana oval, relativamente complicado pero único, origina dos zonas de resonancia de la membrana basilar. Si la diferencia de frecuencia Δf de los dos tonos componentes es lo suficientemente grande, existirá suficiente separación entre las dos zonas de resonancia correspondientes; cada una oscilará con la frecuencia que corresponda al tono correspondiente (curvas finas en Fig. 2.11) y lo que escuchamos serán dos tonos separados, de sonoridad constante, con alturas que corresponderán a cada uno de los tonos originales. Esta propiedad de la cóclea de separar un patrón vibratorio complejo, causado por una superposición de tonos, en los tonos simples originales, se llama discriminación de frecuencia. Es un proceso mecánico, controlado por las propiedades hidrodinámicas y elásticas de los componentes del oído interno. Por otra parte, si la diferencia de frecuencia Δf es menor que cierta cantidad, las zonas de resonancia se superponen y lo que escuchamos será un solo tono de altura intermedia y sonoridad modulada o «batiente». En este caso, la zona de resonancia de la membrana basilar que resulta de la superposición, sigue un patrón de vibración esencialmente idéntico al del tímpano (curva gruesa en Fig. 2.11). La modulación en amplitud del patrón vibratorio (envolvente mostrada con líneas punteadas en la Fig. 2.11) es la que causa la modulación en la sonoridad percibida. Llamamos a este fenómeno «batidos de primer orden». Estos son los batidos ordinarios, bien conocidos por todo músico. La frecuencia del patrón vibratorio resultante de la superposición de dos tonos de frecuencias muy cercanas f_1 y f_2 es igual al valor promedio:

$$f = \frac{f_1 + f_2}{2} = f_1 + \frac{\Delta f}{2} \quad (2.2)$$

El intervalo de tiempo τ_B (Fig. 2.11), al cabo del cual la amplitud resultante alcanza el valor inicial, se llama período de batido. La frecuencia de batido $f_B = 1/\tau_B$ (número de cambios de amplitud por segundo) está dada por la diferencia

$$f_B = f_2 - f_1 = \Delta f \quad (2.3)$$

No importa que f_2 sea mayor que f_1 o viceversa. Los batidos serán percibidos en cualquiera de los dos casos y la frecuencia estará siempre dada por la diferencia de frecuencias de los tonos componentes (la relación 2.3 debe ser tomada en valor absoluto, es decir, siempre positiva). Cuanto más próximas sean las frecuencias f_1 y f_2 , tanto más «lentos» serán los

batidos. Si f_2 llega a ser igual a f_1 , los batidos desaparecen por completo: los dos tonos componentes suenan al *unísono*.

Resumamos las sensaciones evocadas por la superposición de dos tonos simples de igual amplitud y de frecuencias f_1 y $f_2 = f_1 + \Delta f$, respectivamente. Con ese fin, supongamos que mantenemos f_1 fija y que incrementamos lentamente f_2 desde f_1 (desde el unísono, $\Delta f = 0$) a valores más altos. (Nada cambiaría, cualitativamente en lo que vamos a describir, si en lugar de aumentar los valores de f_2 , los bajáramos.) Cuando están al unísono, percibimos un solo tono cuya altura corresponde a f_1 y cuya sonoridad dependerá de la particular diferencia de fase entre los dos tonos simples. Si incrementamos levemente la frecuencia f_2 , continuaremos percibiendo un solo tono, pero de una altura levemente más alta que corresponde a la frecuencia promedio $f = f_1 + \Delta f/2$ (2.2)⁵. La sonoridad de este tono oscilará, «batirá», con una frecuencia Δf (2.3). La frecuencia de batido se incrementa a medida que f_2 se aleja de f_1 (incremento de Δf). Mientras que Δf sea menor que 10 Hz aproximadamente, los batidos son percibidos claramente. Cuando la frecuencia de batido excede, digamos, los 15 Hz, la sensación de batidos desaparece, dando lugar a una característica *rugosidad* o cualidad desagradable de la sensación resultante. Cuando Δf supera el así llamado *límite de discriminación de frecuencia* Δf_D (que no debe ser confundido con el límite de resolución de frecuencia, o DAP de frecuencia, de la Fig. 2.9), distinguimos de repente dos tonos separados, cuyas alturas corresponden a f_1 y f_2 . En ese momento las dos zonas de resonancia de la membrana basilar se han separado lo suficiente como para dar dos señales distintas de altura. Sin embargo, en ese límite, la sensación de rugosidad todavía persiste, especialmente en el rango de alturas más graves. Solo después de sobrepasar una diferencia de frecuencias aun mayor Δf_{CB} , llamada *banda crítica*, la sensación de rugosidad desaparece, y los dos tonos simples suenan puros y placenteros. Esta transición de «rugosidad» a «suavidad» es en realidad más gradual; la banda crítica, tal como la hemos definido aquí, solo representa la diferencia de frecuencias *aproximada* en la que esta transición tiene lugar.

Todos estos resultados son fáciles de verificar usando dos «generadores sinusoidales» electrónicos de frecuencia variable, combinando sus salidas y escuchando en forma monoaural el resultado a través de auriculares. También pueden verificarse, por lo menos cualitativamente, con dos flautas que suenan simultáneamente en el registro agudo, ejecutadas por expertos. Mientras un flautista mantiene una nota fija (sosteniendo la altura bien

⁵ ¿Cómo verificamos que la sensación de altura del sonido resultante corresponde realmente al sonido de frecuencia $f_1 + \Delta f/2$? Esto se hace con experimentos de *apareamiento de altura tonal* ('pitch matching'): se le presenta al sujeto alternativamente un tono de referencia cuya altura es controlable y se le pide que modifique la frecuencia de este último hasta que sienta que su altura «es igual» a la del sonido original.

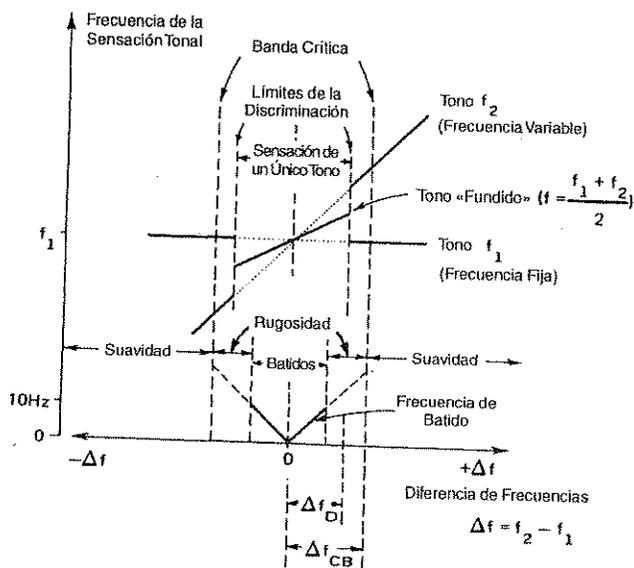


Figura 2.12 Representación esquemática de la frecuencia (líneas gruesas) que corresponde a las sensaciones tonales producidas por la superposición de dos tonos puros de frecuencias cercanas f_1 y $f_2 = f_1 + \Delta f$.

fija), el otro toca la misma nota *escrita* pero desafinándola (alejando o acercando gradualmente la embocadura). Cuestiones tales como batidos, rugosidad y discriminación tonal pueden ser exploradas con relativa facilidad. La figura 2.12 describe en forma comprensiva (no en escala) los resultados arriba mencionados. Las líneas gruesas representan las frecuencias de los tonos (o la de los batidos) que *realmente* se perciben. El tono f_1 es el de la frecuencia fija; f_2 corresponde al tono cuya frecuencia se modifica gradualmente (aumentándola o decreciéndola). El tono «fusionado» corresponde a la sensación de tono único (de frecuencia intermedia) percibido cuando f_2 está dentro del límite de discriminación de frecuencias de f_1 . Obsérvese la extensión de la banda crítica a uno u otro lado del unísono ($\Delta f = 0$). Debemos enfatizar nuevamente que esta transición de la rugosidad a la suavidad no es del todo repentina, como uno podría interpretar a partir de la Fig. 2.12, sino más bien gradual. Un tratamiento más detallado y riguroso está dado en el libro de Zwicker y Fastl (1990).

Los límites para la discriminación de frecuencias y para la banda crítica dependen fuertemente de la frecuencia promedio $(f_1 + f_2) / 2$ de los dos tonos (llamada *frecuencia central* del estímulo de dos tonos). Por otra parte, son independientes de la amplitud, pero pueden variar considerablemente de individuo a individuo. La banda crítica está

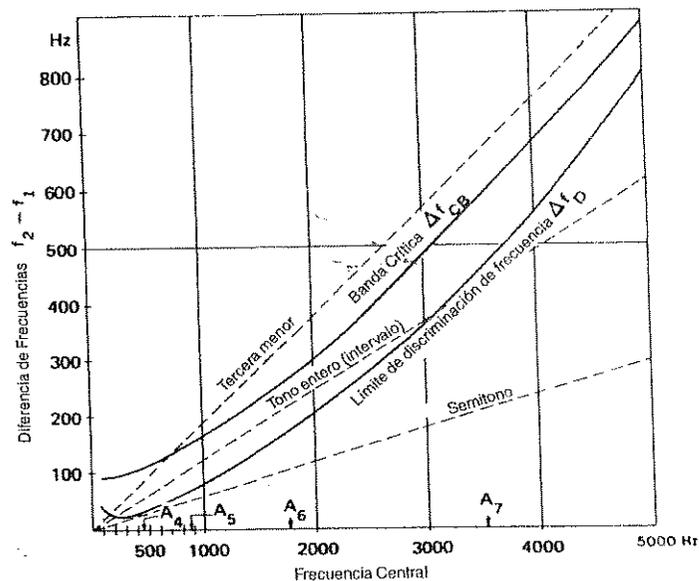


Figura 2.13 Ancho de banda crítica Δf_{CB} (según Zwicker, Flottorp y Stevens, 1957) y límite de discriminación de frecuencia Δf_D (Plomp, 1964), como función de la frecuencia central de un estímulo de dos tonos (escalas lineales). Se muestran, para comparación, las diferencias de frecuencias correspondientes a tres intervalos musicales.

relacionada con otros fenómenos psicoacústicos, y hay otros modos más precisos de definirla experimentalmente (Sec. 3.4). La figura 2.13 muestra la dependencia de la discriminación de altura Δf_D (Plomp, 1964) y de la banda crítica Δf_{CB} (Zwicker, Flottorp, y Stevens, 1957) con la frecuencia central de los tonos componentes. Como dato ilustrativo, las diferencias de frecuencia que corresponden a los intervalos musicales de semitono, tono y tercera menor están indicadas con líneas punteadas. Por ejemplo, en la vecindad de los 2000 Hz, dos tonos deben estar separados un mínimo de 200 Hz para que puedan ser discriminados y de 300 Hz para que suenen suaves. Es notable el hecho de que el límite de discriminación de altura sea más grande que un intervalo de semitono,⁶ y que en ambos extremos de frecuencias muy altas o muy bajas, llega a ser incluso mayor que un

⁶ Esto quizás sea una sorpresa para los músicos: ¡ellos dirán que pueden distinguir muy bien los dos sonidos componentes de una segunda menor tocada con instrumentos musicales! La cuestión es la siguiente: los resultados mostrados en la figura 2.13 solo son válidos para superposiciones de sonidos *simples*, que suenan *de manera uniforme*, con intensidad constante. Cuando un intervalo musical se toca con instrumentos musicales *reales*, los sonidos no son simples, ni suenan todo el tiempo en forma uniforme y además aparece un efecto estereofónico. Todos estos factores dan claves adicionales al sistema auditivo, las cuales son eficientemente utilizadas para la discriminación de los sonidos.

tono. Obsérvese también la importante circunstancia de que para el rango de alturas más agudas, la banda crítica cae entre la diferencia de frecuencias correspondiente a un tono (intervalo considerado «disonante») y la correspondiente a una tercera menor (intervalo considerado «consonante»); es decir que *se extiende por encima de un tercio de octava*. En el rango de frecuencias bajas hay una importante desviación: la discriminación de frecuencias y la banda crítica exceden la tercera menor (incluso una tercera mayor). Este es el motivo por el cual las terceras en general no se usan en el dominio de las alturas graves.

Compárese la Fig. 2.13 con la Fig. 2.9: el límite para la discriminación de frecuencias Δf_D es alrededor de treinta veces más grande que la DAP para la resolución de frecuencia. En otras palabras, podemos detectar cambios de frecuencia muy pequeños en un tono simple *aislado*, pero es necesaria una diferencia de frecuencia apreciable entre *dos* tonos puros que suenen simultáneamente, para que podamos distinguirlos entre sí, separadamente.⁷

¿Qué implican estos resultados para una teoría de la audición? La existencia de un límite finito para la discriminación de tonos es un indicio de que *la región activada sobre la membrana basilar por un tono simple tiene una extensión espacial finita*. Efectivamente, si esta región fuera idealmente «estrecha», dos tonos superpuestos siempre serían percibidos como dos tonos separados, sin importar cuán pequeña sea su diferencia de frecuencias - y nunca aparecería ninguna sensación de batido. El hecho de que la sensación de rugosidad persista aún más allá del límite de discriminación es un indicio de que las dos regiones activadas se superponen o interactúan en cierto grado, por lo menos hasta que se alcanza la diferencia de frecuencias representada por la banda crítica. Un experimento esclarecedor es el siguiente: al recibir *dicóticamente* los dos tonos de frecuencias f_1 y f_2 , es decir, escuchando cada uno de ellos por un oído diferente, la sensación primaria de batido o de rugosidad desaparece completamente; ambos tonos pueden ser discriminados aun cuando la diferencia entre sus frecuencias esté por debajo de Δf_D ¡y la

⁷ Un experimento equivalente simple que demuestra la diferencia entre «resolución» y «discriminación» puede ser realizado con el sentido del *tacto*. Pida a alguien que toque la piel de su antebrazo con un lápiz de buena punta durante un segundo, en un punto dado, *mientras usted mira hacia otro lado*. Luego pídale que repita este proceso en posiciones paulatinamente desplazadas. Se requerirá una cierta distancia, pequeña pero finita, antes de que usted pueda decir que la posición donde se lo ha tocado ha cambiado; esto es la DAP para la localización de una sensación táctil, o «resolución táctil». Ahora pida a la persona que use dos lápices y determine cuán lejos deben estar entre sí los puntos de contacto para que usted empiece a identificar *dos* sensaciones táctiles separadas. Esta es la distancia mínima para la «discriminación táctil», que resulta ser considerablemente mayor que la DAP. El equivalente entre experiencias táctiles y auditivas no es del todo fortuito: la membrana basilar, desde el punto de vista de la evolución biológica, es un trozo de tejido epitelial (piel) con una sensibilidad táctil enormemente amplificadas. Esta analogía ha sido usada profusamente por von Békésy (1960) en los magníficos experimentos que le merecieron el Premio Nobel.

combinación suena suave en todo momento! En cuanto volvemos a escuchar en forma monoaural, los batidos o la rugosidad vuelven. Por supuesto, lo que ocurre en el caso dicótico, es que hay solo *una* región activada en cada membrana basilar, sin posibilidad alguna de superposición de señales en la cóclea⁸; de aquí que no haya batidos ni rugosidad.

El lector puede ahora preguntarse: si la actividad generada en la membrana basilar por *un* tono simple, de *una* única frecuencia, está espacialmente «esparcida», cubriendo un rango finito determinado Δx a lo largo de la membrana, ¿cómo es que llegamos a escuchar solo *una* altura, y no una banda extensa de alturas que corresponden a todos los sensores que han sido activados dentro de Δx ? Lamentablemente, debemos postergar la respuesta a secciones ulteriores (por ej. la 3.6). Solo adelantaremos aquí que existe un proceso de «*agudizamiento*» en el cual la actividad recolectada a lo largo de toda la región Δx de la membrana es «focalizada» en un número mucho menor de neuronas que responden al estímulo.

El fenómeno de los batidos tiene un rol importante en música. Cuando éstos ocurren, son procesados por el cerebro, dando lugar a sensaciones que pueden ir desde lo desagradable o irritante hasta lo placentero y sosegado, según la frecuencia de batido y las circunstancias musicales en que aparecen. El característico y molesto sonido de un instrumento desafinado con respecto a su acompañamiento se debe a los batidos. El sonido feo de una mediocre orquesta de cuerdas se debe, en parte, a los batidos, y el sonido «cómico» de un piano de salón está causado por pares o ternas de cuerdas deliberadamente desafinadas entre sí en los registros medio y superior. El hecho de que los batidos desaparezcan por completo cuando dos tonos tienen exactamente la misma frecuencia ($f_1 = f_2$) tiene un rol práctico para la *afinación* de instrumentos. Si queremos que la frecuencia de un tono dado sea exactamente igual a la de otro establecido como estándar (por ej. *la* del diapasón), podemos hacerlo escuchando los batidos y «ajustando» la frecuencia hasta que los batidos hayan desaparecido completamente.⁹

La banda crítica también tiene un importante rol en la percepción musical. En secciones posteriores discutiremos este concepto con más detalle. Por ahora diremos solamente que la banda crítica representa una especie de «unidad colectora e integradora de información» sobre la membrana basilar. El hecho experimental de que la extensión en frecuencia Δf_{CB} de la banda crítica es más o menos independiente de la amplitud o

⁸ Hay, sin embargo, superposición de señales neurales en las etapas superiores del procesamiento auditivo, que originan los efectos de «segundo orden», que serán considerados en las secciones 2.6 y 2.9.

⁹ En este capítulo hemos analizado los batidos que se originan solo entre sonidos *simples*. Como veremos más adelante, los batidos también aparecen de un modo similar entre los sonidos compuestos de los instrumentos musicales reales.

sonoridad del estímulo es un fuerte indicio de que debe estar relacionada con alguna propiedad inherente del órgano sensorial sobre la membrana basilar, más que con la forma de onda en el fluido coclear. En realidad, si uno convierte el intervalo de frecuencias Δf_{CB} mostrado en la Fig. 2.13 en una extensión espacial a lo largo de la membrana basilar usando la Fig. 2.8, se obtiene un valor casi constante de alrededor de 1,2 mm para la banda crítica. Un resultado aún más significativo es el siguiente: la banda crítica resulta corresponder a una extensión que abarca un número aproximadamente constante de 1300 células receptoras (de un total aproximado de 16000 sobre la membrana) (Zwislocki, 1965), independientemente de la particular frecuencia central (es decir, posición sobre la membrana) involucrada.

Un estímulo sonoro complejo (por ej. dos tonos simples) cuyos componentes tienen frecuencias comprendidas dentro de una banda crítica, causa una sensación subjetiva (rugosidad en nuestro ejemplo) que normalmente es muy diferente de aquella cuando la extensión de frecuencias *excede* la banda crítica (suavidad en el ejemplo de los dos tonos). Esto vale para una variedad de fenómenos. La banda crítica asume un importante rol en la percepción de la cualidad del sonido (Sec. 4.8) y echa las bases para una teoría de consonancia y disonancia de intervalos musicales (Sec. 5.2).

2.5 Otros efectos de primer orden: sonidos resultantes y armónicos aurales

Hasta ahora hemos analizado los efectos de superposición de dos tonos simples cuyas frecuencias son próximas entre sí (Fig. 2.12) ¿Qué ocurre con nuestras sensaciones cuando la frecuencia del tono que varía f_2 se incrementa más allá de la banda crítica, mientras f_1 se mantiene constante? Los efectos resultantes pueden ser clasificados en dos categorías, según su origen esté en el oído o en el sistema nervioso. En esta sección nos ocuparemos de un fenómeno perteneciente a la primera categoría: la percepción de los *sonidos resultantes*. Estos tonos son sensaciones adicionales de altura que aparecen cuando dos tonos puros de frecuencias f_1 y f_2 suenan juntos; se perciben tanto más fácilmente cuanto más alta sea su sonoridad. Estas sensaciones adicionales de altura corresponden a frecuencias que difieren considerablemente de f_1 y de f_2 , lo cual puede comprobarse fácilmente con experimentos de 'apareamiento de altura tonal' y de cancelación de altura (Goldstein, 1970). *Estas frecuencias no están presentes en el estímulo sonoro original* y aparecen como resultado de la así llamada distorsión no lineal de la señal acústica en el oído.

Repitamos el experimento correspondiente a la Fig. 2.12, que consiste

en la superposición de dos tonos simples, pero esta vez incrementando la sonoridad considerablemente, y barriendo la frecuencia f_2 lentamente para arriba y para abajo entre el unísono f_1 y la octava $2f_1$. Prestemos especial atención a las sensaciones de altura que surgen. Por supuesto, oímos el tono de altura constante que corresponde a la frecuencia f_1 , así como el tono variable f_2 . Pero además de éstos se distingue claramente la presencia de uno o más tonos de alturas más bajas que barren para arriba o para abajo, según cómo variemos f_2 . En particular, cuando f_2 asciende, alejándose de f_1 , escuchamos un tono de altura que va subiendo, partiendo desde lo muy profundo. Cuando se varía f_2 en dirección descendente, empezando desde la octava $f_2 = 2f_1$, también escuchamos un sonido cuya altura va subiendo y que parte desde lo muy profundo. Y prestando aún más atención, más de un sonido grave puede ser escuchado al mismo tiempo. Estos, que no existen en absoluto en el estímulo sonoro original, se llaman *sonidos resultantes*.

El sonido resultante más fácilmente identificable, para un nivel *alto* de intensidad de los tonos originales, es aquel cuya frecuencia está dada por la diferencia de las frecuencias componentes

$$f_{C1} = f_2 - f_1 \quad (2.4)$$

Este también se llama *sonido diferencial*. Obsérvese que para valores de f_2 muy cercanos a f_1 , f_{C1} no es otra cosa que la frecuencia de batido (2.3). f_{C1} debe ser por lo menos unos 20 o 30 Hz para ser percibido como un tono. A medida que f_2 sube, f_{C1} también se incrementa. Cuando f_2 está a una octava por encima de f_1 , resulta $f_{C1} = 2f_1 - f_1 = f_1$; es decir, el sonido diferencial coincide con el componente más bajo, f_1 . Cuando el componente f_2 está a mitad de camino entre f_1 y $2f_1$ - es decir, $f_2 = f_1$ (un intervalo musical llamado *quinta*) - el sonido diferencial tiene una frecuencia $f_{C1} = 3/2 f_1 - f_1 = 1/2 f_1$ que corresponde a una altura que suena una octava debajo de aquella correspondiente a f_1 .

Los otros dos sonidos resultantes que son más fáciles de identificar (Plomp, 1965), *incluso a niveles bajos de intensidad* de los sonidos originales, corresponden a las frecuencias

$$f_{C2} = 2f_1 - f_2 \quad (2.5)$$

$$f_{C3} = 3f_1 - 2f_2 \quad (2.6)$$

Ambos sonidos, f_{C2} y f_{C3} , decrecen en altura cuando f_2 se incrementa desde el unísono hacia la quinta, y se escuchan con más facilidad cuando f_2 está aproximadamente entre $1,1f_1$ y $1,3f_1$. Cuando la intensidad de los sonidos originales es alta, estos tonos diferenciales también pueden ser percibidos como tonos graves, cerca de la octava y de la quinta, respectivamente. Nótese que los sonidos f_{C2} y f_{C1} tienen la misma

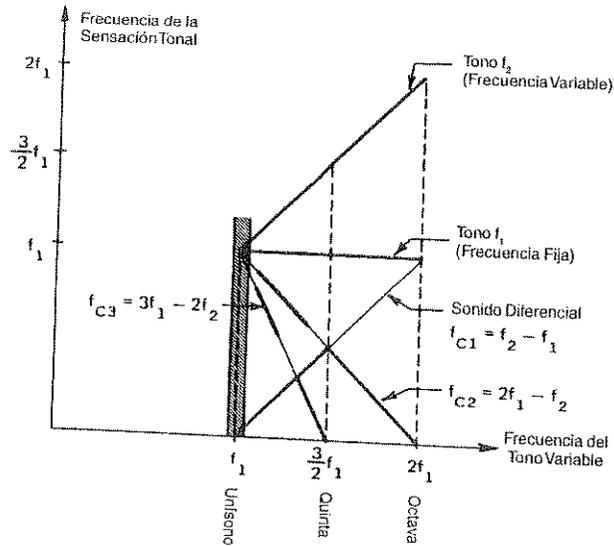


Figura 2.14 Frecuencias de los sonidos resultantes f_{C1} , f_{C2} , f_{C3} , producidos por una superposición de dos tonos (f_1 , f_2). Líneas gruesas: rangos en los cuales los sonidos resultantes se detectan con más facilidad.

frecuencia ($= 1/2f_1$) cuando f_2 está en la quinta, $3/2f_1$. En la Fig. 2.14 resumimos las sensaciones auditivas de primer orden originadas por la superposición de dos sonidos simples de frecuencia f_1 y f_2 . La Fig. 2.12 no es otra cosa que una vista «con lupa» de lo que ocurre cuando la frecuencia f_2 está muy cerca de f_1 (área grisada en la Fig. 2.14). Las porciones de los sonidos resultantes indicadas con trazo más grueso son aquellas más fáciles de percibir (la extensión real depende de la intensidad).

¿Cómo se generan estas sensaciones de tonos adicionales? Como lo señalamos arriba, *no* están presentes en la vibración original del tímpano. Cuidadosos experimentos realizados con animales han revelado que las frecuencias de los sonidos resultantes *no* están presentes ni siquiera a la entrada de la cóclea (membrana de la ventana oval, Fig. 2.6); por otra parte, mediciones directas de los impulsos neurales (Goldstein, 1970) revelan que efectivamente *hay* regiones activadas sobre la membrana basilar en las posiciones correspondientes a las frecuencias de los sonidos resultantes. Se cree que éstos son causados en la cóclea por una distorsión «no lineal» de la forma de onda primaria que actúa como estímulo. Efectivamente, se puede demostrar matemáticamente que cuando dos oscilaciones armónicas (sinusoidales) de diferentes frecuencias f_1 y f_2 pasan a través de un transductor de respuesta distorsionada (no lineal), la salida contendrá, además de las frecuencias originales f_1 y f_2 , todas las

combinaciones lineales del tipo $f_2 - f_1$, $2f_1 - f_2$, $3f_1 - 2f_2$, $f_2 + f_1$, $2f_2 + f_1$, etc. Experimentos posteriores (Smoorenburg, 1972) indican, sin embargo, que el sonido diferencial (2.4) y los otros dos sonidos resultantes (2.5, 2.6) deben originarse en mecanismos cocleares diferentes. El umbral de intensidad para la generación del primero es considerablemente más alto que el de los últimos, y relativamente independiente de la razón de frecuencias f_2/f_1 . Por otra parte, la intensidad de los últimos aumenta a medida que f_2/f_1 se aproxima a la unidad.¹⁰

Es interesante observar, que a causa de la distorsión no lineal, incluso un *único* tono de frecuencia f_1 puede originar sensaciones de altura adicionales, si se produce con gran sonoridad. Estos tonos adicionales, llamados *armónicos aurales*, corresponden a frecuencias que son múltiplos enteros de la frecuencia original: $2f_1$, $3f_1$, $4f_1$, etc.

Aunque para realizar los experimentos pertenecientes a esta sección lo ideal sería contar con generadores electrónicos, es posible, al menos cualitativamente, explorar los sonidos resultantes y los armónicos aurales usando instrumentos musicales capaces de emitir tonos estáticos a alto nivel de intensidad. Tal vez el «instrumento» más adecuado para este propósito sea un silbato de perro cuya altura (aguda) pueda ser variada. Un simple experimento casero sobre sonidos resultantes puede ser hecho soplando al mismo tiempo dos de tales silbatos - uno a altura constante, el otro barriendo la frecuencia de tal manera que se aleje y se acerque al unísono - y escuchando las sensaciones de altura que surgen en el ámbito grave. Los sonidos resultantes no tienen un rol importante en la música, pues para que aparezcan se necesitan niveles de intensidad relativamente altos. Cuando se escucha música de rock a los usuales niveles de exagerada sonoridad (dañosos fisiológicamente), especialmente por auriculares, ¡sonidos resultantes sin duda aparecen copiosamente!

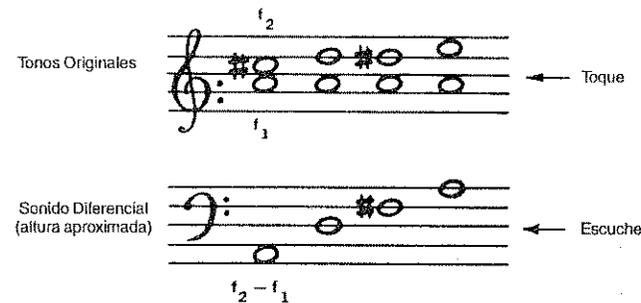


Figura 2.15 Sonidos diferenciales percibidos (pentagrama inferior) cuando se tocan las superposiciones de dos tonos indicadas en el pentagrama superior (¡fuerte!)

¹⁰No se conocen aún claramente las razones de estas diferencias; tampoco se sabe por qué las combinaciones $f_1 + f_2$, $2f_1 + f_2$, etc. no aparecen como sensaciones sonoras.

Sonidos resultantes «falsos» pueden ser fácilmente generados en órganos electrónicos, amplificadores y parlantes de baja calidad. En estos casos, es una distorsión no lineal en el circuito electrónico y en el sistema mecánico de los parlantes, la que crea estas frecuencias parasitarias. En particular, el tono diferencial puede ser producido y percibido con claridad con un órgano electrónico: suba el volumen, use registros de flauta de 8' y toque para atrás y para adelante la secuencia mostrada en el pentagrama superior de la Fig. 2.15, mientras intenta escuchar los tonos indicados en el pentagrama inferior.

Algunos de estos sonidos diferenciales aparecen «desafinados» debido al sistema de afinación del instrumento, basado en el temperamento igual (Sec. 5.3). Debemos enfatizar nuevamente que lo que se escucha en este experimento es un sonido resultante *falso*, en el sentido de que la sensación de altura grave se genera en el parlante y *no* en el oído. Con este ejemplo queda claro por qué se les exige a los circuitos electrónicos y los parlantes de los sistemas de alta fidelidad una respuesta «muy» lineal.

2.6 Efectos de segundo orden: batidos de consonancias desafinadas

Repetimos ahora el experimento de la sección precedente con dos tonos simples generados electrónicamente, pero esta vez ignorando las posibles sensaciones de sonidos resultantes. Enviamos ambos tonos con un bajo nivel de intensidad monoauralmente a un mismo oído; el tono f_1 se mantiene con frecuencia constante, mientras que f_2 puede ser, como antes, variado a voluntad. Las amplitudes de los dos tonos simples se mantienen constantes durante todo el experimento. Cuando barremos f_2 lentamente hacia arriba, advertimos algo peculiar al pasar por las cercanías de la octava $f_2 = 2f_1$: una sensación de batido diferente, distinta a la de los batidos de primer orden cerca del unísono, pero claramente perceptible. Cuando f_2 es exactamente igual a $2f_1$, esta sensación de batido desaparece. Reaparece tan pronto como desafinamos la octava, es decir, cuando $f_2 = f_1 + \varepsilon$, donde ε (épsilon) representa solo unos pocos Hertz. La frecuencia de batido resulta ser igual a ε . Es difícil describir «qué es» lo que está batiendo. La mayoría de los sujetos dicen que se trata de un batido en la «cualidad» del tono. Designaremos a este tipo de batidos de *segundo orden*; algunos prefieren llamarlos «batidos subjetivos». Son el resultado del procesamiento neural.

Es instructivo observar el patrón vibratorio en el osciloscopio mientras se escuchan los batidos de segundo orden. Se ve a este patrón cambiar en exacto sincronismo con las sensaciones de batido. Obviamente, nuestro sistema auditivo de algún modo es capaz de detectar estos cambios de *forma* en un patrón vibratorio. La Fig. 2.16 muestra varios patrones

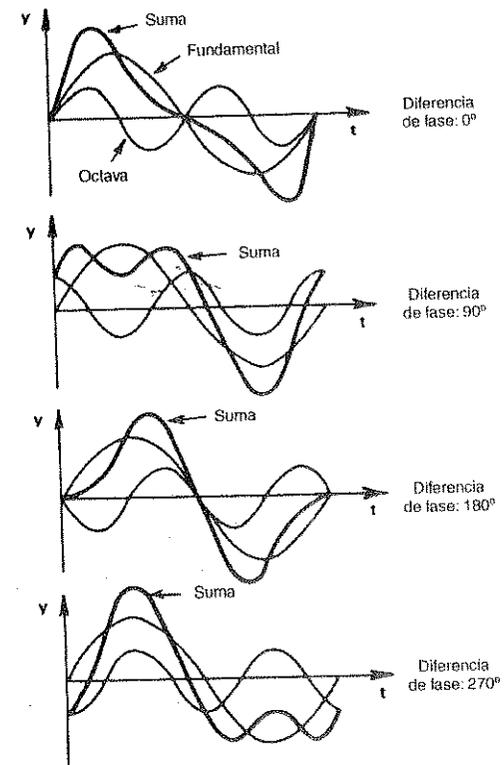


Figura 2.16 Superposiciones de octava de dos tonos puros, graficadas para cuatro diferencias de fase distintas.

vibratorios correspondientes a la superposición de un sonido fundamental de frecuencia f_1 y su octava $f_2 = 2f_1$ (de menor amplitud), para cuatro valores distintos de diferencia de fase. Cuando la octava está perfectamente afinada, la diferencia de fase permanece constante y la imagen en el osciloscopio, estática; cualquiera de las cuatro superposiciones suena igual que las otras; nuestro oído no distingue un caso de otro. Pero cuando desafinamos levemente f_2 : $f_2 = 2f_1 + \varepsilon$, la diferencia de fase cambiará lentamente con el tiempo, y el patrón vibratorio resultante sufrirá un cambio gradual, pasando de una de las formas mostradas en la Fig. 2.16 a la siguiente. Puede demostrarse matemáticamente que este ciclo de cambio del patrón vibratorio se repite con una frecuencia ε , la cantidad por la cual el tono superior está desafinado con respecto a la octava. Esto, obviamente, significa que el oído es sensible a cambios lentos de la

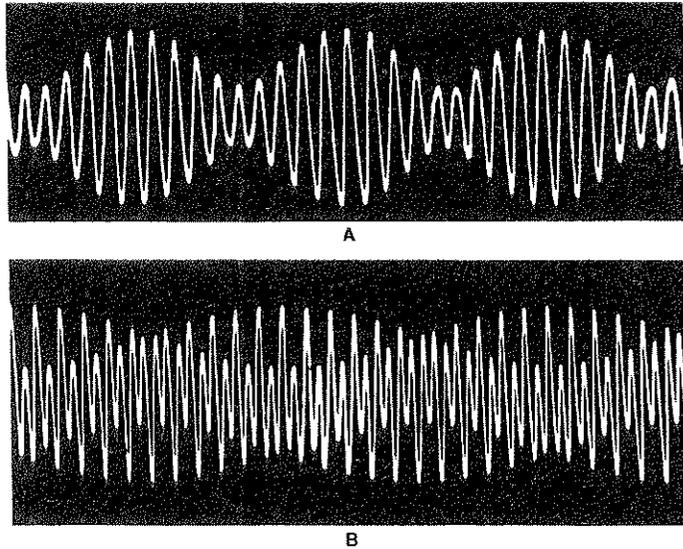


Figura 2.17 Comparación de batidos de primer y segundo orden. (A) Batidos de primer orden (unísono desafinado); modulación en amplitud sin cambio en el patrón vibratorio. (B) Batidos de segundo orden (octava desafinada); modulación del patrón sin cambio en la amplitud total.

diferencia de fase entre dos tonos.¹¹ Dicho en otras palabras: *el sistema auditivo es capaz de detectar cambios cíclicos de la forma de un patrón vibratorio*. Obsérvese con cuidado que no hay ningún cambio macroscópico en la amplitud de un patrón con respecto a otro en la Fig. 2.16, bien al contrario de lo que ocurre con los batidos de primer orden, que son cambios cíclicos en la amplitud del patrón vibratorio (Fig. 2.11). La Fig. 2.17 muestra dos ejemplos, sacados de la pantalla de un osciloscopio, que representan los batidos de primer orden, en las cercanías del unísono, y los batidos de segundo orden de una octava desafinada, respectivamente. Obsérvese la modulación de amplitud en el primero, y la modulación del patrón vibratorio en el segundo. Es importante señalar que los batidos de segundo orden solo aparecen cuando el estímulo original formado por los dos tonos está en el rango de frecuencias bajas. Los batidos de segundo orden no se perciben más cuando f_1 (y f_2) exceden los 1.500 Hz, aproximadamente (Plomp, 1967).

¹¹ Cambios *repentinos* de fase (por ej. presentación de los estímulos de octava de la Fig. 2.16 alternándolos con un estímulo de octava tomado como referencia) también son detectados. El máximo grado de detección ocurre cuando la diferencia de fase con la señal de referencia es de 180° (Raiford y Schubert, 1971).

Volvamos a nuestro experimento y exploremos el rango total de frecuencias entre el unísono y la octava. Descubrimos que hay otros pares de valores para f_2 y f_1 , es decir, otros intervalos musicales, en cuyas vecindades aparecen sensaciones de batido, aunque mucho menos perceptibles que para la octava. Dos «pozos de batido», como podríamos llamarlos, se pueden detectar centrados en las frecuencias $f_2 = \frac{3}{2}f_1$ y $f_2 = \frac{4}{3}f_1$, que corresponden a los intervalos musicales de quinta y cuarta, respectivamente. Al observar el patrón vibratorio en el osciloscopio, al mismo tiempo de escuchar la señal acústica resultante, nos damos cuenta de que para una quinta desafinada ($f_2 = \frac{3}{2}f_1 + \varepsilon$) y para una cuarta desafinada ($f_2 = \frac{4}{3}f_1 + \varepsilon$), la forma del patrón vibratorio no es estática (como ocurre con una quinta o cuarta exactas, o sea para $\varepsilon = 0$), sino que cambia periódicamente en su forma (no en su amplitud). Los batidos de segundo orden de la quinta son más rápidos que aquéllos de la octava (para la quinta la frecuencia de batido es $f_B = 2\varepsilon$, y para la cuarta es $f_B = 3\varepsilon$). No es ésta la única razón por la cual son más difíciles de percibir: el patrón vibratorio mismo se vuelve cada vez más complicado a medida que pasamos de la octava (Fig. 2.16) a la quinta y a la cuarta. Cuanto más complejo es un patrón vibratorio, tanto más difícil será para el sistema auditivo detectar su cambio periódico. Para una discusión detallada sobre los batidos de consonancias desafinadas, véase Plomp (1976).

Existe una relación óptima entre las intensidades de los tonos componentes, para la cual los batidos de segundo orden son más pronunciados; en ella *el tono de frecuencia más alta (f_2) tiene intensidad más baja* (Plomp, 1967). Finalmente, es importante observar que los batidos de segundo orden también son percibidos cuando los tonos componentes es escuchan separadamente, *cada uno por un oído diferente*. En ese caso se experimenta una extraña sensación de rotación de la imagen sonora «dentro» de la cabeza (Sec. 2.9).

Los batidos de segundo orden producidos por consonancias desafinadas de tonos *simples* no tienen importancia en música (principalmente porque los tonos simples no la tienen). Su estudio, sin embargo, es un paso importante para la comprensión del mecanismo de procesamiento del sonido musical (Sec. 2.8).

2.7 Seguimiento de la fundamental [Fundamental Tracking]

Introducimos ahora otra serie de experimentos psicoacústicos que han sido de crucial importancia para las teorías de percepción auditiva. Consideremos dos tonos simples, separados por una quinta justa, de

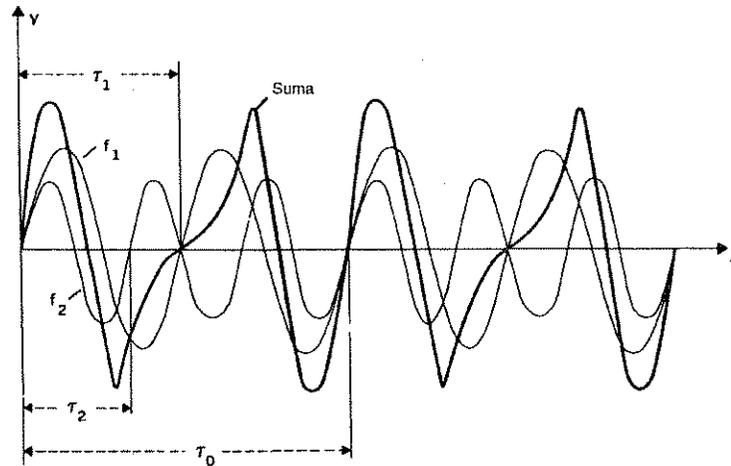


Figura 2.18 Superposición de dos tonos puros separados por un intervalo musical de quinta (con una determinada relación de fase). τ_0 : período de repetición del sonido resultante.

frecuencias f_1 y $f_2 = \frac{3}{2}f_1$. La Fig. 2.18 muestra la vibración resultante (la suma) para una relación de fase particular. Obsérvese que el patrón se repite exactamente con la misma forma después de un tiempo τ_0 , que es un intervalo de duración doble que τ_1 , el período del tono de altura más grave. Esto quiere decir que la frecuencia de repetición $f_0 = 1/\tau_0$ del patrón vibratorio de una quinta es igual a la mitad de la frecuencia del tono más grave:

$$f_0 = \frac{1}{2}f_1 \tag{2.7a}$$

Llamamos a esta frecuencia de repetición «frecuencia fundamental» del patrón vibratorio. En este caso, se encuentra una octava debajo de f_1 . Si tomamos dos tonos que forman una cuarta ($f_2 = \frac{4}{3}f_1$) podemos graficar el patrón vibratorio resultante de la misma manera que para la quinta. La frecuencia de repetición resultante es ahora

$$f_0 = \frac{1}{3}f_1 \tag{2.7b}$$

es decir, una duodécima debajo del tono más bajo. Para una tercera mayor ($f_2 = \frac{5}{4}f_1$), la frecuencia de repetición está dos octavas abajo de f_1 :

$$f_0 = \frac{1}{4}f_1 \tag{2.7c}$$

Nuestro sistema auditivo resulta ser sensitivo a estas frecuencias de repetición. Efectivamente, se ha realizado un cuidadoso experimento en el cual se presentaban cortas secuencias de estímulos a título de melodía; cada estímulo consistía de pares de tonos simples sonando simultáneamente y formando intervalos de quinta, cuarta, y tercera. (Houtsma y Goldstein, 1972). Se pidió a los sujetos que identificaran una única altura básica para los tonos de la «melodía» presentada. ¡La mayoría identificó una altura cuya frecuencia correspondía a las definidas por las relaciones (2.7a), (2.7b) o (2.7c), respectivamente!¹² Es importante insistir que este experimento de identificación de altura requiere que los pares de tonos sean presentados como *secuencia temporal* o melodía. (Cuando uno de estos pares de tonos simples se mantiene constante, nuestro sistema auditivo falla en su «búsqueda» de una sensación de altura única; reorienta su atención con el fin de discriminar las alturas espectrales de los dos tonos componentes por separado, como fue explicado en la Sec. 2.4).

Obsérvese que las frecuencias de repetición (2.7a) - (2.7c) de las combinaciones de dos tonos son idénticas a las frecuencias de los sonidos resultantes correspondientes (por ej. vea los casos cuarto, segundo y primero en la Fig. 2.15). Sin embargo, los experimentos han demostrado que la percepción de la frecuencia de repetición es posible aún cuando las intensidades de los dos tonos f_1 y f_2 son bajas, por debajo del umbral de intensidad para la producción de sonidos resultantes. Debemos descartar, pues, un sonido diferencial (2.4) (Plomp, 1967a). En realidad, el fenómeno de percepción de la frecuencia de repetición ha sido usado en música durante muchos siglos (y equivocadamente atribuido al efecto de un sonido resultante). Por ejemplo, desde fines del siglo XVI, muchos órganos incluyen un juego (registro de $5\frac{1}{3}$) compuesto por tubos que suenan una quinta más alta que la altura de la nota escrita. El propósito es estimular o reforzar el bajo, una octava debajo (ecuación [2.7a]) de la altura de la nota escrita (es decir, reforzar el sonido de $16'$ del órgano). Más antiguo aun es el uso del juego de $10\frac{2}{3}$, que en combinación con los registros de $16'$, refuerza el bajo de $32'$ (dos octavas debajo de la nota escrita).

El tono de frecuencia f_0 (2.7) no está presente en el estímulo original. Este tono es llamado la *fundamental ausente* (por razones que se aclararán posteriormente); la sensación de altura correspondiente se llama *altura de periodicidad* («periodicity pitch»), *altura subjetiva*, *tono residual* o *altura virtual*. Esta sensación de altura debe distinguirse claramente de la altura primaria o espectral de cada uno de los dos tonos simples

¹² Téngase en cuenta que este experimento debe ser realizado con pares de sonidos sinusoidales, generados electrónicamente; el experimento no funcionará con el piano o cualquier otro instrumento musical. Ver, sin embargo, las observaciones subsecuentes sobre los órganos.

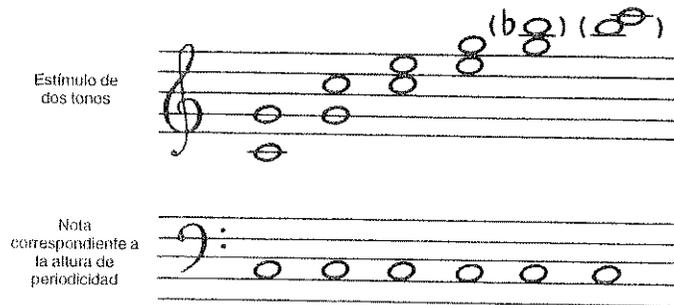


Figura 2.19 Estímulos de dos tonos simples (*pentagrama superior*) que originan la misma altura de periodicidad (*pentagrama inferior*). El Sib que aparece entre paréntesis debe afinarse más bajo que el correspondiente a la escala en uso (Sec. 5.3), con el fin de producir un Do como altura de periodicidad.

originales. Se ha demostrado por medio de experimentos que para niveles normales de sonoridad, la frecuencia f_0 no está presente ni siquiera en las oscilaciones del fluido coclear (mientras que la de los sonidos resultantes sí lo está). Efectivamente, la región de la membrana basilar que corresponde a la frecuencia f_0 (Fig. 2.8) puede ser saturada (enmascarada) con una banda de ruido (sonido de un infinito número de frecuencias componentes dentro de un rango determinado) de manera que cualquier excitación adicional de aquella región pasaría inadvertida; aún así, la fundamental ausente se percibe (Small, 1970). O se puede introducir un tono adicional levemente desafinado con respecto a f_0 ; si la fundamental ausente, tono f_0 , realmente existiera en la cóclea, deberían aparecer batidos de primer orden; sin embargo, no aparece ninguna sensación de batido. Un efecto todavía más notable es que *la fundamental ausente se percibe aun cuando los dos tonos componentes son enviados por separado*, uno a cada oído (Houtsma y Goldstein, 1972). Todo esto indica que la fundamental ausente, o altura de periodicidad, debe ser el resultado de un procesamiento neural a un nivel más alto.

La detección de altura subjetiva, es decir la capacidad de nuestro sistema auditivo de identificar la frecuencia de repetición de un patrón vibratorio que llega a nuestros oídos, solo funciona en el rango más bajo de frecuencias (pero musicalmente más importante), por debajo de los 1500 Hz aproximadamente. Cuanto más complejo es el patrón vibratorio, es decir, cuanto más pequeño es el intervalo entre los tonos componentes, tanto más difícil es para el sistema auditivo identificar la fundamental ausente, y la sensación de altura subjetiva se torna más ambigua.

Volvamos ahora a las relaciones (2.7) e investiguemos qué pares de frecuencias de tonos simples tienen la *misma* frecuencia de repetición o *frecuencia fundamental* f_0 . Obtenemos

$$\begin{array}{cc} \underbrace{2f_0 \text{ y } 3f_0}_{\text{quinta}} & \underbrace{3f_0 \text{ y } 4f_0}_{\text{cuarta}} \\ \underbrace{4f_0 \text{ y } 5f_0}_{\text{tercera mayor}} & \underbrace{5f_0 \text{ y } 6f_0}_{\text{tercera menor}} \quad \text{etc.} \end{array}$$

En otras palabras, si f_0 corresponde a la nota mostrada en el pentagrama inferior de la Fig. 2.19, los intervalos musicales mostrados en el pentagrama superior producen la misma nota como sensación de altura subjetiva. Es importante recordar que las notas de la Fig. 2.19 representan tonos *simples*, de una sola frecuencia, y *no* tonos compuestos producidos por instrumentos musicales reales.

Los componentes individuales de frecuencias $2f_0, 3f_0, 4f_0, 5f_0, \dots$, etc., se llaman *armónicos superiores* de la frecuencia fundamental f_0 . Las frecuencias de los armónicos superiores son múltiplos enteros de la frecuencia fundamental. Dos tonos sucesivos cualesquiera de la serie de armónicos superiores forma un par con la misma frecuencia de repetición, o frecuencia fundamental f_0 . De aquí que todos los armónicos superiores, al sonar juntos, produzcan una única sensación de altura que corresponde a f_0 - ¡incluso si esta última frecuencia está ausente en el conjunto de sonidos que sirve de estímulo! Esta es la razón por la cual en los ejemplos anteriores f_0 fue también llamada «fundamental ausente», y por la cual el proceso de percepción de esta frecuencia de repetición se llama *seguimiento de la fundamental*. Obsérvese una vez más la sorprendente propiedad de este conjunto de tonos puros de frecuencias $2f_0, 3f_0, 4f_0, \dots, nf_0 \dots$. Entre la infinita variedad de superposiciones posibles de tonos simples, este conjunto, y solo éste, tiene componentes que, tomados por pares contiguos, producen una misma frecuencia de repetición. Recíprocamente, ésta es la razón por la cual *cualquier* sonido periódico, con un patrón vibratorio complejo pero repetitivo (con una frecuencia de repetición f_0), está formado por una superposición de tonos simples de frecuencias nf_0 ($n =$ número entero) (véase Sec. 4.2).

Los experimentos psicoacústicos arriba mencionados con pares de tonos han sido extendidos a secuencias o melodías con estímulos complejos *multitónicos*, empezando por el n ésimo armónico (es decir superposiciones de tonos puros de frecuencias $nf_0, (n+1)f_0, (n+2)f_0$, etc.). Aunque aquí también el tono de frecuencia fundamental está ausente, la altura subjetiva asignada a estos complejos sonoros es siempre f_0 . De hecho, cuanto más armónicos se incluyen, la altura de periodicidad es percibida con mayor claridad (excepto cuando el orden de armónico con que empieza el estímulo sea muy alto, es decir para un valor alto de n). Los pares de armónicos vecinos más importantes para la percepción de la altura de periodicidad son aquellos alrededor de $n = 4$ (Ritsma, 1967). Dado que los sonidos musicales reales están contruidos con

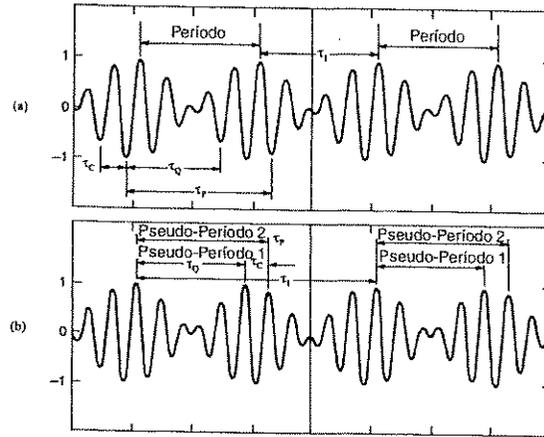


Figura 2.20 Patrón vibratorio de dos tonos puros simultáneos. (a) Los tonos son armónicos vecinos ($n = 4$). (b) Los tonos tienen la misma diferencia de frecuencia que en (a), pero no son armónicos vecinos. τ_1 : período que corresponde a la repetición exacta del patrón vibratorio. f_c : frecuencia central. τ_p y τ_q : pseudoperíodos (ver texto).

superposiciones de armónicos (Cap. 4), *el seguimiento de la fundamental es el mecanismo auditivo que nos permite asignar una sensación de altura única a un tono compuesto de un instrumento musical*.¹³

Es importante comprender todo lo que implica el seguimiento de la fundamental para la teoría de la audición. Un breve análisis de los clásicos experimentos sobre apareamiento de altura tonal [pitch-matching experiments] de Smoorenburg (1970) servirá ese propósito.¹⁴ Considérese un estímulo de corta duración que consiste de dos tonos, cuyas frecuencias f_a y f_b difieren entre sí en una cierta cantidad fija $\Delta f = f_b - f_a$. Al ser presentado este estímulo en algún contexto específico, como una corta melodía, aproximadamente la mitad de los sujetos del experimento percibe una altura «residual» subjetiva, baja y claramente identificable (los demás sujetos solo son capaces de identificar una o ambas alturas del estímulo original). Los experimentos muestran que si f_a y f_b corresponden a *dos armónicos vecinos* de un tono compuesto (por ej. de orden n y $n + 1$), la altura subjetiva, cuando se la percibe, es aquella correspondiente a la

¹³Tal vez el ejemplo más convincente de seguimiento de la fundamental de tonos compuestos esté dado por el hecho de que al escuchar una radio a transistores barata, somos todavía capaces de percibir correctamente la altura de los sonidos graves, ¡a pesar de que todas las frecuencias inferiores a los 100 o 150 Hz están siendo eliminadas por lo inadecuado del parlante y los circuitos!

¹⁴Lamentablemente, estos interesantes experimentos no pueden ser demostrados fácilmente, ni siquiera en un laboratorio destinado a la enseñanza de la física bien equipado.

fundamental ausente $f_1 (= f_a/n = \Delta f)$. Por ejemplo si $f_a = 800$ Hz y $f_b = 1000$ Hz ($n = 4$ y $\Delta f = 200$ Hz), la altura residual escuchada es la que corresponde a una nota de 200 Hz. La figura 2.20(a) muestra el patrón vibratorio del estímulo de estos dos tonos, que nos recuerda el fenómeno de los batidos de primer orden (Sec. 2.4). En este caso, sin embargo, la modulación en amplitud («envolvente» de la curva) es muy rápida (200 veces por segundo) y no se percibe como batido. Más bien, lo que se percibe (por el 50% de los sujetos del experimento) es una altura que corresponde a la frecuencia de repetición del patrón vibratorio, que es exactamente de 200 Hz. El período de repetición correspondiente es $\tau_1 = 1/\Delta f = 1/(f_b - f_a)$ (el período de la fundamental ausente), indicado en la figura (también marcamos el otro período, mucho más corto, presente en la estructura fina temporal del patrón vibratorio, que corresponde a la llamada frecuencia central del estímulo de los dos tonos, $f_c = (f_a + f_b)/2$). Obviamente, nuestro sistema auditivo extrae información del cambio periódico del patrón vibratorio, como lo hace cuando se perciben batidos (lentos) de consonancias desafinadas (Fig. 2.17). Con otro par de armónicos vecinos, como ser, por ejemplo, $f_a = 2000$ Hz y $f_b = 2200$ Hz ($n = 10$), se percibe la *misma* altura subjetiva (con mayor dificultad a medida que n aumenta - Houtsma, 1970); el patrón vibratorio tiene exactamente la misma envolvente que la mostrada en la Fig. 2.20(a), pero dado que ahora la frecuencia central es mucho más alta, la curva de la vibración tiene más oscilaciones (diez, el orden armónico de f_a) dentro de un período de repetición.

Una complicación interesante surge cuando el par f_a, f_b no corresponde a dos armónicos vecinos de una fundamental. Considérese un estímulo de dos tonos $f_a = 900$ Hz y $f_b = 1100$ Hz. No hay ningún sonido musical del cual éstos pudieran ser armónicos *vecinos*; más bien, ellos son el 9º y el 11º armónico de un tono de frecuencia fundamental $f_1 = 100$ Hz. ¿Es percibida esta altura? ¡No! La percepción resulta ambigua: *¿dos alturas posibles* pueden detectarse (dependiendo del contexto en el cual el estímulo es presentado), que corresponden aproximadamente a 180 o 220 Hz! La figura 2.20(b) muestra el patrón vibratorio para este caso. En primer lugar, obsérvese que el patrón vibratorio muestra una envolvente con el mismo período de modulación que el de la figura 2.20(a) (es decir, correspondiente a una frecuencia de 200 Hz). En segundo lugar, examínese cuidadosamente el gráfico de picos y valles y nótese que la estructura se repite en forma *exacta* con un período que es el *doble* de largo (es decir, correspondiente a 100 Hz). La frecuencia de repetición resulta ser 100 Hz; pero ni ésta ni la frecuencia de modulación de 200 Hz son percibidas. ¡En realidad, resulta que las dos alturas que se perciben corresponden exactamente a los dos «seudoperíodos» τ_p y τ_q marcados en la figura! Más aun, los experimentos de Smoorenburg muestran que incluso en el primer ejemplo (cuando el estímulo consiste de dos armónicos vecinos)

pueden escucharse alturas adicionales ambiguas, que corresponden a los pseudoperíodos definidos por los intervalos de tiempo entre el pico central y los picos secundarios del período de modulación subsecuente (Fig. 2.20 [a]). Todo esto es indicio de un proceso de extracción de altura mucho más sofisticado que el de detección de frecuencias de repetición: el mecanismo de audición es capaz de detectar rasgos característicos muy detallados del patrón vibratorio.

Sin embargo, hay una explicación alternativa a este mecanismo de extracción de altura basado en lo temporal, que puede ser invocada para explicar los resultados antedichos. Es un proceso espectralmente basado, en el cual el patrón *espacial* producido por el estímulo de los dos tonos a lo largo de la membrana basilar es analizado y apareado lo más posible a configuraciones «familiares» (por ej. a las posiciones de las regiones de resonancia de armónicos vecinos). Cuando se logra una correspondencia «aceptable» es posible, pueden resultar alturas ambiguas. Puede probarse matemáticamente que este proceso conduce a los mismos resultados cuantitativos que el mecanismo de análisis basado en claves temporales. Continuaremos desarrollando este tema en la Sec. 4.8 y en el Apéndice II.

Finalmente, la disparidad entre distintos individuos en lo que concierne a la habilidad de percibir alturas residuales subjetivas ha sido invocada por algunos psicoacústicos como un llamado de atención, poniendo en duda la interpretación de los experimentos de Smoorenburg. Sin embargo, la consistencia de los resultados cuantitativos (ya verificados por varios grupos de investigación independientes) para aquellos sujetos que *sí* escuchan la altura residual es tan notable, que la mencionada disparidad simplemente puede estar indicando diferentes *estrategias* de audición, sin mayores consecuencias para las conclusiones extraídas en estos experimentos sobre el mecanismo de extracción de altura *per se*. Para una excelente reseña histórica de los experimentos más importantes sobre percepción de altura, véase Plomp (1976).

2.8 Codificación auditiva en el sistema nervioso periférico

El descubrimiento de los efectos de segundo orden en el procesamiento auditivo, tales como la percepción de batidos en consonancias desafinadas y el seguimiento de la fundamental, han tenido un gran impacto sobre la teoría de la audición. Por un lado, la percepción de batidos en consonancias desafinadas (Sec. 2.6) es un indicio de que el sistema auditivo obtiene y utiliza información sobre los detalles de la estructura temporal del patrón de vibración acústico. Por otro lado, el seguimiento de la fundamental podría, en principio, implicar dos alternativas: 1) un mecanismo que ejecuta un análisis detallado del patrón *temporal* de vibración, con la instrucción

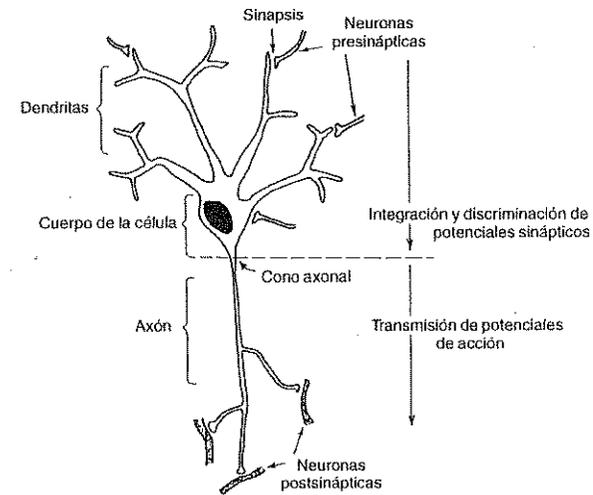


Figura 2.21 Diagrama de una neurona modelo.

de detectar rasgos repetitivos cuya frecuencia conduce luego a la sensación de altura (Fig. 2.20); o 2) un mecanismo que analiza información sobre los detalles del patrón de excitación *espacial* suscitado a lo largo de la membrana basilar, con la instrucción de producir una sensación de altura única si ese patrón se corresponde, al menos en parte, con la excitación característica producida por un sonido musical.¹⁵ Debemos anticipar que la segunda alternativa debería funcionar mejor en la región de los armónicos más bajos (bajos valores de n), donde las máximas excitaciones espaciales correspondientes son más pronunciadas. Sea cual fuere, ambas alternativas implican que los detalles de información acústica codificada en la periferia deben ser analizados a un nivel más alto en el sistema nervioso central.

Con el fin de comprender los mecanismos subyacentes que han sido propuestos para explicar estos fenómenos psicoacústicos, es necesario analizar primero algunos aspectos operativos del sistema nervioso auditivo.

La unidad fundamental de procesamiento y transmisión del sistema nervioso es la célula nerviosa o *neurona*. Una neurona «modelo» se muestra en la Fig. 2.21. Se distingue el *cuerpo* de la célula o soma, una serie de ramificaciones llamadas *dendritas*, y una fibra alargada, el *axón*,

¹⁵ En visión, un proceso equivalente de correspondencia espacial nos permite interpretar los siguientes símbolos incompletos correctamente como letras del alfabeto: AFDJ. Pero un símbolo parcial como puede ser A podría conducir a diferentes interpretaciones, dependiendo del contexto en que dicho símbolo aparece: Andrés de Hungría.

que también puede dividirse en múltiples ramas. Las dendritas y el cuerpo de la célula son los *receptores* de señales neurales entrantes y el axón es el *transmisor*, que pasa señales a otras neuronas. Estas señales neurales son impulsos eléctricos (variaciones en el voltaje eléctrico) de varias decenas de milivoltios; se originan en el flujo de átomos cargados eléctricamente (iones) a través de la membrana de la célula. Estos impulsos pueden ser registrados implantando microelectrodos dentro de la neurona (un procedimiento que no inhibe el normal funcionamiento de la célula). En el axón cada impulso, llamado *potencial de acción*, tiene casi la misma forma y duración (varias decenas de milisegundos), y se propaga desde el cuerpo de la célula (desde el cono axonal) hacia las terminaciones del axón. El potencial de acción representa el mensaje de salida neural fundamental, elemental. Un mensaje neural «integrado» está dado por la *frecuencia* o la *distribución en el tiempo* con la cual los impulsos individuales son disparados a lo largo del axón.

El axón está en contacto con dendritas o cuerpos de células de otras neuronas. El punto de contacto se llama *sinapsis*. Un axón determinado puede estar en contacto sináptico con muchas otras células; recíprocamente, una célula dada puede estar conectada a axones que llegan de otras cientos o miles de células.¹⁶ Un impulso que llega a un contacto sináptico causa la liberación de una sustancia química desde la célula presináptica hacia el espacio entre las membranas de ambas células (espacio sináptico). La presencia de esta sustancia dispara un impulso eléctrico en la célula postsináptica: el «potencial postsináptico». Los potenciales sinápticos son señales analógicas de forma variada y de duración más larga que los potenciales de acción digitales de tamaño standard, que se propagan a lo largo del axón. Hay dos tipos distintos de sinapsis, *excitatoria* e *inhibitoria*, que evocan potenciales sinápticos de polaridad mutuamente opuesta. Si dentro de un corto intervalo de tiempo dado, una neurona recibe un número de estimulaciones excitatorias que supera en un cierto valor *umbral* al de las señales inhibitorias que llegan simultáneamente, ésta responderá disparando un impulso a través de su axón. De lo contrario quedará en silencio. Concluimos, pues, que las dendritas y el cuerpo de la célula funcionan como el sistema de recolección e integración de información de cada neurona, representando de este modo la unidad de procesamiento de información fundamental en el sistema nervioso. Es importante señalar que el hecho de que una neurona dispare una señal de salida dependerá de las distribuciones espacial y temporal de las señales que llegan desde las neuronas presinápticas. Por otra parte, muchas neuronas disparan espontáneamente potenciales de acción a una frecuencia determinada. Una misma neurona solo puede impartir órdenes

¹⁶ Son excepciones las neuronas unipolares que forman parte del sistema de fibras nerviosas aferentes.

excitatorias o bien inhibitorias a otras neuronas. Cuando una neurona inhibitoria dispara un impulso a otra neurona inhibitoria, el efecto inhibitorio de la última se *cancela*. Existen indicios de que la respuesta de las neuronas conectadas a la hilera interna de células pilosas es excitatoria, mientras que la de las conectadas a la hilera externa es inhibitoria (Sokolich y Zwislocki, 1974).

Un retraso temporal característico (usualmente menor al milisegundo) se da entre la llegada de un impulso a una sinapsis y la formación de la respuesta en la célula postsináptica. Esta demora hace posible que la activación nerviosa producida por un único estímulo externo pueda subsistir o «reverberar» un tiempo considerable cuando se propaga por una serie de miles de pasos sinápticos sucesivos en el tejido cerebral. Tal «reverberación» puede ser un proceso clave en el proceso de formación de patrones de actividad nerviosa que dependen del tiempo (por ej. Sec. 4.10). Después de cada activación, una neurona tiene un período refractario durante el cual no puede ser reexcitada, o durante el cual su umbral de disparo es mayor. Una neurona sola es solo un componente «microscópico» de un conglomerado de alrededor de diez mil millones de neuronas en el cerebro humano. Es en el «esquema de cableado» entre estos miles de millones de neuronas donde yacen los misterios de la inteligencia, de la conducta, de los sentimientos y las angustias de los seres humanos. En un sentido macroscópico, el desarrollo de este esquema de cableado está gobernado por el código genético de las especies. Sin embargo, en una parte importante del sistema nervioso, especialmente en la corteza cerebral, la distribución específica de los contactos sinápticos activos es el resultado de la acción repetida de ciertos modos de estimulación; en otras palabras, de la experiencia y del aprendizaje (véase Sec. 4.10).

Cuando se observa con un electrodo una neurona individual, perteneciente a la región nerviosa de algún órgano sensorial determinado, generalmente, se constata una correlación entre la frecuencia de disparo y la magnitud de ciertos parámetros físicos del estímulo externo original. Los disparos individuales no ocurren, habitualmente, a intervalos de tiempo iguales, separados regularmente. Lo que cuenta es el *hecho* de estar disparando o bien la *frecuencia promedio* (véanse, sin embargo, los razonamientos que se exponen más abajo). Una neurona puede disparar espontáneamente impulsos a frecuencias que llegan a varias decenas de Hz. Lo que constituye el mensaje nervioso de esa neurona es el *cambio* en la frecuencia de disparo espontáneo (incremento o inhibición). En general, un estímulo aplicado de manera constante producirá una frecuencia de descarga que va decreciendo gradualmente en función del tiempo, hasta quedar nivelada en cierto valor inferior al inicial. Este fenómeno se llama *adaptación*. Se ha observado que algunas neuronas responden solamente a cambios temporales de los estímulos sensoriales y que otras lo hacen frente a un conjunto de configuraciones espaciales o

patrones temporales muy complejos del estímulo (*detectoras de rasgos*). Como regla general, cuanto más avanzamos en la trayectoria nerviosa desde los receptores hasta la corteza cerebral, tanto más complejos y elaborados son los rasgos del estímulo a los cuales responde una neurona determinada (véase Fig. 2.26)

Estamos ahora en mejores condiciones para discutir cómo el sistema nervioso puede recoger y codificar información sobre patrones vibratorios acústicos. Cuando la señal acústica de un tono puro aislado de determinada frecuencia llega al oído, las oscilaciones de la membrana basilar estimulan las células pilosas que están en la zona de resonancia que corresponde a esa frecuencia (Sec. 2.3). En los seres humanos, hay una hilera de 4000 *células pilosas «internas»* (hilera que va a lo largo de la membrana basilar desde la base hasta el ápex, del lado del modiolo, hueso central de la cóclea), y tres hileras de un total aproximado de 12.000 *células pilosas «externas»* (ver Fig. 2.7[a] y [b]). Cuando las cilias de una de estas células se flexionan en una cierta dirección, se disparan impulsos eléctricos en las neuronas del ganglio espiral que están en contacto sináptico con la célula pilosa en cuestión (en caso de una sinapsis inhibitoria, la frecuencia de descarga espontánea de una neurona podría ser inhibida). Los axones de estas neuronas forman las fibras aferentes del nervio auditivo; sus potenciales de acción (pág. 62) llevan colectivamente información digitalmente codificada sobre el movimiento de la membrana basilar al sistema nervioso central.

Una característica importante es la disposición de las terminaciones de los nervios aferentes. Mientras que una única fibra nerviosa está habitualmente en contacto con una sola célula pilosa interna, recibiendo de ahí mensajes provenientes de una región extremadamente limitada de la membrana basilar, cada una de las fibras nerviosas aferentes que inervan las hileras externas hacen contacto con 10-50 células pilosas esparcidas por varios milímetros, siendo de este modo capaces de coleccionar información de una región de resonancia que corresponde a un dominio de frecuencia mucho más ancho. Las células pilosas internas responden a la *velocidad* del movimiento de la membrana basilar, porque la flexión de sus cilias es proporcional a la velocidad del fluido endolinfático de los alrededores (la fuerza sobre un obstáculo inmerso en un fluido viscoso es proporcional a la velocidad del flujo). Las células pilosas externas, contrariamente, envían señales según el *desplazamiento* (las fuerzas mecánicas interactivas dependen de la distorsión momentánea de la membrana); su respuesta se satura para altos niveles de intensidad. El hecho de que cerca del noventa y cinco por ciento de las fibras aferentes en el nervio acústico terminen sobre las células pilosas internas, con 10-50 de fibras individuales haciendo contacto sináptico con cada célula, confiere claramente a la hilera interna el rol de receptor sensorial principal. Por otra parte, las células pilosas externas también reciben terminaciones

de fibras *eferentes* que transmiten impulsos nerviosos que llegan desde el sistema nervioso central¹⁷; este hecho, más la notable *motilidad* de las células pilosas externas descubierta recientemente, coloca a estas últimas en un rol dual como efectores y receptores (esto será discutido en detalle en la Sec. 3.6). Mientras es claro que cada tipo de células pilosas asume un rol muy específico en la transducción de la señal acústica, ambos tipos deben trabajar juntos: un daño en las células pilosas externas perjudica severamente la audición aun cuando la hilera interna funcione perfectamente. Continuaremos tratando este tema en la Sec. 3.6; para una información más detallada y actualizada ver por ejemplo Gelfand (1990) y Zwicker y Fastl (1990).

Se ha observado, implantando microelectrodos en fibras nerviosas cocleares activadas acústicamente, que una fibra determinada tiene un umbral de disparo más bajo para aquella frecuencia acústica f que evoca una oscilación máxima en la posición x de la membrana basilar (Fig. 2.8) inervada por esa fibra. Esta frecuencia de máxima respuesta se llama *frecuencia característica* de la neurona (Kiang et al., 1965). Yendo ahora a la distribución real en el tiempo de los impulsos individuales, mediciones (Zwislocki y Sokolich, 1973) han mostrado que la máxima frecuencia de descarga está asociada con la máxima *velocidad* de la membrana basilar cuando ésta se mueve hacia la rampa timpánica; mientras que durante el movimiento en la dirección opuesta hacia la rampa vestibular, se produce inhibición de la frecuencia de descarga. Además, la *posición momentánea* de la membrana basilar tiene un efecto (menos pronunciado) excitatorio o inhibitorio, que depende de si la membrana está momentáneamente torcida hacia la rampa timpánica o hacia el otro lado, respectivamente. Ambos efectos se suman para determinar la respuesta total. La Fig. 2.22 muestra una distribución hipotética en el tiempo de impulsos nerviosos en una fibra nerviosa del oído interno conectada a la región de resonancia apropiada de la membrana basilar, en el caso que fuera excitada por un tono de baja frecuencia cuyo patrón vibratorio tiene forma trapezoidal (según Zwislocki y Sokolich, 1973).

El estudio atento de esta figura revela cómo la información sobre la frecuencia de repetición (en realidad, sobre el período de repetición) de la señal acústica original puede ser codificada en forma de «trenes» de impulsos nerviosos. La Fig. 2.22 correspondería a un caso ideal de baja frecuencia. En realidad, las frecuencias acústicas son habitualmente más altas que las frecuencias de descarga nerviosa, y la situación *real* es una donde los intervalos «on» y «off» no están tan claramente definidos, o delineados, por un lado a causa de su corta duración (comparada con el período

¹⁷ Una pequeña proporción de fibras eferentes también actúan sobre el 'output' de una célula ciliada interna, pero sólo indirectamente: ellas hacen sinapsis con las fibras aferentes que están en contacto con la célula ciliada interna.

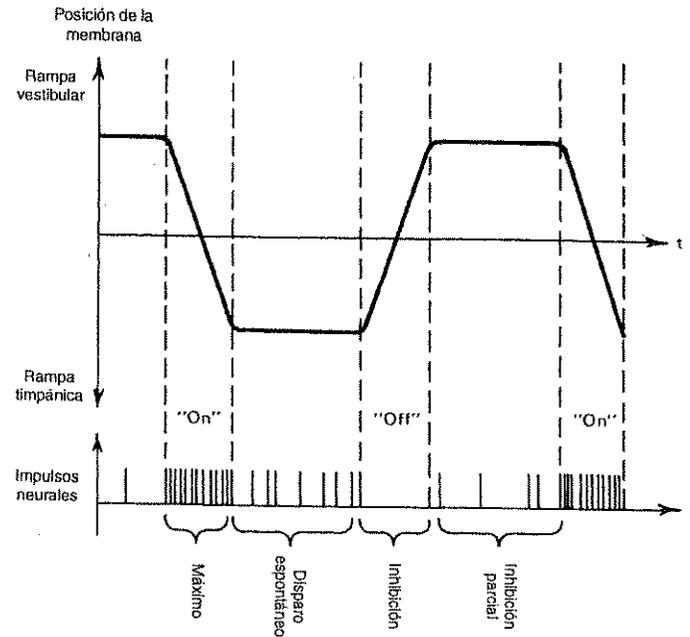


Figura 2.22 Bosquejo de la distribución de impulsos neurales en una fibra del nervio acústico durante distintas fases de un patrón trapezoidal de vibración de la membrana basilar.

refractario de una neurona típica) y por otro debido al carácter aleatorio de la distribución de los impulsos. La propiedad que realmente importa estadísticamente es que habrá más impulsos que caen en los intervalos «on» que en los intervalos «off».

Por consiguiente, para los tonos simples, *el intervalo de tiempo entre impulsos sucesivos tenderá a ser un múltiplo entero del período τ de la vibración sonora* (Kiang et al., 1965). Es obvio que cuanto más alta sea la frecuencia del tono, tanto menos definidos serán estos agrupamientos. Para frecuencias superiores a unos pocos miles de Hz esto ya no funciona más. Cuando varias fibras que reciben estímulos desde una estrecha región de la membrana basilar están agrupadas en un mismo paquete (como ocurre en el nervio auditivo), la suma de sus impulsos (como se detecta con un macroelectrodo que hace contacto con muchas fibras simultáneamente) aparecerá en sincronía con el estímulo auditivo. Estos grupos de señales neurales sincronizadas colectivamente han sido llamadas *volleys*.

2.9 Altura subjetiva y el rol del sistema nervioso central

Se desprende de la sección precedente que una fibra nerviosa dada, del nervio auditivo, puede transmitir dos tipos de información.

1. El simple hecho de que *esté disparando* (o que su frecuencia de descarga espontánea haya sido inhibida) le comunica al sistema auditivo central que la membrana basilar ha sido activada en o cerca de la región enervada por esa fibra; la distribución espacial (u «organización tonotópica») de las fibras que están disparando señales codifica la información primaria sobre la altura. Este proceso funciona en todo el rango de frecuencias.

2. La *distribución temporal* de los impulsos lleva información con respecto a la frecuencia de repetición o periodicidad y también sobre detalles más finos del patrón vibratorio mismo (véase más abajo). Esto sólo funciona en el rango de frecuencias más bajas.

No hay duda de que la información referida al *lugar* de excitación es usada por el sistema auditivo a todos los niveles. Pero ¿utiliza este sistema la información contenida en la distribución *temporal* de los impulsos nerviosos, tal como se muestra esquemáticamente en la Fig. 2.22?

Retornemos primero, por un momento, a la percepción de tonos simples (sinusoidales) aislados. Varios argumentos señalan el hecho de que la distribución en el tiempo de los impulsos nerviosos *no* es utilizada en la percepción de la altura de un tono *puro*. Por ejemplo, cálculos teóricos (Siebert, 1970) muestran que si la codificación de la altura primaria estuviera basada en rasgos temporales, la DAP de resolución de frecuencia (por ej., véase Fig. 2.9) sería independiente de la frecuencia (no lo es), y a su vez debería decrecer a medida que la amplitud del estímulo aumenta (cosa que no ocurre).

Que las claves temporales sean en gran parte ignoradas en la percepción de altura primaria de tonos puros no es gran sorpresa. ¿Pero qué decir acerca de la percepción de los batidos en consonancias desafinadas, y de la altura de periodicidad en complejos de armónicos? Es difícil encontrar una explicación de los batidos en consonancias desafinadas, y de otros efectos que demuestran sensibilidad a la fase, sin tener que suponer que en algún nivel exista un mecanismo que analice la estructura fina temporal del patrón vibratorio del estímulo. Efectivamente, podemos aludir al efecto mostrado en la Fig. 2.22 para intentar una explicación de cómo podría ser codificada la información sobre el patrón vibratorio y sus variaciones (batidos de segundo orden). Consideremos una superposición de dos tonos simples formando una octava. Supongamos que el patrón vibratorio resultante es el graficado en la parte inferior de la Fig. 2.16. Aparecerán

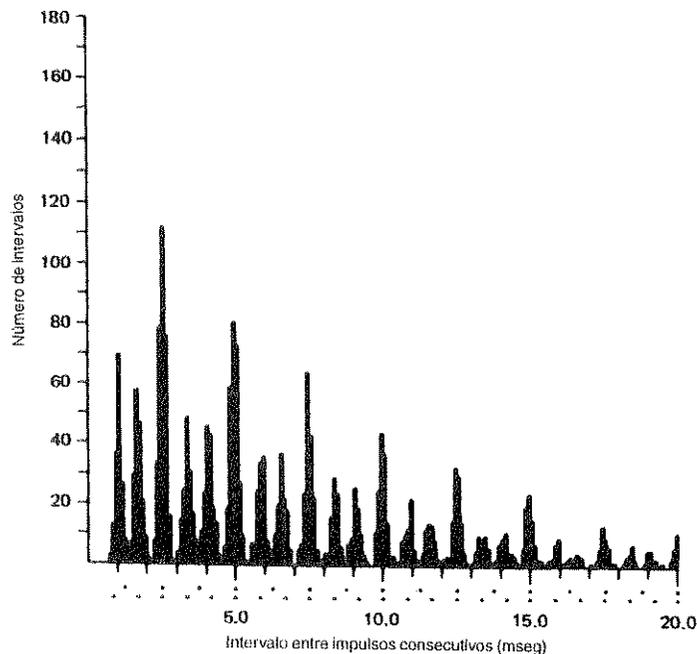


Figura 2.23 Histograma que muestra el número de veces (eje vertical) que ocurre un dado intervalo de tiempo entre impulsos de respuesta neural (eje horizontal), en una fibra estimulada por una superposición de dos tonos (quinta) en cierta relación de fase (Rose et al., 1969). Con permiso de los autores.

dos zonas de resonancia sobre la membrana basilar, centradas en las posiciones x_1 y x_2 , que corresponden a las frecuencias componentes f_1 y $f_2 = 2f_1$ (Fig. 2.8). En el nervio coclear tendremos dos focos principales de actividad, centrados en las fibras que tienen como frecuencias características f_1 y f_2 , que llevan a dos sensaciones de altura, separadas por una octava. Sin embargo, las zonas de resonancia sobre la membrana basilar son bastante anchas, produciéndose cierta superposición en la región entre x_1 y x_2 , donde los puntos de la membrana vibrarán de acuerdo con un patrón superpuesto que de algún modo está relacionado con el movimiento original del tímpano¹⁸. Las fibras conectadas a la región de superposición por lo tanto responderán con disparos que se agrupan en intervalos «on» y que corresponden, digamos, a las porciones descendentes (pendiente negativa) del segundo gráfico de la Fig. 2.16. Obsérvese que

¹⁸ Las ondas en el fluido coclear cambian su relación de fase y, desde luego, su amplitud a medida que se propagan, alterando de este modo la *real* forma del patrón vibratorio en diferentes puntos de la membrana basilar.

en este caso los intervalos «on» no son de igual duración, sino que responden a una secuencia que alterna «corto - más largo - corto - más largo». Si los dos tonos estuvieran separados por una quinta, el patrón vibratorio correspondiente a la zona de superposición podría ser el de la Fig. 2.18, que conduciría a una diferente secuencia de intervalos «on». La periodicidad de esta secuencia representaría así la información sobre la frecuencia de repetición, mientras que la estructura de la secuencia (suerte de «código Morse») daría información sobre detalles del patrón vibratorio. En efecto, tal estructura fina ha sido identificada estadísticamente a través de mediciones electrofisiológicas.

La Fig. 2.23 es un ejemplo (un histograma) de la distribución de los intervalos de tiempo entre impulsos nerviosos en una fibra nerviosa (Rose et al., 1969), para un estímulo que corresponde a un intervalo musical de quinta, con una relación de fase determinada.

Obsérvese la diferencia en el número relativo de veces (eje vertical) en que aparece un intervalo dado entre impulsos sucesivos (eje horizontal). Esto representa el código Morse (estadístico) que mencionamos arriba, y que lleva información sobre el patrón vibratorio. Cuanto mayor sea la complejidad del patrón vibratorio original y cuanto más alta sea la frecuencia de los tonos componentes, más «borroneada» será la información transmitida por la secuencia de impulsos, es decir, tanto más difícil de ser interpretada en los niveles superiores del cerebro. El detallado análisis de la distribución temporal de los impulsos nerviosos requeriría la acción, en algún nivel, de un mecanismo llamado de *autocorrelación temporal* (propuesto inicialmente por Licklider, 1959), en el cual un «tren» de impulsos es comparado con trenes de impulsos previos, los rasgos repetitivos similares son amplificados (tales como los períodos marcados en la Fig. 2.20[b]) y los otros (no periódicos) son suprimidos. Sin embargo, no existe hasta ahora evidencia anatómica de la existencia de este mecanismo en el sistema nervioso auditivo aferente.

Las claves temporales también son utilizadas por el mecanismo responsable de la sensación de direccionalidad o *localización del sonido*¹⁹ (por ej., Molino, 1974). Esta audición baural debería involucrar un proceso llamado *correlación cruzada temporal** de las señales nerviosas que llegan de ambas cócleas, en el cual se determina la diferencia temporal entre las mismas. Hay evidencias fisiológicas de que tal mecanismo existe (en el complejo olivar superior, Fig. 2.26). Licklider propuso un modelo del mecanismo nervioso de correlación cruzada (1959). En este modelo (Fig. 2.24) se supone que una neurona ascendente solo puede producir una

¹⁹ Las claves dadas por la intensidad (diferencia de amplitud entre las ondas sonoras que llegan a los dos oídos) y las dadas por el espectro (diferencia tímbrica) también contribuyen a la localización del sonido, especialmente a frecuencias elevadas y en entornos cerrados.

* En el original: temporal *crosscorrelation* (N. del T.).

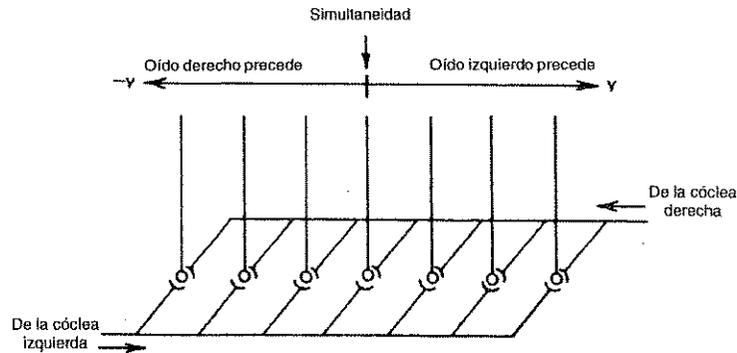


Figura 2.24 Modelo de un mecanismo de correlación cruzada (detector de diferencia temporal interaural). Según Licklider (1959).

descarga cuando es excitada simultáneamente por ambas fibras que traen señales del exterior. Dado que una señal nerviosa se propaga con una velocidad finita a lo largo de cada fibra, la llegada simultánea a una terminación de una neurona ascendente dada requerirá una cierta *diferencia temporal* entre las señales de ambas cócleas. Por ejemplo, una perfecta simultaneidad de las señales de ambas cócleas dispararía la neurona ascendente que está exactamente en el centro, porque ése es el lugar donde ambas señales, la del oído derecho y la del oído izquierdo, se encontrarían. Si, en cambio, la señal original fuera detectada primero en el *oído derecho*, su impulso viajará más allá del punto medio hasta encontrarse con el impulso demorado del oído izquierdo. Es fácil de ver que la ubicación y de la neurona ascendente activada dependerá del retraso temporal interaural, el que a su vez dependerá de la dirección de llegada del sonido. El esquema de la Fig. 2.24 es un modelo muy simplificado. Detallados estudios neuroanatómicos del complejo olivar superior (Fig. 2.26) revelan una estructura más complicada. En realidad, la información binaural es codificada a través de una compleja interacción entre señales entrantes excitatorias e inhibitorias que son el resultado de diferencias de tiempo (fase) y de intensidad entre los estímulos que alcanzan ambos oídos (Goldberg y Brown, 1969). Sea cual fuere el mecanismo real de correlación cruzada, su capacidad de resolución temporal es asombrosa: ¡Los seres humanos podemos localizar fuentes sonoras en el espacio en base a diferencias temporales interaurales menores de 20 microsegundos! Esto es solo una pequeña fracción de la duración de un potencial de acción. La característica inervación de las células pilosas internas, en la que cada una de éstas es contactada por más de 20 fibras nerviosas, puede ser responsable de la transmisión coherente de información nerviosa a la frecuencia requerida (Hudspeth, 1989).

Dos tonos binaurales, separados por un intervalo desafinado, pueden «engañar» al correlador cruzado: la diferencia de fase que cambia gradualmente entre los dos tonos (por ej., Fig. 2.16) será interpretada por este mecanismo como una diferencia cambiante en el *tiempo de llegada* entre la señal auditiva de la derecha y la de la izquierda, enviando de aquí mensajes al cerebro que llevan a una sensación de cambio cíclico en la dirección de llegada del sonido (cambio físicamente inexistente). Esta es la razón por la cual dos tonos puros que forman un intervalo consonante desafinado, presentado dicóticamente con auriculares, da la extraña sensación de una imagen sonora que parece estar «rotando dentro de la cabeza» (pág. 53).

Un asunto aún no resuelto es si el análisis de la secuencia temporal de impulsos neurales puede explicar de por sí la percepción de la altura de periodicidad (por ej., Yost y Watson, 1987). Un mecanismo de autocorrelación temporal, con su potencial capacidad de detectar la frecuencia de repetición de señales neurales, efectivamente podría explicar algunos importantes resultados relacionados con el seguimiento de la fundamental (pero *no todos*). Si no es un análisis de las claves temporales, ¿cuál es el mecanismo que nos permite asignar una sola altura a un tono musical compuesto de muchos armónicos, incluso cuando la fundamental no esté presente? ¿Por qué percibimos las alturas que corresponden a las frecuencias dadas por las relaciones (2.7a)-(2.7c) cuando se toca una melodía con los complejos de dos tonos armónicos correspondientes?

Algunas ideas que pueden conducir a una explicación de estos fenómenos (por ej., Terhardt, 1972; Wightman, 1973; Goldstein, 1973) son presentadas aquí en forma muy simplificada. Los sonidos «naturales» usados en la comunicación acústica entre seres humanos y animales contienen una importante proporción de tonos que son *armónicos* (vocales, canto de pájaros, llamadas de animales). Tales tonos tienen una propiedad en común - están formados por una superposición de armónicos, de frecuencias nf_1 , múltiplos enteros de una fundamental f_1 (pág. 57). Estos tonos producen un complicado patrón de resonancia sobre la membrana basilar, con múltiples picos de amplitud, uno para cada armónico (Fig. 2.25). A pesar de su complejidad, este patrón conlleva algunas

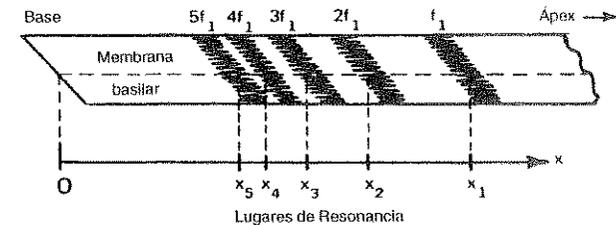


Figura 2.25 Bosquejo de las regiones de resonancia producidas en la membrana basilar por un tono compuesto.

características invariantes. Una de estas invariancias es la relación particular que existe entre las distancias de zonas vecinas de máxima resonancia²⁰. Para armónicos altos (n más grande que 7 u 8) esta relación pierde su definición física a causa de las superposiciones mutuas de las zonas de resonancia²¹. O bien aprendemos a temprana edad (Terhardt, 1972, 1974) o tenemos un mecanismo innato (Wightman, 1973; Goldstein, 1973), para reconocer esta característica invariante como perteneciente «a una y la misma cosa». Llamaremos a este mecanismo de reconocimiento *procesador central de altura*. La principal función de esta unidad neural es transformar el patrón de actividad periférica en otro patrón, de manera tal que todos los estímulos con la misma periodicidad sean representados de una manera similar. El resultado es una única sensación de altura, a pesar de los muchos armónicos presentes y de la consiguiente complejidad del patrón de excitación inicial. Esta sensación única de altura corresponde a la de la fundamental f_1 , que en los sonidos naturales es habitualmente la más destacada (en lo que a intensidad se refiere). Todo esto debería funcionar de manera muy análoga al reconocimiento de patrones visuales. Por ejemplo, al mirar usted el símbolo III , puede que éste no le transmita ningún significado particular o «único» (su interpretación probablemente dependerá de la orientación espacial del símbolo y del contexto en el que se muestra). Pero cualquiera que esté familiarizado con el alfabeto cirílico lo percibirá como siendo solo «una cosa» (la letra «shch»), no importa dónde en el campo visual o en qué orientación esté proyectado.

Se supone que dentro de nuestro sistema de procesamiento central hemos construido *moldes* [templates] básicos con los cuales comparamos las complejas estructuras de los patrones de excitación espacial de la membrana basilar. Cada vez que logramos establecer una correspondencia con uno de estos moldes, surge una sensación única de altura. Este proceso de correspondencia también funciona si solo una parte del patrón de excitación está presente. Si en lugar de un sonido natural complejo somos expuestos a uno en el cual algunos de los elementos normalmente esperados están ausentes (por ej., fundamental ausente), el patrón de excitación, al entrar en el mecanismo de reconocimiento del procesador de alturas, dentro de ciertas limitaciones puede ser apareado con un molde y dar una sensación de altura, a pesar de haber sido parcialmente truncado. Volvemos a encontrar muchas analogías en el terreno del reconocimiento

²⁰ Otra característica invariante es la elevada coherencia de las variaciones temporales macroscópicas de este complicado patrón de excitación sobre todo el dominio espacial de la membrana basilar.

²¹ A causa de la particular relación entre lugar de resonancia y frecuencia (Fig. 2.8), las regiones de resonancia se van amontonando cada vez más en la medida en que ascendemos la serie de los armónicos (a medida que n aumenta).

de patrones visuales. Obsérvese los contornos no existentes -pero esperados- en las siguientes letras:

Música

El proceso de correspondencia o ajuste de moldes arriba mencionado funciona incluso si los componentes armónicos de un tono son enviados alternadamente a oídos diferentes (por ej., Houtsma y Goldstein, 1972). Esto obviamente significa que el procesador central de alturas debe estar localizado en un nivel avanzado del procesamiento auditivo, después que la información procedente de ambas cócleas ha sido combinada. Aun más, el proceso de correspondencia funciona incluso en el caso de que se presenten solo dos armónicos vecinos de un tono compuesto, como vimos en la Sec. 2.8. En ese caso, sin embargo, el mecanismo de correspondencia puede cometer errores y ajustarse a una de *varias* posiciones aceptables (percepción de altura ambigua).

Hay tres teorías que predicen correctamente muchos de los resultados cuantitativos de mediciones psicoacústicas concernientes a la altura subjetiva. Una de ellas (Goldstein, 1973) supone que la información neural sobre las posiciones de máxima resonancia de la membrana basilar no está perfectamente definida, que fluctúa estadísticamente dentro de ciertos límites. Se espera que la correspondencia, o «ajuste de un molde», sea aquella que minimice las diferencias con la señal real dentro del marco de las fluctuaciones esperadas (una estimación de probabilidad máxima de frecuencia fundamental y orden de armónico - más sobre esto en la Sec. 4.8 y en el Apéndice II). Una teoría casi equivalente (Wightman, 1973) supone la operación de un mecanismo de autocorrelación en el dominio espacial (a diferencia de la autocorrelación temporal mencionada en la pág. 69). La distribución de la actividad espacial originada en un conjunto de fibras nerviosas que reciben la información procedente de ambas cócleas es enviada a una red neural en la cual se compara cuantitativamente la activación de una región espacial determinada con la existente en otras regiones, en cada instante. En este proceso, ciertos aspectos de la información entrante se convierten en un foco de actividad cuya localización espacial codifica una cualidad (es decir, la altura), relacionada con una característica invariante de la información entrante (las relaciones de distancia entre las zonas de máxima excitación sobre la membrana basilar), sin tener en cuenta otras posibles variables de la información de entrada (como ser las intensidades o las fases de los distintos armónicos). Debería señalarse que lo más probable es que este mecanismo de autocorrelación espacial también implique un proceso de aprendizaje. Efectivamente, el condicionamiento es necesario para lograr que el sistema responda a las configuraciones espaciales definidas por las

regiones de máxima actividad (Fig. 2.25) que aparecen con más frecuencia al escuchar sonidos naturales (armónicos).²²

Ninguna de estas dos teorías intenta describir *cómo* los algoritmos clave (ajuste a moldes [template matching] o autocorrelación espacial, respectivamente) son realmente ejecutados por el procesador de altura en el sistema nervioso. Sin embargo, en el cerebro existen redes neurales que son perfectamente capaces de ejecutar las operaciones de suma y multiplicación de impulsos neurales que se requieren para la ejecución de estos algoritmos. La teoría de Terhardt (1974) es la que más se acerca a proponer un esquema de conexión neuronal. En la práctica, sus computaciones están basadas en una *matriz que aprende*, un circuito analógico que «aprende a responder»²³ a los rasgos característicos de las configuraciones de entrada que ocurren más frecuentemente (es decir, a las relaciones de distancia entre zonas de máxima excitación causadas por un tono compuesto). Retornaremos a estas teorías más adelante, cuando discutamos explícitamente la percepción de tonos compuestos (Sec. 4.8 y Apéndice II) y de consonancia y disonancia (Sec. 5.2). (Ver también Terhardt et. al., 1982).

Finalmente, *no se puede excluir* la posibilidad de un uso, por lo menos parcial, de la distribución temporal de los impulsos nerviosos en la percepción de la altura de tonos compuestos.

¡Es difícil creer que el sistema nervioso, siempre orientado a trabajar con tan asombrosa eficiencia, con tantos sistemas de 'backup', no saque ventaja de la información dada por el «código Morse» (Fig. 2.22) que realmente existe en los canales periféricos de transmisión auditiva! Ya hemos mencionado que uno debería esperar que el mecanismo de altura basado en claves espaciales funcionara mejor para los armónicos más bajos (donde las zonas de resonancia vecinas están bien separadas entre sí), mientras que un mecanismo basado en lo temporal debería funcionar de manera óptima en el rango de armónicos más altos. Algunos experimentos psicoacústicos parecen exigir una explicación basada en el análisis de claves temporales. Por ejemplo, tonos simples de frecuencias bajas y de muy corta duración (dos o tres ciclos de vibración reales) pueden originar una clara sensación de altura (Moore, 1973). O, si una señal acústica (ruido blanco) se presenta a un oído, y la misma señal es enviada al otro oído con un intervalo de demora τ (unos pocos milisegundos), se percibe, muy débilmente, una *altura* correspondiente a una frecuencia $1/\tau$ (Bilsen y Goldstein, 1974). Ninguno de estos resultados puede explicarse

²² Es improbable que el sistema de «medición» requerido (cuasi logarítmico y dependiente de las dimensiones de la cóclea) venga dado genéticamente.

²³ Esto fue realizado originalmente en un modelo de laboratorio, disminuyendo apropiadamente las resistencias de las líneas de transmisión (filas y columnas de la matriz) que son simultáneamente activadas (que conducen corriente) de acuerdo con cierta configuración de entrada, presentada repetidamente.

satisfactoriamente con una teoría basada en el análisis de patrones espaciales.

Del análisis precedente surge claramente que la percepción de alturas subjetivas requiere la ejecución de operaciones de extracción de altura de «un orden más alto» en el sistema nervioso, *después* de haber sido combinada la información procedente de ambas cócleas. Por esa razón concluimos este capítulo con una descripción resumida de algunas de las características más importantes del trayecto de procesamiento auditivo (Whitfield, 1967; Brodal, 1969; Gelfand, 1990). Esto también servirá de referencia para discusiones en capítulos posteriores. La exploración anatómica de los trayectos nerviosos y de sus interconexiones es una tarea experimental extremadamente difícil. Las neuronas son células cuyas ramificaciones pueden tener muchos centímetros de largo; cada neurona, especialmente en el tejido cerebral, puede recibir información de miles de otras células, al tiempo que distribuye información a cientos de otras. Es casi imposible seguir microscópicamente el patrón de conexión en el caso de una célula. Solo pueden hacerse estimaciones globales a través del uso de diversas técnicas de tñido, inyección de sustancias trazadoras, degeneración axonal o siguiendo en el tejido el curso de patrones específicos de actividad eléctrica nerviosa.

La Fig. 2.26 muestra el trayecto nervioso auditivo desde la cóclea hasta el área de recepción auditiva en la corteza cerebral en forma de *diagrama de flujo*. Este muestra los canales de transmisión de información y las etapas de procesamiento y retransmisión; no guarda ninguna relación a escala con la estructura nerviosa real. El *ganglio espiral* es la red nerviosa que se encuentra en la cóclea, un primer estadio de procesamiento en este camino. Es aquí donde las neuronas que están en contacto con las células pilosas externas e internas tienen la primera oportunidad de interactuar, determinando la particular distribución espacio-temporal de actividad en el *nervio acústico* (el VIII nervio craneano), que transmite esta información al cerebro. La siguiente etapa de procesamiento, ubicada en el bulbo raquídeo, comprende los *núcleos cocleares*, y está compuesta por tres subdivisiones de cuya elaborada estructura dependen los primeros pasos en las tareas de resolución y discriminación sonora. La información es canalizada en tres haces principales. Uno cruza directamente hacia el lado *contralateral* y entra al *lemnisco lateral*, el canal principal a través del tronco cefálico. Algunas fibras terminan en la *formación reticular*, una red difusa en el tronco cefálico que tiene el rol de «consola telefónica» principal.²⁴ Otro paquete (el *cuerpo trapezoidal*) envía fibras desde el

²⁴ Esta estructura, que recibe datos de los sentidos y del organismo, así como también elaborada información desde la corteza, es responsable de la activación o inhibición de operaciones de procesamiento cerebral según las necesidades momentáneas. Además controla el sueño, el despertar y el estado consciente, e influye sobre muchas funciones viscerales.

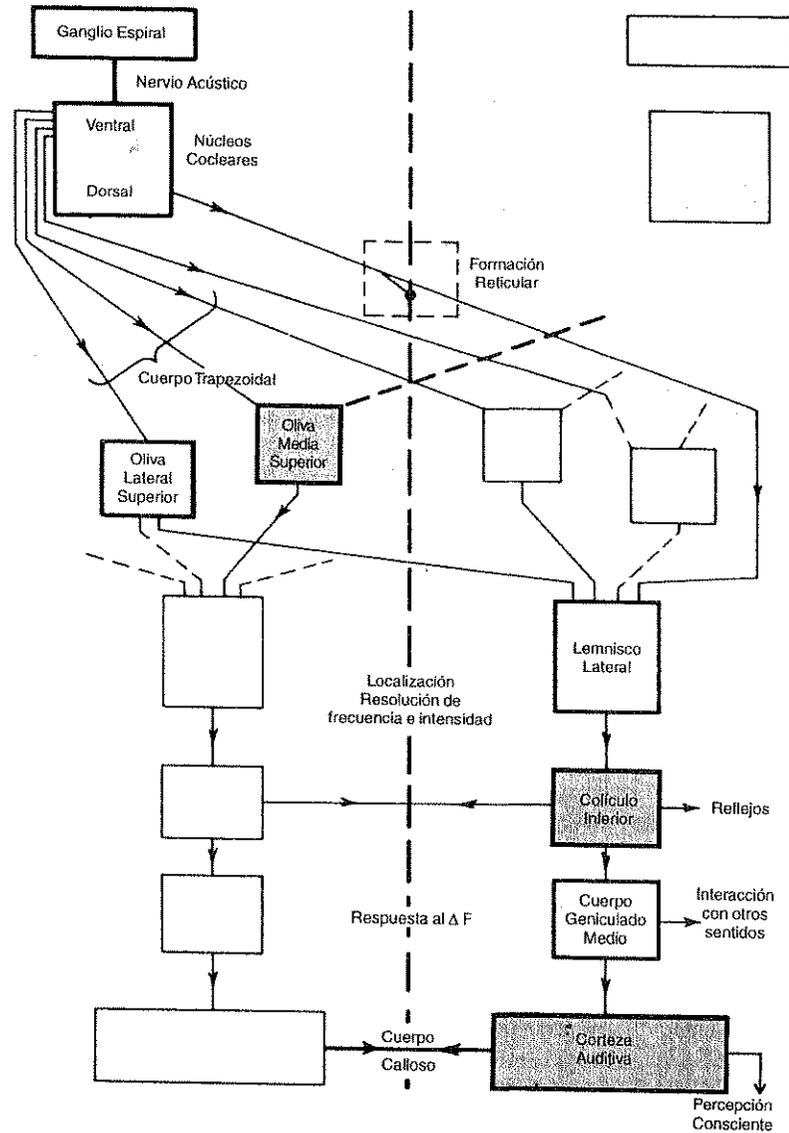


Figura 2.26 Diagrama de flujo de las señales neurales en las vías auditivas que parten desde el oído y llegan, a través del tronco cefálico, a las áreas corticales.

núcleo coclear ventral a importantes estaciones de procesamiento y retransmisión, las *olivas superiores lateral y media*. De éstas, la oliva superior media es el primer centro de mezcla de señales interaurales. Es en este lugar donde el mecanismo de correlación cruzada (Fig. 2.24) produce la información necesaria para la localización de la fuente sonora. Finalmente, un tercer paquete va del núcleo coclear ventral al complejo olivar contralateral.

Las tres etapas superiores están constituidas por el *colículo inferior*, el *cuerpo geniculado medio* y, finalmente, la *corteza auditiva* (Fig. 2.26). Algunas fibras están conectadas al *colículo superior*, el cual también está inervado por fibras relacionadas con la visión. Obsérvese en cada uno de estos estadios las interconexiones con el lado contralateral y con otras vías sensoriales y centros cerebrales.

Existe una red de *fibras eferentes*, no mostrada en la Fig. 2.26, que lleva información desde las etapas superiores hacia las inferiores, terminando en la cóclea. Este sistema interviene en el control de la información entrante aferente. La parte inferior de la red, el *paquete olivococlear* parece participar de manera importante en el proceso de agudizamiento (pág. 45). Aunque hay solo alrededor de 1600 fibras aferentes alcanzando cada cóclea, las fibras más anchas inervan profusamente las células pilosas externas, y de este modo pueden ejercer un control central sobre las operaciones eléctricas y/o mecánicas de las últimas (ver Sec. 3.6).

Por último, señalemos algunas generalidades que pueden ser útiles para capítulos posteriores. En la fase inicial existe una correspondencia geométrica muy específica entre las fibras nerviosas activadas y la posición del estímulo-fuente sobre la membrana basilar. Efectivamente, la distribución espacial de la excitación a lo largo de la membrana basilar proyecta un mapa definido de actividad neural en los haces de las fibras. Esto es particularmente evidente en cada uno de los núcleos cocleares. A medida que avanzamos en el trayecto neural, sin embargo, esta correspondencia «tonotópica» se va perdiendo gradualmente (excepto en un estado de anestesia). El número de neuronas participantes crece enormemente y la respuesta neural va representando cada vez más los rasgos complejos de la fuente sonora, siendo además crecientemente influenciada por la conducta y el estado mental del individuo. Los canales contralaterales (es decir los que cruzan) son «mejores» portadores de información que los ipsilaterales (del mismo lado); si se envía información contradictoria a ambos oídos, la información en el canal contralateral tiende a imponerse a la que llega por el canal ipsilateral (Milner et al., 1968).

A nivel del colículo inferior ya existe una buena resolución de frecuencia, de intensidad y de direccionalidad sonora; también hay una respuesta selectiva a barridos de frecuencia hacia arriba y hacia abajo. En

esta etapa, los reflejos ya funcionan, pero no hay evidencia alguna de percepción consciente del sonido, como ha sido demostrado con experimentos de ablación cerebral. En el cuerpo genicular medio (y probablemente en el colículo superior) ya existe alguna capacidad de reconocimiento de patrones sonoros. En este nivel existe información sobre *dónde está* una fuente sonora y *adónde está yendo*, en el espacio y en el tiempo. Aquí tiene lugar la primera integración con información de otros sentidos.

La última etapa de procesamiento de información entrante se ejecuta en la *corteza auditiva*. Aquí la función primaria es identificar el estímulo, integrarlo a la representación ambiental del momento y ponerlo a disposición del estado consciente del cerebro. Efectivamente, de aquí en adelante la información es distribuida a otros centros cerebrales, donde es almacenada, analizada e integrada a la totalidad de las funciones cognitivas del cerebro, o descartada por irrelevante. El *cuerpo calloso* (Fig. 2.26), una gigantesca comisura de alrededor de 200 millones de fibras que conectan a ambos hemisferios cerebrales, tiene un rol fundamental en la tarea de procesamiento de información global, especialmente si se tiene en cuenta la notable especialización de ambos hemisferios, ya mencionada en la Sec. 1.5. Retomaremos este tópico en la Sec. 5.7.

3 Ondas sonoras, energía acústica y percepción de la sonoridad

«Antes que nada debo cerciorarme de que el instrumento tenga buenos pulmones»

Broma usual de Johann Sebastian Bach (1685-1750)
antes de tirar todos los registros en la prueba
de un nuevo órgano.

En el capítulo precedente estudiamos *vibraciones* sonoras simples y sus efectos subjetivos, sin investigar cómo ellas llegan al oído. Nos hemos referido a experimentos en los cuales la fuente sonora (auriculares) estaba situada muy cerca del tímpano. En este capítulo discutiremos el proceso de *propagación* de energía sonora desde una fuente distante hasta el oyente y analizaremos cómo este flujo de energía acústica determina la sensación de sonoridad. Terminaremos el capítulo con una descripción más detallada de esa maravilla «electromecánica» que es la cóclea.

3.1 Ondas elásticas, fuerza, energía y potencia

Cuando el sonido se propaga a través de un medio, los puntos de éste vibran. Si no hay ningún sonido ni ningún tipo de perturbación, cada punto del medio¹ permanecerá en reposo. La posición en el espacio de un punto del medio en ausencia de perturbaciones, se llama *posición de equilibrio* de ese punto.

Las ondas sonoras son un caso particular de las así llamadas *ondas elásticas*. Cuando producimos una deformación repentina en un lugar dado del medio (por ejemplo al golpear una cuerda del piano con el martillo o al desplazar repentinamente el aire poniendo en movimiento la caña de un clarinete), aparecen fuerzas elásticas que harán que los puntos cercanos a la perturbación inicial empiecen a desplazarse. Estos puntos, a su vez, ejercerán fuerzas sobre otros puntos vecinos, transmitiéndoles la «orden» de ponerse en movimiento, y así sucesivamente. Esta «reacción en cadena» representa una onda elástica que se propaga desde el lugar de la

¹ Un «punto» del medio debe entenderse aquí en el sentido macroscópico, todavía englobando miles de millones de moléculas.

perturbación inicial. Lo que se propaga con esta onda no es materia sino *energía*: aquella energía necesaria para poner en movimiento cada punto alcanzado por la onda. Las ondas sonoras que interesan a la música son ondas elásticas en las cuales los puntos ejecutan movimientos periódicos. Al vibrar, cada punto del medio se mantiene muy, pero muy cerca de su posición de equilibrio. Una onda sonora se propaga a una velocidad perfectamente definida, alejándose de la fuente en línea recta, hasta que es absorbida o reflejada. La calidad acústica de un recinto, como ser una sala de conciertos, está determinada por el modo en que las ondas sonoras se propagan y son reflejadas y absorbidas.

Hemos mencionado los conceptos de fuerza y energía. Debemos ahora precisar el significado exacto que estos términos tienen en física. Todos tenemos una noción intuitiva de *fuerza*: la acción de tirar o empujar necesaria para cambiar la forma de un cuerpo, para ponerlo en movimiento, para sostener algo en nuestra mano, para ir frenando un móvil, etc. Pero en física no basta con conceptos intuitivos. Debemos dar una definición clara de fuerza, así como también la «receta» para medirla. Tanto la receta como la definición deben estar basadas en experimentos cuyos resultados puedan ser condensados o sintetizados en una ley física.

En nuestra experiencia cotidiana observamos que para cambiar la forma de un cuerpo tenemos que ejercer una acción bien específica: «aplicar una fuerza» sobre él. Cuando una fuerza actúa sobre un cuerpo, la deformación, es decir un cambio en la forma, no es el único efecto posible. En efecto, observamos que para alterar el movimiento de un objeto debemos aplicar una fuerza. Experiencias demuestran que, en general, la aceleración a de un cuerpo, que representa el cambio de su velocidad en función del tiempo, causada por una fuerza determinada F , es proporcional a esta última. Recíprocamente, la fuerza es proporcional a la aceleración producida: $F = ma$. Esto se conoce como la ecuación de Newton. La constante de proporcionalidad m es la *masa* del cuerpo. Esta representa su «inercia» o «resistencia» a cambiar su movimiento. Si más de una fuerza está actuando sobre un cuerpo, la aceleración resultante estará dada por la *suma* (vectorial) de todas las fuerzas. Esta suma puede ser cero; en tal caso, se dice que las fuerzas actuantes están en *equilibrio*.

La unidad de fuerza está representada por aquella fuerza necesaria para acelerar un cuerpo de 1 kg a razón de 1 m/s^2 . Esta unidad se llama Newton. Dado que la aceleración de la gravedad es de $9,8 \text{ m/s}^2$, el peso de un cuerpo cuya masa es 1 kg resulta ser 9,8 Newton.

Podemos calcular una fuerza midiendo la aceleración que ésta imparte a un cuerpo de masa determinada o equilibrándola (es decir, cancelando su efecto) con una fuerza conocida, por ejemplo la tensión de un resorte calibrado².

² «Calibrado» significa que previamente hemos establecido cuánto se alargará el resorte bajo una fuerza dada, por ejemplo, el peso de una masa determinada.

Hay muchas situaciones físicas en que una fuerza se aplica uniformemente sobre toda la superficie de un cuerpo. Por ejemplo, en un avión con cabina presurizada, volando a gran altura, el aire del interior ejerce una fuerza F considerable hacia afuera sobre cada ventana (y sobre toda otra parte del fuselaje), que es proporcional a la superficie S de la ventana. La relación $p = F/S$ representa la *presión del aire* dentro de la cabina. En general, definimos la presión del aire como la razón entre la fuerza que actúa sobre una superficie S que separa el aire del vacío. Si en lugar del vacío tenemos simplemente una presión diferente p' del otro lado de la superficie, la fuerza F que actúa sobre la superficie estará dada por

$$F = (p - p') S \quad (3.1)$$

Todo esto es muy importante para la música. Las ondas sonoras en el aire son *oscilaciones de la presión*. De este modo, si en la relación (3.1) S corresponde a la superficie del tímpano, p' es la presión constante en el oído medio y p la presión oscilante en el meato (Fig. 2.6), F será la fuerza oscilante que actúa sobre el tímpano, responsable de su movimiento y del de la cadena de huesos del oído medio. La presión se expresa en Newton por metro cuadrado, o Pascal. La presión atmosférica normal a nivel del mar es de alrededor de $100.000 \text{ Newton/m}^2$ (= 1000 hectopascal).

Pasemos ahora al concepto de *energía*. Como antes, tenemos una idea intuitiva al respecto, pero nuestra intuición puede engañarnos fácilmente en este caso. Por ejemplo algunos dirían que «se requiere mucha energía para sostener un objeto pesado durante un largo tiempo». Sin embargo, para el físico no hay ninguna energía involucrada en este caso (excepto durante el acto de levantar el objeto o de bajarlo). Por otra parte, para el fisiólogo, hay un flujo continuo de energía química hacia los músculos, para que las fibras de éstos puedan mantener el requerido estado de contracción continua. Para evitar confusiones es necesario introducir el concepto de energía de manera más precisa y cuantitativa.

El concepto de fuerza, por sí solo, no basta para la solución de problemas prácticos en física. Por ejemplo, si queremos determinar la velocidad final de un cuerpo acelerado por una fuerza dada, necesitamos saber durante cuánto tiempo, o a lo largo de qué distancia, esa fuerza ha estado actuando (hasta la fuerza más intensa puede tener solo un pequeño efecto *final*, si la duración o el camino de su acción fue muy corto). En efecto, lo que realmente cuenta para determinar un cambio dado de velocidad, digamos, entre 0 y v , es el producto de la *fuerza por la distancia recorrida* en la dirección de la fuerza. Si llamamos x a esa distancia, puede demostrarse matemáticamente, basándose en la ecuación de Newton, que $F \cdot x = \frac{1}{2} mv^2$. El producto $F \cdot x$ se llama *trabajo* y se considera positivo si el desplazamiento x tiene la misma dirección que la fuerza F . El producto $\frac{1}{2} mv^2$ se llama *energía cinética* del cuerpo de masa m . Si $F \cdot x$ es positivo,

la relación nos dice que el trabajo de la fuerza aplicada ha incrementado la energía cinética del cuerpo o, lo que es equivalente, que «trabajo ha sido entregado al sistema», incrementando su energía cinética de cero a $\frac{1}{2}mv^2$.

El trabajo y la energía cinética se expresan en Newton por metro. Esta unidad se llama Joule (J), en honor a un físico e ingeniero británico. Un cuerpo de 1 kg (unidad de masa), moviéndose a una velocidad de 1 m/s tiene una energía cinética de 0,5 J. Si su velocidad es el doble, su energía cinética será cuatro veces mayor: 2 J. Una persona promedio (de 70 kg), corriendo a una velocidad de 3 m/s (10,8 km/h) tiene una energía cinética de 315 J; la energía cinética de un auto de 2.000 kg viajando a 30 m/s (108 km/h) es de 900.000 J.

La energía cinética no es la única forma de energía. Consideremos un cuerpo sujeto a un resorte. Para comprimir el resorte debemos suministrar una cantidad determinada de trabajo. Si hacemos esto de manera muy, muy lenta, podremos despreciar la energía cinética involucrada. El trabajo suministrado será convertido en *energía potencial*; en este caso, energía potencial *elástica* del cuerpo ligado al resorte comprimido. Al soltar el resorte, el cuerpo será acelerado por la fuerza del resorte en expansión y la energía potencial será convertida en energía cinética. Podemos decir que la energía potencial es *energía de posición* de un cuerpo, mientras que la energía cinética es su *energía de movimiento*.

La suma de la *energía potencial más la energía cinética* de un cuerpo representa su *energía mecánica* total (hay muchas otras formas de energía que no consideraremos: térmica, química, electromagnética, etc.). Existen casos importantes en los cuales la energía mecánica de un cuerpo se mantiene constante. Uno «musicalmente» significativo es el ejemplo previo de un cuerpo unido a un resorte, oscilando por la acción de la fuerza elástica de este último. Puede demostrarse que la vibración resultante alrededor de la posición de equilibrio es *armónica* (siempre y cuando la amplitud sea pequeña). Cuando el cuerpo se suelta desde una posición de resorte extendido, su energía cinética inicial es cero. El cuerpo sin embargo posee una energía elástica potencial que, a medida que la oscilación comienza, se va convirtiendo en energía cinética. Cada vez que el cuerpo pasa por su posición de equilibrio, su energía elástica potencial es cero en ese instante, mientras que su energía cinética es máxima. Durante la oscilación armónica hay una conversión periódica de energía potencial en cinética y viceversa.

La energía mecánica total se mantiene constante mientras no entren en juego otras fuerzas «disipadoras». La fricción produce una disminución continua de la energía total, y por ende, de la amplitud de la oscilación. El movimiento resultante se llama *oscilación amortiguada*. Esta es extremadamente importante en música. Efectivamente, muchos instrumentos musicales involucran oscilaciones amortiguadas; la cuerda vibrante de un piano es un ejemplo típico. Se dan otros casos en que

fuerzas externas pueden actuar de manera tal que aumenten gradualmente la energía mecánica, pudiendo usarse para compensar las pérdidas disipativas y mantener de este modo una oscilación de amplitud constante. Un ejemplo es el de una cuerda de violín frotada: las fuerzas que aparecen en el mecanismo de frotamiento suministran energía a la cuerda vibrante en un monto equivalente a la pérdida de energía causada por fricción interna y radiación acústica (Sec. 4.2).

Ahora llegamos a un último punto, pero de máxima importancia, concerniente a la energía. Las máquinas (y los seres humanos) suministran energía a una *razón determinada*. Cualquier máquina (o ser humano) es capaz de ejecutar una cantidad casi arbitrariamente grande de trabajo, ¡pero puede llevarle un tiempo muy largo! Lo que realmente define la cualidad o potencia de una máquina es la «rapidez» o *razón* a la cual puede entregar energía (o sea, ejecutar trabajo). Esta razón, en el caso de que sea constante, está dada por:

$$P = \frac{\text{trabajo realizado}}{\text{tiempo empleado}} = \frac{W}{(t_2 - t_1)} \quad (3.2)$$

W es el trabajo entregado entre los tiempos t_1 y t_2 . P se llama la *potencia* mecánica. Se mide en unidades de J/s, llamadas Watt* (según otro ingeniero británico). Al subir una escalera, el cuerpo humano está suministrando una potencia de aproximadamente 300 W; la energía eléctrica consumida por segundo por una plancha eléctrica es de 1.000 W, y la potencia máxima entregada por un motor de auto pequeño es de 30 kW (1 caballo de fuerza = 0,735 kW). Un trombón tocando 'fortissimo' emite una potencia acústica total de alrededor de 6 W.

El concepto de potencia es de máxima importancia para la física de la música. En realidad, nuestro oído no se interesa en absoluto por la energía acústica total que llega al tímpano; a lo que responde es la *razón* con que esta energía llega por unidad de tiempo, es decir, la *potencia* acústica. Ella es la que determina la sensación de *sonoridad*.

3.2 Velocidad de propagación, longitud de onda y potencia acústica

Después de nuestra incursión en el campo de la física pura en la sección precedente, estamos en mejores condiciones de comprender el fenómeno de propagación de ondas. Con ese fin usaremos un *modelo* del medio de propagación formado por pequeños cuerpos de masa determinada, ligados entre sí por resortes comprimidos (que representan las fuerzas elásticas

* En castellano, *watts* (N del T).

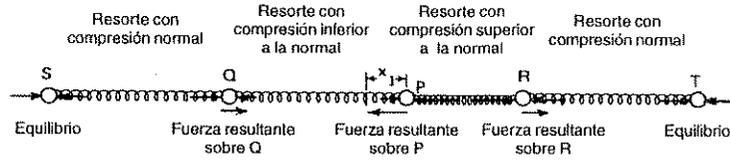


Figura 3.1 Modelo unidimensional de un medio elástico (resortes en estado de compresión), en el cual el punto P ha sido desplazado longitudinalmente.

del aire). Inicialmente, las fuerzas del resorte están en equilibrio y todos los puntos en reposo. La Fig. 3.1 muestra la situación cuando el punto P ha sido repentinamente desplazado una cantidad x_1 hacia la derecha.

Considerando las fuerzas mostradas en la Fig. 3.1, advertimos que los puntos Q y R - que inicialmente estaban en reposo en sus respectivas posiciones de equilibrio - se hallan sometidos a la acción de una fuerza resultante que actúa hacia la derecha. En otras palabras, de acuerdo con la ecuación de Newton, esos puntos serán acelerados hacia la derecha y empezarán a moverse en la misma dirección en la cual el punto P había sido desplazado originalmente. Este punto P , por otra parte, estará moviéndose hacia su posición de equilibrio, acelerado por un fuerza resultante que actúa hacia la izquierda (Fig. 3.1). Un momento más tarde, cuando los puntos Q y R estén en camino hacia la derecha, la compresión del resorte entre R y T empezará a aumentar, mientras que aquella entre Q y S comenzará a disminuir. Es fácil apreciar que los puntos S y T serán sometidos a una fuerza neta dirigida hacia la derecha, que hará que empiecen a desplazarse en esa dirección, mientras que los puntos Q y R pueden estar moviéndose hacia la izquierda, de vuelta hacia su posición de equilibrio. Este proceso continúa, de punto a punto, representando una onda que se propaga a partir del punto P en ambas direcciones. El «frente» de la onda no es otra cosa que una «orden» que pasa de un punto a otro diciéndole a este último: «empieza a moverte hacia la derecha». Esta orden está dada por los resortes comprimidos (por sus fuerzas elásticas). Como todos los puntos vuelven a sus posiciones de equilibrio, en ningún momento hay un transporte real de materia. El caso examinado representa una *onda longitudinal*, porque los desplazamientos de los puntos ocurren en dirección paralela a la dirección de propagación de las ondas. En el caso real de una onda sonora que se propaga en el aire, la acción concertada de las fuerzas de los resortes que actúan sobre los puntos P, Q, R, \dots corresponde a la presión del aire; las variaciones de estas fuerzas (por las variaciones de las distancias entre los puntos) corresponden a las variaciones de la presión del aire en la onda sonora.

El modelo unidimensional de la Fig. 3.1 muestra cómo una onda elástica también representa transporte de energía. En primer lugar, debemos

entregar trabajo desde «afuera» para producir el desplazamiento inicial x_1 del punto P , dado que tenemos que modificar las longitudes de los dos resortes PQ y PR . En otras palabras, necesitamos una fuente de energía. En este caso, la energía inicial se presenta bajo la forma de energía potencial (posicional) del punto P . Luego, a medida que transcurre el tiempo, los puntos a la derecha y a la izquierda de P empiezan a moverse, y la longitud de sus resortes a cambiar. Todos estos procesos involucran energía, tanto cinética (movimiento de los puntos materiales) como potencial (compresión o expansión de los resortes). La energía entregada inicialmente a P es transferida de punto a punto del medio a medida que la onda se propaga: tenemos un *flujo* o transporte de energía que parte desde la fuente.

Examinemos ahora un caso en el cual los resortes del modelo se encuentran bajo tensión (expandidos) en lugar de estar comprimidos, con puntos vecinos *bajo tracción* entre sí. Físicamente, esto corresponde a una cuerda de violín tensa. Para los desplazamientos longitudinales (en dirección a lo largo de los resortes) obtenemos una situación cualitativamente similar a la anterior; solo que ahora todas las fuerzas indicadas en la Fig. 3.1 aparecen invertidas. Pero además tenemos una nueva posibilidad, que no existe para el caso de resortes comprimidos: podemos desplazar el punto P *perpendicularmente* a la dirección de x (Fig. 3.2) y obtener así otro tipo de onda. Dado que todas las fuerzas de los resortes ahora tiran de los puntos, de acuerdo con la Fig. 3.2, la fuerza resultante F_P acelerará a P en dirección descendente, hacia su posición de equilibrio O . Los puntos Q y R , a su vez, estarán sometidos a fuerzas netas que los acelerarán hacia arriba, en una dirección esencialmente perpendicular a x . Esto representa una onda elástica *transversal*, que se propaga hacia la derecha e izquierda de P . En una onda transversal, los desplazamientos de los puntos son perpendiculares a la dirección de la propagación. En un medio bajo tensión, como una cuerda de violín, pueden ocurrir simultáneamente los dos modos de propagación de una onda elástica: transversal y longitudinal.

Examinemos ahora la expresión para calcular la velocidad de propagación de ondas transversales. Aplicando la ley de Newton a los

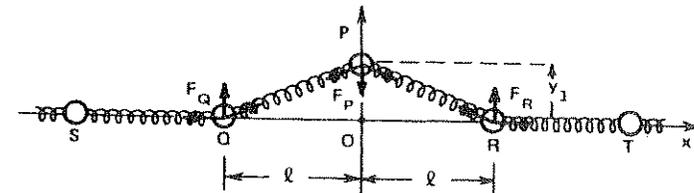


Figura 3.2 Modelo unidimensional de un medio elástico (resortes en estado de extensión), en el cual el punto P ha sido desplazado transversalmente.

43

puntos individuales del modelo unidimensional de la Fig. 3.2 puede demostrarse que, para una cuerda bajo tensión T (expresada en Newton), la velocidad V_T de ondas elásticas transversales está dada por

$$V_T = \sqrt{\frac{T}{d}} \quad (\text{m/s}) \quad (3.3)$$

d es la «densidad lineal» del medio, es decir, la masa por unidad de longitud (kg/m). Obsérvese que cuanto más tensa esté la cuerda, tanto más rápidamente viajarán las ondas transversales. Por otra parte, cuanto más densa sea, tanto más lentamente se propagarán las ondas.

Una relación físicamente equivalente existe para el cálculo de la velocidad de propagación de ondas longitudinales en un medio de densidad δ (en kg/m³) y en el que la presión es p (en Newton/m²):

$$V_L = \sqrt{\frac{p}{\delta}} \quad (\text{m/s}) \quad (3.4)$$

Para un gas ideal, la razón p/δ resulta proporcional a la temperatura absoluta t_A , definida en términos de temperatura t_C en grados centígrados por

$$t_A = 273 + t_C \quad (\text{grados Kelvin}) \quad (3.5)$$

Adviértase que en el punto de congelamiento ($t_C = 0^\circ\text{C}$), la temperatura absoluta es $t_A = 273^\circ$. Aunque el aire ordinario no es un «gas ideal» cien por cien, prácticamente se comporta como si lo fuera, y la velocidad de las ondas sonoras puede expresarse en la forma:

$$V_L = 20,1 \sqrt{t_A} \quad (\text{m/s}) \quad (3.6)$$

Esto resulta ser 331,5 m/s a 0°C y 344 m/s a 21°C . El factor numérico en la ecuación (3.6) vale solamente para el aire. En general, depende de la composición del medio a través del cual se propaga el sonido. Para el hidrógeno puro, por ejemplo, es 74. Por lo tanto, las ondas sonoras se propagan con una velocidad casi cuatro veces mayor en hidrógeno que en el aire. Esto lleva a efectos acústicos cómicos cuando una persona habla o canta después de haber inhalado hidrógeno (¡NO FUMAR!).

El sonido se propaga «rápidamente», pero no con velocidad infinita. Esto, por ejemplo, lleva a pequeñas pero perceptibles diferencias en el tiempo de llegada entre las ondas sonoras provenientes de diferentes instrumentos de una gran orquesta, que pueden causar serios problemas de sincronización rítmica. Un pianista que toca por primera vez un gran órgano, en el cual la consola está lejos de los tubos, puede sentirse confundido al comienzo por la tardía llegada del sonido, desincronizada con la acción de sus dedos. La reverberación de una sala se basa en la

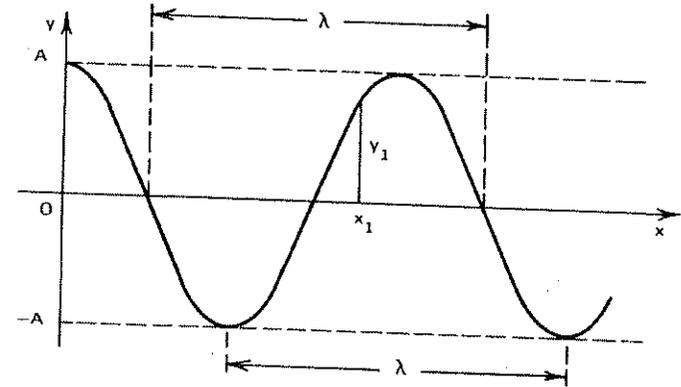


Figura 3.3 «Foto instantánea» de los desplazamientos y de una cuerda, cuando una onda transversal de una sola frecuencia se propaga a lo largo de ésta en dirección x .

superposición de ondas sonoras retrasadas, que han sufrido múltiples reflexiones sobre las paredes (Sec. 4.7).

Consideremos ahora una cuerda muy larga, cuyo extremo izquierdo empieza a vibrar con un movimiento armónico simple por la acción de alguna fuerza externa. Después de un rato, todos los puntos de la cuerda estarán vibrando con el mismo movimiento armónico simple. Si en un instante dado un punto cualquiera está, digamos, en la posición de su máximo desplazamiento, los puntos vecinos a su derecha todavía no llegaron a es extremo, mientras aquellos a su izquierda acaban de pasar por ahí. La figura 3.3 muestra los desplazamientos transversales de todos los puntos de la cuerda en un instante dado. Esta curva representa una «foto instantánea» de la forma de la cuerda durante el pasaje de una onda transversal sinusoidal. El gráfico de la Fig. 3.3 no debe confundirse con la curva mostrada en la Fig. 2.4, la cual representa la historia de un único punto dado. Este último gráfico representa un patrón vibratorio en el tiempo, mientras que el primero es una representación en el espacio. La distancia mínima entre dos puntos de la cuerda que estén vibrando en forma paralela (vibrando en fase, es decir, teniendo en todo momento idénticos desplazamientos y), se llama longitud de onda. Habitualmente se la indica con la letra griega λ . Otro modo de definir la longitud de onda es considerándola como la distancia mínima después de la cual el patrón espacial de vibración de la cuerda se repite. Compárese esta definición con aquella del período, el cual representa el intervalo temporal mínimo, después del cual el patrón vibratorio de un punto dado se repite (Fig. 2.3[b]).

A medida que el tiempo transcurre, la curva de la «instantánea» se ve mover hacia la derecha con la velocidad de la onda (Fig. 3.4); a pesar de

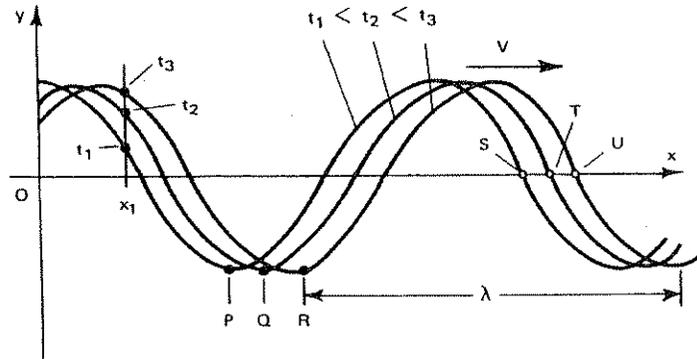


Figura 3.4 Tres «instantáneas» consecutivas de una onda transversal tomadas a los tiempos t_1 , t_2 y t_3 . Los puntos de la cuerda solo se mueven para arriba y para abajo (en dirección y). Lo que se mueve hacia la derecha, a lo largo de x , es el perfil de la onda (y la energía de la misma).

esto, cada punto de la cuerda se mueve en realidad solo para arriba y para abajo (por ejemplo, considérese el punto x_1 en el Fig. 3.4). Lo que se mueve hacia la derecha es la configuración, es decir, la real forma de la cuerda, pero no la cuerda misma. En otras palabras, lo que se mueve hacia la derecha es una cualidad, tal como la de «estar en la posición de máximo desplazamiento» (por ej. los puntos P, Q, R en la Fig. 3.4), o la cualidad de «pasar justo por la posición de equilibrio» (puntos S, T, U). Y, desde luego, lo que también se mueve hacia la derecha es *energía*, la energía cinética y potencial involucrada en la oscilación transversal de los puntos de la cuerda.

Existe una importante relación entre la velocidad de propagación V de una onda sinusoidal, su longitud de onda λ , y la frecuencia f de oscilación de los puntos individuales. Al observar la Fig. 3.3, nos damos cuenta que la onda se habrá propagado exactamente una longitud de onda λ durante el tiempo que le lleva al punto del extremo izquierdo (o a cualquier otro punto) ejecutar una oscilación completa, o sea durante un período τ . Podemos, por ello, escribir la siguiente relación para la velocidad de la onda:

$$V = \frac{\text{distancia recorrida}}{\text{tiempo empleado}} = \frac{\lambda}{\tau}$$

Dado que la inversa del período es la frecuencia f (relación [2.1]), también podemos escribir

$$V = \lambda \cdot f \tag{3.7}$$

Esta expresión relaciona cuantitativamente la «representación espacial» de la Fig. 3.3 con la «representación temporal» de la Fig. 2.4. La relación

(3.7) nos permite expresar la longitud de onda de una onda transversal en una cuerda en función de la frecuencia de oscilación de los puntos individuales y de la velocidad de propagación (3.3):

$$\lambda = \frac{1}{f} \sqrt{\frac{T}{d}} \tag{3.8}$$

Es interesante observar que las relaciones (3.3) y (3.8) pueden ser aplicadas también, hasta cierto punto, a la membrana basilar, reemplazando la tensión T con un parámetro apropiado de rigidez. Dado que la rigidez decrece un factor de aproximadamente 10.000 desde la base hacia el ápex (pág. 33), según las ecuaciones (3.3) y (3.8) la velocidad de propagación y longitud de onda *locales* de las ondas de una dada frecuencia en la membrana basilar decrecerán en un factor 100 la medida que viajan hacia el ápex. La frecuencia de resonancia de la membrana basilar, por otra parte, también es proporcional a la raíz cuadrada del parámetro de rigidez. Como consecuencia de todo esto, cuando la onda se propaga, su amplitud crece (la energía se «apila» porque la onda va propagándose con velocidad decreciente). Cuando llega a la zona de resonancia, tanto la amplitud como la disipación de energía alcanzan un máximo, lo que causa que la onda decaiga rápidamente más allá de ese punto. La figura 3.5 muestra esquemáticamente cómo la onda producida por un sonido simple, de una única frecuencia, se propaga a lo largo de la membrana basilar.

Podemos anticipar que cuando dos o más tonos simples llegan al oído (como ocurre con los sonidos reales de la música), aparecerán varios

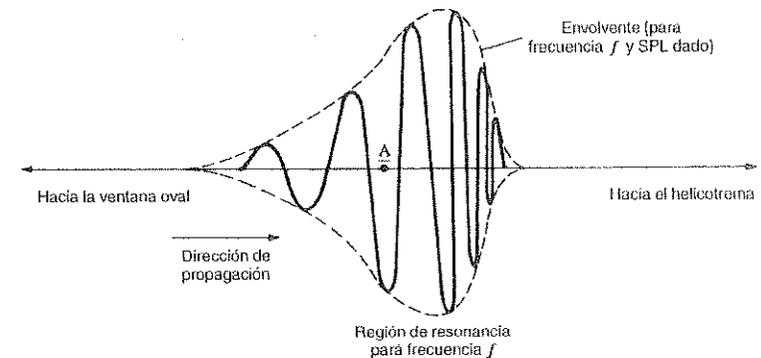


Figura 3.5 Bosquejo de una onda propagándose a lo largo de la membrana basilar, generada por un tono simple. Curva gruesa: instantánea de los desplazamientos transversales de la membrana (¡no en escala!). Imagínese esta curva avanzando hacia la derecha dentro de las líneas punteadas y reduciendo su velocidad a medida que su amplitud decae hacia la derecha. Curva punteada: envoltorio de amplitud (que permanece fija, salvo que cambie la amplitud y/o la frecuencia del tono).

45

«paquetes de ondas» como el mostrado en la figura, centrados en distintas zonas de resonancia, una para cada frecuencia componente (véase Fig. 2.25). Así es, en términos extremadamente simplificados, cómo trabaja el mecanismo hidromecánico de análisis de frecuencia en la cóclea. Recuérdese que la Fig. 3.5 representa una onda transversal: los puntos individuales vibran verticalmente, pero el patrón de la onda (y la energía asociada) se propaga de izquierda a derecha, con la amplitud de las oscilaciones individuales manteniéndose dentro de la «envolvente» de la onda. Obsérvese cómo la longitud de onda decrece a medida que la onda avanza en dirección a la zona de máxima resonancia. Todos los puntos de la membrana basilar, incluso aquellos que están fuera de la zona de resonancia máxima, oscilan con la misma frecuencia que la del tono puro original. Si cambiamos la frecuencia del estímulo tonal hacia arriba o abajo, el perfil mostrado en la Fig. 3.5 se desplazará hacia la base (izquierda) o hacia el ápex (derecha), respectivamente.³

Volvamos al caso de ondas longitudinales, tal como una onda sonora en el aire. Los puntos del medio vibran en sentido paralelo a la dirección de la propagación, y no es fácil representar gráficamente sus posiciones reales. Por esta razón es más conveniente representar las ondas sonoras como oscilaciones de presión. La figura 3.6 (la hilera inferior) muestra los desplazamientos de los puntos en un modelo de medio de propagación unidimensional, cuando una onda longitudinal pasa a través de él.

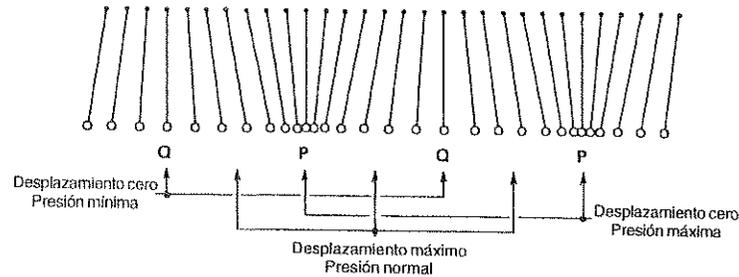


Figura 3.6 Onda longitudinal en un medio uni-dimensional. Para graficar los desplazamientos reales, cada punto se muestra como el menearse de un péndulo.

³ Así es como surge la «frecuencia característica» de una fibra nerviosa en el nervio acústico (pág. 65); considérese una neurona conectada con las células ciliadas ubicadas en la posición A de la Fig. 3.5. Su respuesta estará dada por la amplitud local de las oscilaciones de la membrana basilar. A medida que la frecuencia del tono de prueba ascienda gradualmente desde un valor muy bajo, la totalidad del patrón mostrado en la Fig. 3.5 se desplazará desde la extrema derecha al extremo izquierdo: cuando la envolvente de la oscilación pase por el punto A, la respuesta de aquella neurona incrementará gradualmente hasta un máximo (precisamente, la frecuencia característica) y luego decaerá bruscamente cuando el patrón se aleje hacia la izquierda.

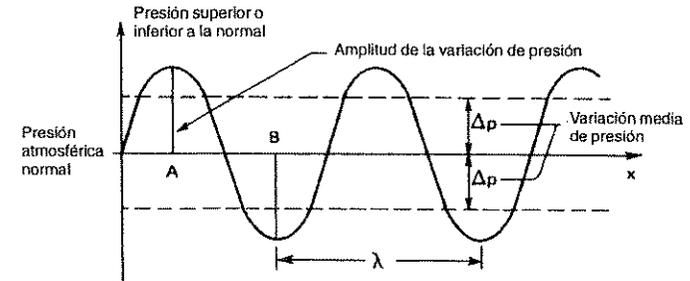


Figura 3.7 Variaciones de presión para un instante dado t en una onda sinusoidal que se propaga a lo largo del eje x .

Obsérvese que los puntos exhiben su máxima acumulación (es decir, máxima presión) y su máxima rarefacción (o sea, mínima presión) en los lugares donde sus desplazamientos son cero (puntos P y Q, respectivamente). Por otra parte, en los lugares donde sus desplazamientos son máximos, las variaciones de presión son cero. Esto significa que las variaciones de presión de una onda sonora están 90° fuera de fase con relación a la oscilación de los puntos: las variaciones máximas de presión (positivas o negativas) ocurren en lugares donde los desplazamientos de los puntos del medio son nulos; recíprocamente, los desplazamientos máximos de los puntos ocurren donde las variaciones de presión pasan por cero.

En una onda sonora sinusoidal la presión en cada punto oscila armónicamente alrededor de un valor normal (la presión atmosférica normal en ausencia de perturbación) (Fig. 3.7). En un lugar como el A, todos los puntos del medio se han acercado lo más posible entre sí (máximo incremento de la presión, puntos P en la Fig. 3.6); en un punto como B ellos se han separado entre sí al máximo (máxima disminución de la presión, puntos Q en la Fig. 3.6). La variación media de presión Δp es igual a la amplitud de la variación de presión dividida por $\sqrt{2}$ ($\approx 1,41$). Tomando en cuenta las relaciones (3.6) y (3.7) obtenemos para la longitud de onda de ondas sonoras sinusoidales en el aire:

$$\lambda = \frac{20,1}{f} \sqrt{t_A} \quad (\text{en metros})$$

t_A es la temperatura absoluta dada por la ecuación (3.5). En la Fig. 3.8 se muestran valores típicos de longitud de onda a temperatura normal.

Las ondas elásticas pueden transmitirse de un medio a otro: por ejemplo, del aire al agua, o del aire a una pared y luego de vuelta al aire, o de una cuerda a una caja de madera y de ahí al aire circundante. La naturaleza de la onda puede cambiar en cada transición (por ejemplo, la transición de una onda transversal en la cuerda y en la caja a una onda sonora longitudinal en el aire). Sin embargo, en cada transición la frecuencia

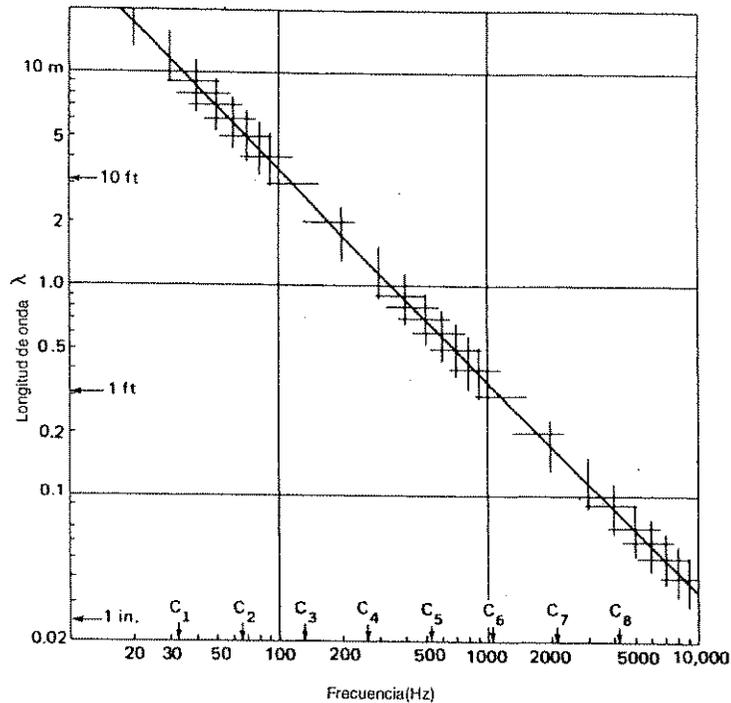


Figura 3.8 Longitud de onda de una onda sonora en el aire, a temperatura normal, como función de la frecuencia (escala logarítmica).

sigue invariante. La longitud de onda, por otra parte, cambiará según la relación (3.7): $\lambda = v / f$. En esta relación, v cambia de medio a medio, mientras que f estará determinada exclusivamente por la vibración original (por la fuente de las oscilaciones).

Cuando una onda elástica arriba al límite entre dos medios, parte de la onda es *reflejada* hacia el medio original. Algunos límites son reflectores casi perfectos (por ej., paredes lisas de cemento para las ondas sonoras en aire, los extremos fijos de una cuerda bajo tensión para las ondas transversales). Este fenómeno ocurre por el hecho de que en el límite reflector los puntos del medio están obligados a permanecer en reposo por fuerzas externas que cancelan el efecto de las fuerzas elásticas que «comandan» la propagación de la onda. En una reflexión la frecuencia se mantiene invariable, mientras que la dirección de propagación es invertida en el caso de una incidencia perpendicular (o, en general, dirigida en un ángulo de reflexión igual al ángulo de incidencia). También la amplitud seguiría invariable en el caso de que no hubiera absorción alguna.

Por último, consideremos el *flujo de energía* asociado a una onda sonora. Lo definiremos como la cantidad total de energía mecánica (potencial y cinética, asociada con las oscilaciones elásticas de los puntos del medio) que es transferida durante cada segundo a través de la unidad de superficie (1 m^2) perpendicular a la dirección de la propagación (Fig. 3.9). Este flujo de energía se expresa en J por m^2 y segundo o, tomando en cuenta la definición y las unidades de potencia (ecuación [3.2]), en W/m^2 . Esto se llama, comúnmente, la *intensidad* de la onda, y se la indica con la letra I . Puede demostrarse que existe una relación entre la intensidad de una onda sonora sinusoidal y el valor de la *variación (absoluta) media de presión* asociada con la onda (véase Fig. 3.7), que indicamos con Δp (amplitud de la variación de presión, dividida por $\sqrt{2}$):

$$I = \frac{(\Delta p)^2}{V \delta}$$

En esta relación, V es la velocidad de la onda sonora (3.6) y δ es la densidad del aire. Para condiciones normales de temperatura y presión, se obtiene la siguiente relación numérica:

$$I = 0,00234 \times (\Delta p)^2 \quad (\text{en } \text{W}/\text{m}^2) \quad (3.10)$$

Δp debe expresarse en Newton/m^2 . Como veremos en la Sec. 3.4, el tono más débil que puede ser percibido a una frecuencia de 1.000 Hz tiene una intensidad de solo $10^{-12} \text{ W}/\text{m}^2$. Según la relación 3.10, esto representa una variación de presión promedio de $2,0 \times 10^{-5} \text{ Newton}/\text{m}^2$, es decir, sólo $2,0 \times 10^{-10}$ veces la presión atmosférica normal. ¡Esto da una idea de cuán sensitivo es el oído!

Una fuente sonora dada (un instrumento musical o un parlante) emite ondas sonoras en todas las direcciones. En general, la cantidad de energía emitida por segundo depende de la particular dirección considerada. Supongamos que I_1 sea la intensidad de la onda en el punto A_1

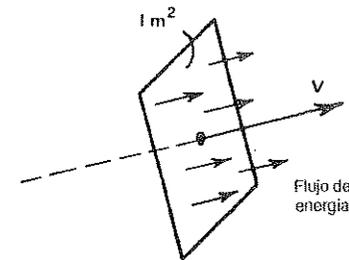


Figura 3.9 Flujo de energía a través de una unidad de superficie perpendicular a la dirección de propagación.

47

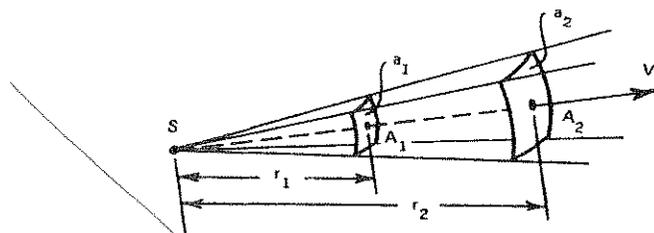


Figura 3.10 Flujo radial de energía sonora.

propagándose en la dirección indicada en la Fig. 3.10. Esto significa que una cantidad de energía $I_1 a_1$ pasa por la superficie a_1 durante cada segundo. Suponiendo que en el trayecto no haya pérdidas, la misma cantidad de energía pasará por segundo a través de la superficie a_2 en el punto A_2 . De ahí,

$$I_1 a_1 = I_2 a_2$$

Dado que las áreas de las superficies a_1 y a_2 son proporcionales al cuadrado de sus respectivas distancias r_1 y r_2 a la fuente, la intensidad de una onda sonora varía en forma inversamente proporcional al cuadrado de la distancia a la fuente:

$$\frac{I_1}{I_2} = \left(\frac{r_2}{r_1}\right)^2 \quad (3.11)$$

Esta ley ya no vale si se consideran las reflexiones y absorciones del sonido. Si imaginamos a la fuente sonora rodeada por completo de una superficie esférica, la cantidad total de energía que pasa por segundo a través de esa superficie se llama *potencia acústica de salida* de la fuente. Representa la energía que la fuente emite por segundo en todas las direcciones en forma de ondas sonoras. Su valor está dado en W (J/s). Los instrumentos típicos irradian entre 0,01 W (clarinete) y 6,4 W (trombón tocando fortissimo).

3.3 Superposición de ondas; ondas estacionarias

En ausencia de paredes reflectoras, las ondas sonoras se propagan en línea recta desde la fuente. Como se demostró en la sección anterior, su intensidad decrece rápidamente, proporcional a $1/r^2$, donde r es la distancia a la fuente. Si tenemos más de una fuente, las ondas emitidas por cada una se propagarán individualmente como si las otras no existieran, y el efecto resultante en un punto dado del medio (por ejemplo en el canal auditivo de un oyente) será una oscilación de presión dada por la suma

algebraica de las oscilaciones de presión de las ondas individuales⁴. En otras palabras, *las ondas sonoras se superponen linealmente*⁵. Esto incluso ocurre con las ondas de la membrana basilar para sonidos de baja intensidad, dando lugar a zonas de resonancia independientes entre sí (con envolturas del tipo de la mostrada en la Fig. 3.5).

Consideremos la superposición de dos ondas de tonos simples de frecuencias f_1 y f_2 y, según la relación (3.7), de longitudes de onda $\lambda_1 = V/f_1$ y $\lambda_2 = V/f_2$, propagándose en la misma dirección. Con el fin de obtener una «instantánea» de las variaciones de presión resultante, simplemente debemos sumar los valores correspondientes a las variaciones de presión individuales de cada una de las ondas, en cada punto x a lo largo de la dirección de propagación. Dado que la velocidad de las ondas sonoras no depende de la frecuencia (ni del patrón vibratorio en sí), todos los puntos del medio repetirán exactamente el mismo patrón vibratorio complejo, sólo sujetos a una diferente fase temporal. El flujo de energía, es decir, la intensidad de la superposición de dos (o más) ondas que viajan en la misma dirección con fases mezcladas al azar será simplemente la suma de los flujos de energía aportados individualmente por cada una de las ondas componentes:

$$I = I_1 + I_2 + I_3 + \dots \quad (3.12)$$

Un caso particularmente importante es aquel de dos ondas sinusoidales de igual frecuencia y amplitud viajando en direcciones opuestas. Esto ocurre, por ejemplo, cuando una onda sinusoidal se refleja en un punto dado (sin absorción), superponiéndose con la onda que va llegando. Estudiemos primero el caso de ondas transversales en una cuerda (Fig. 3.11). Sumando las contribuciones de cada componente, obtenemos otra onda sinusoidal de la misma frecuencia pero de diferente amplitud. El hecho sorprendente, sin embargo, es que el perfil de esta onda resultante *no se propaga en absoluto!* Queda anclada en ciertos puntos N_1, N_2, N_3, \dots llamados *nodos*, los cuales no vibran. Los puntos entre los nodos vibran con diferentes amplitudes, según sea su posición. En particular, los puntos A_1, A_2, A_3, \dots (a mitad de camino entre los nodos), llamados *vientres*^{*}, vibran con el máximo de amplitud (el doble de aquella de cada onda componente). La Fig. 3.12 muestra sucesivas formas de una cuerda, cuando dos ondas sinusoidales de la misma amplitud y frecuencia viajan en direcciones opuestas. Esto se llama *onda estacionaria*. Los puntos oscilan, pero no hay ninguna evidencia de propagación. El perfil de la

⁴ Obsérvese con atención que lo que se suma aquí son *variaciones* de presión, y no valores absolutos de presión.

⁵ Esto no se cumple en el caso de ondas que corresponden a sonidos muy intensos (estruendosos), como los de una explosión.

* Menos frecuentemente llamados *antinodos* (N. del T.).

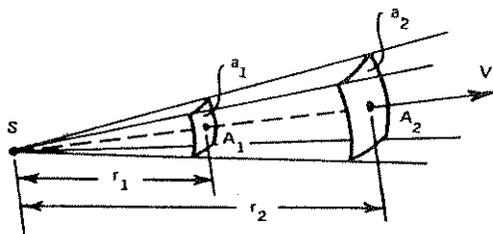


Figura 3.10 Flujo radial de energía sonora.

propagándose en la dirección indicada en la Fig. 3.10. Esto significa que una cantidad de energía $I_1 a_1$ pasa por la superficie a_1 durante cada segundo. Suponiendo que en el trayecto no haya pérdidas, la misma cantidad de energía pasará por segundo a través de la superficie a_2 en el punto A_2 . De ahí,

$$I_1 a_1 = I_2 a_2$$

Dado que las áreas de las superficies a_1 y a_2 son proporcionales al cuadrado de sus respectivas distancias r_1 y r_2 a la fuente, la intensidad de una onda sonora varía en forma inversamente proporcional al cuadrado de la distancia a la fuente:

$$\frac{I_1}{I_2} = \left(\frac{r_2}{r_1}\right)^2 \quad (3.11)$$

Esta ley ya no vale si se consideran las reflexiones y absorciones del sonido. Si imaginamos a la fuente sonora rodeada por completo de una superficie esférica, la cantidad total de energía que pasa por segundo a través de esa superficie se llama *potencia acústica de salida* de la fuente. Representa la energía que la fuente emite por segundo en *todas* las direcciones en forma de ondas sonoras. Su valor está dado en W (J/s). Los instrumentos típicos irradian entre 0,01 W (clarinete) y 6,4 W (trombón tocando fortissimo).

3.3 Superposición de ondas; ondas estacionarias

En ausencia de paredes reflectoras, las ondas sonoras se propagan en línea recta desde la fuente. Como se demostró en la sección anterior, su intensidad decrece rápidamente, proporcional a $1/r^2$, donde r es la distancia a la fuente. Si tenemos más de una fuente, las ondas emitidas por cada una se propagarán individualmente como si las otras no existieran, y el efecto resultante en un punto dado del medio (por ejemplo en el canal auditivo de un oyente) será una oscilación de presión dada por la suma

algebraica de las oscilaciones de presión de las ondas individuales⁴. En otras palabras, *las ondas sonoras se superponen linealmente*⁵. Esto incluso ocurre con las ondas de la membrana basilar para sonidos de baja intensidad, dando lugar a zonas de resonancia independientes entre sí (con envolventes del tipo de la mostrada en la Fig. 3.5).

Consideremos la superposición de dos ondas de tonos simples de frecuencias f_1 y f_2 y, según la relación (3.7), de longitudes de onda $\lambda_1 = V/f_1$ y $\lambda_2 = V/f_2$, propagándose en la *misma* dirección. Con el fin de obtener una «instantánea» de las variaciones de presión resultante, simplemente debemos sumar los valores correspondientes a las variaciones de presión individuales de cada una de las ondas, en cada punto x a lo largo de la dirección de propagación. Dado que la velocidad de las ondas sonoras no depende de la frecuencia (ni del patrón vibratorio en sí), todos los puntos del medio repetirán exactamente el mismo patrón vibratorio complejo, solo sujetos a una diferente fase temporal. El flujo de energía, es decir, la intensidad de la superposición de dos (o más) ondas que viajan en la misma dirección con fases mezcladas al azar- será simplemente la suma de los flujos de energía aportados individualmente por cada una de las ondas componentes:

$$I = I_1 + I_2 + I_3 + \dots \quad (3.12)$$

Un caso particularmente importante es aquel de dos ondas sinusoidales de igual frecuencia y amplitud *viajando en direcciones opuestas*. Esto ocurre, por ejemplo, cuando una onda sinusoidal se refleja en un punto dado (sin absorción), superponiéndose con la onda que va llegando. Estudiemos primero el caso de ondas transversales en una cuerda (Fig. 3.11). Sumando las contribuciones de cada componente, obtenemos otra onda sinusoidal de la misma frecuencia pero de diferente amplitud. El hecho sorprendente, sin embargo, es que el perfil de esta onda resultante *no se propaga en absoluto!* Queda anclada en ciertos puntos N_1, N_2, N_3, \dots llamados *nodos*, los cuales no vibran. Los puntos entre los nodos vibran con diferentes amplitudes, según sea su posición. En particular, los puntos A_1, A_2, A_3, \dots (a mitad de camino entre los nodos), llamados *vientres**, vibran con el máximo de amplitud (el doble de aquella de cada onda componente). La Fig. 3.12 muestra sucesivas formas de una cuerda, cuando dos ondas sinusoidales de la misma amplitud y frecuencia viajan en direcciones opuestas. Esto se llama *onda estacionaria*. Los puntos oscilan, pero no hay ninguna evidencia de propagación. El perfil de la

⁴ Obsérvese con atención que lo que se suma aquí son *variaciones* de presión, y no valores absolutos de presión.

⁵ Esto no se cumple en el caso de ondas que corresponden a sonidos muy intensos (estruendosos), como los de una explosión.

* Menos frecuentemente llamados *antinodos* (N. del T.).

48

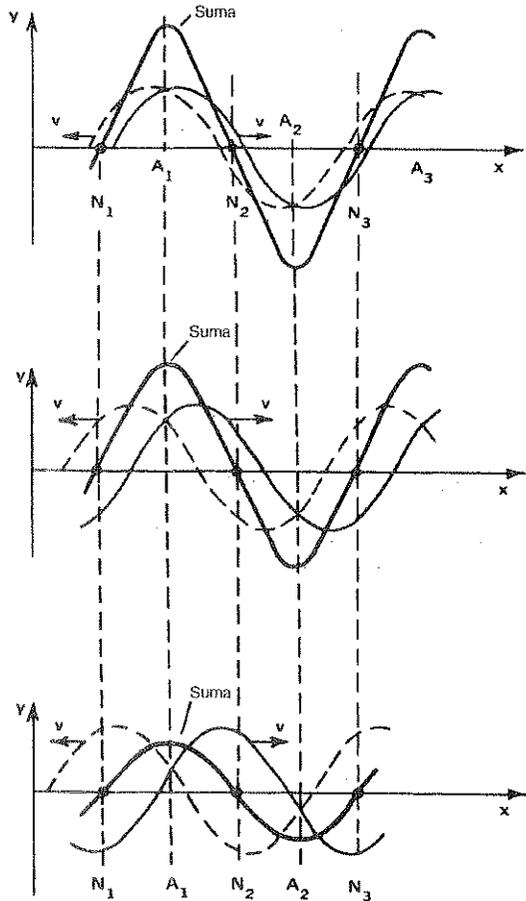


Figura 3.11 Superposición de dos ondas transversales de la misma amplitud y frecuencia, propagándose en direcciones opuestas, +V y -V. La onda resultante no se propaga: queda «anclada» en los nodos N, cambiando solo de amplitud.

onda cambia en amplitud, pero no se desliza, ni hacia la derecha ni hacia la izquierda. En un momento dado (t_1), la cuerda exhibe una deformación máxima; en otro (t_5) la cuerda no tiene deformación alguna. Como veremos en el próximo capítulo, las ondas estacionarias tienen un rol clave en música, especialmente en los mecanismos de generación sonora de los instrumentos musicales.

En una onda estacionaria *no hay propagación neta de energía*. La cuerda en su totalidad actúa casi como un resorte elástico en vibración: en un

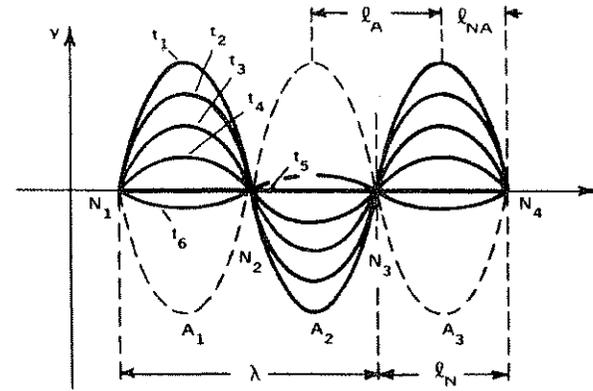


Figura 3.12 Aspectos sucesivos de una cuerda oscilando en forma estacionaria.

instante dado (por ej. t_5 en la Fig. 3.12), todos los puntos están pasando a través de su posición de equilibrio, y la energía de la cuerda aparece toda en forma cinética (energía de movimiento). En otro instante (por ej. t_1 en la Fig. 3.12), todos los puntos están en su máximo desplazamiento y toda la energía es potencial. Obsérvese con atención que esto no ocurre con una onda que se propaga: en la Fig. 3.3, por ejemplo, en cada instante hay puntos que tienen desplazamiento máximo (solo poseen energía potencial), desplazamiento cero (solo energía cinética), o que se hallan en cualquier situación intermedia (ambas formas de energía). Lo que es más: en una onda viajera todos los puntos tienen la misma amplitud; lo que varía es el instante de tiempo en el que alcanzan el desplazamiento máximo (los puntos están fuera de fase).

Un análisis de la Fig. 3.12 revela que la distancia l_N entre dos nodos vecinos N_1, N_2 o la distancia l_A entre dos vientres A_1, A_2 es exactamente una mitad de longitud de onda λ :

$$l_N = l_A = \lambda/2 \tag{3.13}$$

Por otra parte, la distancia l_{NA} entre un nodo N_1 y un vientre A_1 es un cuarto de longitud de onda:

$$l_{NA} = \lambda/4 \tag{3.14}$$

Por supuesto, también hay ondas estacionarias longitudinales. Aparecen cuando dos ondas sonoras de la misma frecuencia y amplitud de variación de presión viajan en direcciones opuestas. Esto ocurre, por ejemplo, cuando una onda sonora se propaga a lo largo de un tubo y se refleja en el otro extremo. Ondas estacionarias también surgen como resultado de reflexiones en las paredes de salas y recintos. Tienen las mismas propiedades que las ondas estacionarias transversales, y la discusión

anterior también es válida para ellas. Sin embargo, debemos hacer una importante observación. Como señalamos en la sección anterior, el modo más conveniente de representar las ondas sonoras es como oscilaciones de presión. Vimos que los puntos del medio (aire) que exhiben una máxima variación de presión tienen cero desplazamiento longitudinal (Fig. 3.6), mientras que los lugares con cero variación de presión corresponden a los puntos con máximo desplazamiento. Podemos traducir esto al caso de una onda longitudinal estacionaria: los *nodos de presión* (es decir, puntos en que las variaciones de presión son constantemente nulas) son *vientres de vibración* (puntos que oscilan con máxima amplitud), en tanto que los *vientres de presión* (puntos en que la presión oscila con máxima amplitud) son *nodos de vibración* (puntos que están permanentemente en reposo).

3.4 Intensidad, nivel de intensidad sonora y sonoridad

En la Sec. 2.3 dijimos que, al escuchar un tono simple, la amplitud de las oscilaciones del tímpano conduce a la sensación de sonoridad. Esta amplitud está directamente relacionada con la variación media de presión Δp de la onda sonora, y por tanto, al flujo de energía acústica o intensidad I que llega al oído (relación [3.10]). Pasamos a investigar ahora el rango de intensidades a las cuales el oído es sensitivo, empezando por ondas correspondientes a tonos simples. Hay dos límites de sensibilidad a un tono de frecuencia dada: 1) un límite inferior o *umbral de la sensación auditiva*, que representa la intensidad mínima perceptible; 2) un *límite superior de la sensación auditiva*, más allá del cual se siente dolor, que eventualmente puede conducir a un daño físico del mecanismo auditivo. Estos dos límites varían de individuo a individuo y dependen de la frecuencia considerada. En general, para un tono de aproximadamente 1000 Hz (una altura entre las notas Si₅ y Do₆), el intervalo entre estos dos límites es máximo. El enorme rango de intensidades comprendido entre los dos límites es sorprendente. En efecto, para un tono de 1000 Hz el umbral de intensidad promedio es de aproximadamente 10^{-12} W/m², mientras que el límite del dolor se sitúa alrededor de 1 W/m². ¡Esto representa una razón de intensidades de un billón a uno, a las cuales el oído es sensitivo! La tabla 3.1 muestra las relaciones entre intensidad sonora y sensación musical de sonoridad para un tono de 1000 Hz⁶.

⁶Es muy arbitrario considerar la notación musical como representativa de valores absolutos de sonoridad. Los músicos afirmarán, por ejemplo, que son perfectamente capaces de percibir *fortísimos* y *pianísimos* al escuchar música de una radio cuyo volumen ha sido bajado al nivel de un susurro. Lo que ocurre en ese caso es que se usan *otras* características, además de la intensidad, para abrir juicios subjetivos sobre una sonoridad «relativa». Por otra parte, se ha encontrado en experimentos sistemáticos (Patterson, 1974), que la interpretación de la notación de la sonoridad musical en un contexto dinámico depende mucho del instrumento y del rango de alturas cubierto.

TABLA 3.1 Comparación entre intensidad sonora y sensación musical de sonoridad.

Intensidad (Watt/m ²)	Sonoridad
1	Límite de dolor
10 ⁻³	<i>fff</i>
10 ⁻⁴	<i>ff</i>
10 ⁻⁵	<i>f</i>
10 ⁻⁶	<i>mf</i>
10 ⁻⁷	<i>p</i>
10 ⁻⁸	<i>pp</i>
10 ⁻⁹	<i>ppp</i>
10 ⁻¹²	Umbral de audición

A 1000 Hz, el rango de intensidades de interés *musical* se extiende de alrededor de 10^{-9} W/m² a 10^{-2} W/m². ¡Esto todavía representa un factor de variación del orden de los 10 millones!

A causa de este rango tremendo, la unidad W/m² no resulta práctica. Existe otra razón por la cual no lo es. En Psicofísica, la diferencia apenas perceptible (DAP) de un estímulo dado es habitualmente una buena unidad para la magnitud física correspondiente. Experimentos psicoacústicos muestran que la *DAP en intensidad sonora* es más o menos proporcional a la intensidad del tono. Esta proporcionalidad sugiere que una unidad más apropiada para *I* debería aumentar gradualmente, a medida que se incrementa la intensidad del tono que queremos describir. Esto llevaría a una gran complicación. Es más práctico introducir una magnitud física diferente que sea una función apropiada de la intensidad *I*. Esta nueva magnitud deberá cumplir simultáneamente con tres objetivos: 1) «comprimir» la escala de intensidades audibles en un rango de valores mucho más pequeño; 2) reemplazar el uso de valores absolutos por valores relativos (por ejemplo, relativos al umbral de audición); y 3) permitir la introducción de una unidad más conveniente, cuyo valor sea del orden del mínimo cambio perceptible en intensidad sonora.

La introducción de esta nueva cantidad se hace de la siguiente manera. Obsérvese en la tabla 3.1 que la característica que está más relacionada con el efecto de sonoridad es el *exponente* al cual es elevado el número 10 cuando citamos el valor de la intensidad sonora (columna izquierda); -12 para el umbral de la sensación sonora, -9 para un sonido *ppp*; -7 para *p*, -5 para *forte*, -3 para *forte-fortissimo*, y 0 para el límite del dolor ($10^0 = 1$). Esto sugiere claramente que deberíamos usar lo que en matemáticas se llama función logarítmica para representar la intensidad.

Los *logaritmos decimales*⁷ sirven para definir una magnitud más

⁷El *logaritmo decimal* de un número dado es el exponente al cual tiene que ser elevado 10 para obtener ese número. Por ejemplo, el logaritmo de 100 es 2 porque $10^2 = 100$; el

apropiada para describir la intensidad del sonido. En primer lugar, adoptamos el umbral de la sensación sonora (a 1000 Hz) de 10^{-12} W/m^2 como nuestra intensidad de referencia I_0 . Por lo tanto introducimos la cantidad

$$IL = 10 \times \log \frac{I}{I_0} \quad (3.15)$$

Esto se llama *nivel de intensidad sonora**. La unidad de IL se llama *decibel* y se abrevia «db». Para el umbral de la audición, $I/I_0 = 1$ y $IL = 0$ db. Para el límite superior de la audición tenemos $I/I_0 = 10^{12}$ y $IL = 120$ db. Un sonido típicamente «forte» (tabla 3.1) tiene un nivel de intensidad sonora de 70 db; un sonido *ppp* corresponde a 30 db.

Es importante tener presente que cuando una cantidad dada se expresa en decibeles, se está dando una medida *relativa* con respecto a algún valor de referencia (por ejemplo, en la definición de IL , el umbral de la sensación auditiva). Cuando la intensidad se multiplica por 10, se *suman* diez decibeles al IL ; cuando la intensidad se multiplica por 100, se *suman* 20 db, etc. Del mismo modo, cuando la intensidad se divide por 100, se *restan* 20 db al IL . La tabla 3.2 muestra algunas relaciones útiles.

TABLA 3.2 Comparación de cambios en el nivel de intensidad sonora IL (en db) con los correspondientes cambios en la intensidad I .

Cambio en el IL	Qué ocurre con la intensidad
suma (resta) 1db	Se multiplica (divide) por 1,26
+ (-) 3db	$\times (+) 2$
+ (-) 10db	$\times (+) 10$
+ (-) 20db	$\times (+) 100$
+ (-) 60db	$\times (+) 1.000.000$

Podemos usar la relación (3.10) para expresar la intensidad en términos de la variación media de presión Δp , que es mucho más fácil de medir. Encontramos que el umbral de sensación I_0 a 1000 Hz corresponde aproximadamente a una variación media de presión $\Delta p_0 = 2 \times 10^{-5}$

logaritmo de 10.000 es 4 porque $10^4 = 10.000$; el logaritmo de 1 es cero porque $10^0 = 1$; y el logaritmo de 0,000001 es -6 porque $10^{-6} = 0,000001$. Estas relaciones se escriben simbólicamente: $\log 100=2$; $\log 10.000=4$; $\log 1=0$; $\log 0,000001 = -6$. El logaritmo de cualquier número cuyo valor esté ubicado entre las potencias enteras de diez puede ser encontrado mediante el uso de tablas.

Una propiedad importante es que el logaritmo del *producto* de dos números es igual a la *suma* de los logaritmos de los números individuales. Por ejemplo, el logaritmo de 10^4 por 10^3 es 4 *más* 3 (o sea 7), porque $10^4 \times 10^3 = 10^{4+3} = 10^7$. En general para dos números cualesquiera a y b , tenemos la relación $\log (a \times b) = \log a + \log b$. Para el logaritmo de una división a/b tenemos en cambio $\log (a/b) = \log a - \log b$.

* *Sound intensity level* (N. del T.).

Newton/m² (20 micropascal). Dado que de acuerdo con la relación (3.10) I es proporcional al cuadrado de Δp , tenemos que

$$\log \frac{I}{I_0} = \log \left(\frac{\Delta p}{\Delta p_0} \right)^2 = 2 \times \log \frac{\Delta p}{\Delta p_0}$$

De aquí, podemos introducir la cantidad

$$SPL = 20 \times \log \frac{\Delta p}{\Delta p_0} \quad (3.16)$$

llamada *nivel de presión sonora* (SPL)*. Para el caso de una onda de propagación, los valores de las ecuaciones (3.15) y (3.16) son idénticos: IL y SPL representan la misma cosa. Pero en el caso de ondas *estacionarias*, en vista de que no hay ningún *flujo* de energía (Sec. 3.3), la intensidad I tal como es usada en la ecuación (3.15) no puede ser definida; de ahí que el IL pierda su sentido físico. Pero el concepto de variación media de presión Δp en un punto determinado del espacio (por ejemplo, en el interior de un tubo de órgano) sigue vigente y lo mismo ocurre con SPL, el nivel de presión sonora. Esta es la razón por la cual la relación (3.16) se usa con más frecuencia que la ecuación (3.15). Obsérvese cuidadosamente que las definiciones de IL y SPL no involucran de ningún modo la frecuencia de la onda sonora. Aunque hicimos referencia a un tono de 1000 Hz, nada impide definir el IL y el SPL según las relaciones (3.15) y (3.16), respectivamente, para tonos de frecuencia arbitraria. Lo que sí depende de la frecuencia, y mucho, son los límites subjetivos de la audición (por ej., I_0 y Δp_0) y, en general, la sensación *subjetiva* de sonoridad, como veremos más adelante.

Cuando superponemos dos tonos de la misma frecuencia, cosas curiosas parecen ocurrir con el nivel de intensidad sonora, o con el nivel de presión sonora. Véase la tabla 3.2: el sumar dos tonos de la misma intensidad, que según la ecuación (3.12) significa duplicar la intensidad, agrega meramente 3 db al nivel sonoro del sonido original, cualquiera haya sido el valor del IL . El superponer *diez* tonos iguales (en fase) solo incrementa en 10 db el IL resultante. Para elevar el IL de un tono dado en 1 db, debemos multiplicar su intensidad por 1,26, lo que significa que es menester sumarle un tono cuya intensidad sea 0,26 (alrededor de 1/4) la del tono original.

El cambio mínimo de SPL requerido para un cambio detectable en la sensación de sonoridad (DAP de nivel sonoro) es más o menos constante y del orden de 0,2 - 0,4 db en el rango musicalmente relevante de altura y sonoridad. La unidad de IL o de SPL, el decibel, resulta de este modo ser

* SPL: *sound pressure level* (N. del T.).

una «unidad razonable», cercana a la DAP.

Existe una alternativa para enfocar la cuestión de la DAP de la intensidad o del nivel sonoro. En lugar de preguntarnos cuánto debemos cambiar la intensidad de un tono dado para producir un cambio mínimo perceptible, podemos plantear la cuestión de un modo totalmente equivalente: ¿Cuál es la mínima intensidad I_2 que un *segundo* tono de la misma frecuencia y fase debe tener para que sea advertido en presencia del primero (cuya intensidad I_1 se mantiene constante)? Esa mínima intensidad I_2 se llama el *umbral de enmascaramiento*. El tono original de intensidad constante I_1 se llama «tono enmascarador», y el tono adicional, «tono enmascarado». El enmascaramiento tiene un rol muy importante en música. Aquí solo hemos mencionado el enmascaramiento de tonos de idéntica frecuencia y fase; más adelante trataremos el enmascaramiento de sonidos de distinta frecuencia. La relación entre el nivel de enmascaramiento ML (IL del tono enmascarado en el umbral), la DAP del nivel sonoro, y el IL del tono enmascarador pueden hallarse partiendo de las expresiones que los definen (basadas en la ecuación (3.15)):

$$ML = 10 \log (I_2/I_0); \quad DAP = 10 \log [(I_1 + I_2)/I_1]; \quad IL = 10 \log (I_1/I_0)$$

Hasta ahora estuvimos trabajando con las cantidades físicas IL y SPL. En lo que sigue, examinaremos la magnitud psicológica *sonoridad*, asociada a un SPL determinado. En las secciones 1.4 y 2.3 mencionamos la capacidad que poseemos para establecer un orden entre dos sensaciones del mismo tipo, y señalamos que surgen dificultades cuando queremos hacer comparaciones cuantitativas, absolutas. En el caso de la sonoridad, los juicios emitidos por diferentes individuos acerca de si dos tonos simples suenan igual de 'fuerte' son coincidentes. Pero un juicio sobre *cuánto* más 'fuerte' o 'sonoro' es un tono que otro, requiere previo entrenamiento y los resultados, en general, difieren mucho de individuo a individuo.

Tonos de igual SPL pero de diferente frecuencia son juzgados en general como de distinta sonoridad. Por lo tanto, el SPL no es una buena medida de la sonoridad o 'fuerza' del sonido cuando se trata de comparar tonos de diferente frecuencia. Se han realizado experimentos para establecer *curvas de igual sonoridad*,⁸ tomando el SPL a 1.000 Hz como cantidad de referencia. Estas están graficadas en la Fig. 3.13 (Fletcher y Munson, 1933). Partiendo del eje vertical centrado en 1.000 Hz se trazan curvas correspondientes a los SPL de tonos que son juzgados como «igualmente fuertes» que el tono de 1.000 Hz de referencia. Obsérvese, por ejemplo, que mientras un SPL de 50 db (una intensidad de 10^{-7} W/m^2) es considerado un sonido «piano» a 1000 Hz, el mismo SPL es apenas audible

⁸Obtenidas a través de experimentos en 'apareamiento de sonoridad', conducidos de manera muy similar a los de 'apareamiento de altura tonal'.

a 60 Hz. En otras palabras, para producir una sensación de sonoridad determinada en frecuencias bajas - digamos un «forte» -, se necesita una intensidad mucho mayor (un mayor flujo de energía o mayor variación media de presión) que el requerido a 1000 Hz. Esta es la razón por la cual los tonos graves parecen «extinguirse» mucho antes que los agudos, cuando nos alejamos gradualmente de una fuente sonora fija. Por la misma razón tenemos que pagar tanto más por un equipo de audio de alta fidelidad, especialmente por los parlantes, si queremos bajos bien balanceados.

La curva más baja de la Fig. 3.13 representa el umbral de la sensación auditiva a distintas frecuencias. De nuevo se puede ver cómo el oído va siendo mucho menos sensitivo hacia las frecuencias bajas (y también hacia las muy altas). La mayor sensibilidad se da alrededor de los 3000 Hz. La forma de la curva que representa el umbral está influenciada por las propiedades acústicas del canal auditivo (meatus) y por las propiedades mecánicas de la cadena de huesos del oído medio. Finalmente, debe quedar claro que las curvas de la Fig. 3.13 se refieren solo a *tonos simples aislados, sonando de manera continua*. Más adelante discutiremos qué pasa con la sensación de sonoridad en el caso de que el tono sea de corta duración (fracción de un segundo). Estudios detallados (Molino, 1973) muestran que las curvas de igual sonoridad dependen de la frecuencia del tono de referencia (que fue de 1000 Hz en la Fig. 3.13).

Llegamos ahora a una cuestión que, a veces, resulta confusa. Introducimos una nueva cantidad, llamada *nivel de sonoridad*, que indicaremos LL^* . Se define del modo siguiente: el LL de un tono de frecuencia f está dado por el SPL de un tono de 1000 Hz que se juzga de la misma sonoridad. Esto significa que las curvas de la Fig. 3.13 son curvas de LL constante. La unidad de LL se llama *fon*. La figura 3.13 puede ser usada para encontrar el LL de un tono de un SPL determinado, a cualquier frecuencia f . Por ejemplo, considérese un tono de un SPL de 70 db ($I = 10^{-5} \text{ W/m}^2$) a 80 Hz. Vemos que la curva que pasa por ese punto corta la línea de los 1000 Hz a 50 db. El LL de ese tono entonces será de 50 fones. En general los números a lo largo de la línea de los 1000 Hz representan el LL en fones de la curva correspondiente de sonoridad constante. Nótese que el LL es una cantidad física, y no psicológica (a pesar de su nombre). Representa aquellas intensidades o valores de SPL que suenan igual de 'fuerte', pero no intenta representar la sonoridad de manera absoluta: un tono cuyo LL es el doble de otro, simplemente no suena el doble de 'fuerte'. Se han realizado muchos estudios estadísticos para definir una escala subjetiva de sonoridad. La figura 3.14 (la línea gruesa) muestra resultados que relacionan la «sonoridad subjetiva» L con el *loudness level* LL , en el rango de interés musical. La cantidad L que

* LL : *loudness level* (N. del T.).

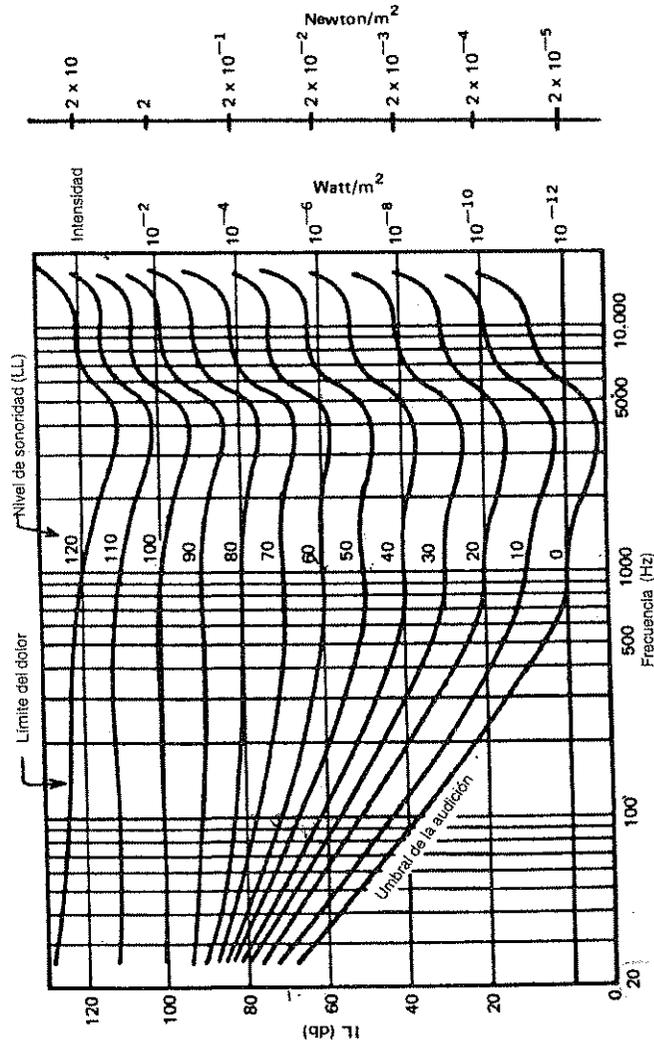


Figura 3.13 Curvas de igual sonoridad (Fletcher y Munson, 1933) en un diagrama de nivel de intensidad sonora (LL) y frecuencia. También se muestran las escalas de intensidad sonora y variación media de presión correspondientes. Reimpreso con permiso del *Journal of the Acoustical Society of America*.

describe sonoridad subjetiva se expresa en unidades llamadas *sones*. Obsérvese que la relación no es lineal (la escala de sonoridad de la Fig. 3.14 es logarítmica). Cuando se incrementa el LL en 10 fones la sonoridad *L* se *duplica*. Esto significa, por ejemplo, ¡que *diez* instrumentos que tocan una nota determinada con un mismo LL serán juzgados meramente como sonando con *dobles* sonoridad que uno de los instrumentos tocando solo!

Se ha demostrado que la relación entre *L* y la intensidad de la onda *I*, o la variación media de presión Δp , puede describirse aproximadamente por la siguiente simple función (Stevens, 1970):

$$L = C_1 \sqrt[3]{I} = C_2 \sqrt[3]{(\Delta p)^2} \tag{3.17}$$

donde C_1 y C_2 son parámetros que dependen de la frecuencia. Esto da la línea punteada en la Fig. 3.14, que se encuentra bien dentro de la fluctuación estadística de las mediciones reales (no indicada). Obsérvese que la relación logarítmica (como en las ecuaciones (3.15) y (3.16)) ha desaparecido. Sin embargo, queda todavía una «compresión» apreciable de la sonoridad subjetiva: para variar *L* entre 1 y 200, la intensidad *I* debe incrementar un factor de 8 millones.

Cuando superponemos dos tonos de la misma frecuencia (y fases mezcladas aleatoriamente), el tono resultante tiene una intensidad (flujo de energía) que es la suma de las intensidades de los tonos componentes: $I = I_1 + I_2 + I_3 + \dots$ (3.12). Dado que en este caso los tonos individuales no pueden ser discriminados en forma separada, esta intensidad total determina la sonoridad resultante a través de la relación (3.17). *L* no será, obviamente, la suma de las sonoridades de los tonos individuales. La situación es diferente cuando las ondas superpuestas tienen *frecuencias distintas*. Podemos distinguir tres casos: 1) Si las frecuencias de los tonos componentes caen *dentro de la banda crítica* de la frecuencia central (Sec. 2.4), la sonoridad resultante sigue estando directamente relacionada con la intensidad total (flujo de energía total), suma de las intensidades individuales:

$$L = C_1 \sqrt[3]{(I_1 + I_2 + I_3 + \dots)} \tag{3.18}$$

Esta propiedad efectivamente conduce a una determinación de la banda crítica más precisa (Zwicker, Flottorp y Stevens, 1957) que la dada en la Sec. 2.4. 2) Cuando la diferencia entre las frecuencias que componen el estímulo *excede la banda crítica*, la sonoridad subjetiva resultante es *más grande* que la obtenida por simple sumatoria de intensidades (3.18), incrementándose con la diferencia de frecuencias y tendiendo a un valor dado por la suma de las sonoridades individuales aportadas por bandas críticas adyacentes (Zwicker y Scharf, 1965):

$$L = L_1 + L_2 + L_3 + \dots \tag{3.19}$$

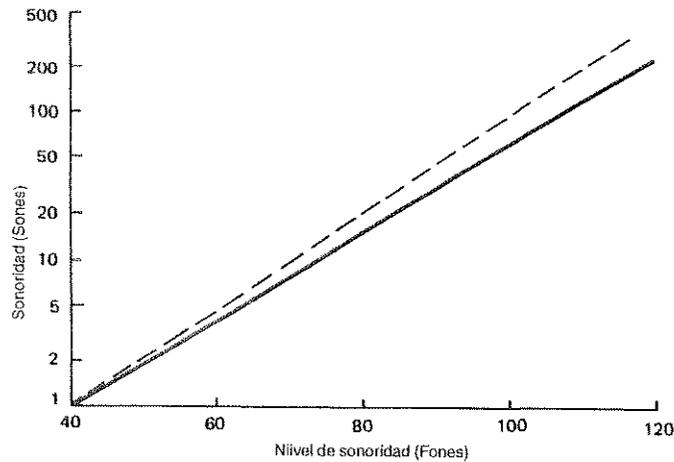


Figura 3.14 Línea continua: relación experimental entre la magnitud psicológica sonoridad y la magnitud física nivel de intensidad (según Stevens, 1955). Línea quebrada: relación potencial (3.17) (Stevens, 1970).

Si las sonoridades individuales L_1 , L_2 , etc., difieren mucho entre sí, hay que tener en cuenta los efectos de enmascaramiento. El límite de integración de sonoridades (3.19) nunca es alcanzado en la práctica. 3) Cuando la diferencia de frecuencias entre los tonos individuales es grande, la situación se torna complicada. Ante todo surgen dificultades con el concepto mismo de sonoridad total. El oyente tiende a concentrarse en sólo *uno* de los tonos componentes (por ejemplo, el más 'fuerte', o el de altura más alta) y asignar la sensación de sonoridad total a ese único componente:⁹

$$L = \text{máxima o más alta} \} \text{ de } (L_1, L_2, \dots) \quad (3.20)$$

Todo esto tiene consecuencias muy importantes para la música. Por ejemplo, dos tubos de órgano del mismo tipo y de la misma altura suenan sólo 1,3 veces más 'fuerte' que un tubo solo. Cuando sus alturas difieren entre sí en un semitono o un tono entero, la sonoridad total seguirá siendo aproximadamente 1,3 veces mayor que aquella correspondiente a un tubo solo (un semitono o un tono caen dentro de la banda crítica, Fig. 2.13).

⁹ Una situación similar, muy familiar, ocurre con la sensación de dolor. Si a usted lo pinchan en dos lugares muy cercanos el uno del otro, el dolor será «dos veces» (o más bien $\sqrt{2}$ veces) el causado por un solo pinchazo (equivalente al caso 1) de arriba). Pero cuando los lugares están bien separados, usted tendrá dificultad en decir cuánto es lo que podríamos llamar el «dolor total» (caso 3)); usted tenderá a percibir la sensación de mayor dolor, e ignorar la otra!

Pero dos tonos que estén separados por una tercera mayor sonarán más 'fuerte' que la combinación previa. Estos hechos han sido bien conocidos durante siglos por los constructores de órganos y por los compositores. Dado que no hay ninguna posibilidad de controlar por el toque la sonoridad de los tonos individuales en el órgano, contrariamente a lo que ocurre en los instrumentos de cuerda o en las maderas, la sonoridad en el órgano sólo puede alterarse cambiando el número de tubos que suenan simultáneamente. Pero, como acabamos de ver, la suma de sonoridades resulta más efectiva cuando los tonos componentes difieren en frecuencia de manera apreciable (relación [3.19]), los juegos que suenan una (registro de 4'), dos (registro de 2'), o más octavas arriba (y debajo [16'; 32']) de la nota escrita, son usados con ese propósito¹⁰. Por otra parte, la sonoridad puede ser controlada también por medio del número de notas tocadas simultáneamente, sin agregar o quitar registros. Cada nueva voz que entra en una fuga aumenta la sonoridad subjetiva de la pieza; lo mismo ocurre con cada nota adicional que se agrega a un acorde. Algunos organistas tocan los acordes finales de las fugas de Bach tirando registros adicionales. Esto es absurdo: ¡Bach mismo programó el incremento deseado en la sonoridad simplemente escribiendo más notas que el número de voces usadas en la fuga!¹¹

Hemos examinado la adición de sonoridades de dos o más tonos simultáneos, pero todavía no discutimos lo que ocurre con el umbral de la sensación de un tono cuando éste suena simultáneamente con otro. Si las frecuencias de ambos tonos coinciden, este umbral está dado por el nivel de enmascaramiento analizado antes (pág. 102). Si sus frecuencias difieren, también podemos determinar un nivel de enmascaramiento, definido como el nivel mínimo de intensidad que el tono enmascarado debe superar para poder ser 'detectado' y escuchado individualmente en presencia del tono enmascarador. El umbral de intensidad de tonos simples aislados (curva inferior en la Fig. 3.13) cambia apreciablemente, es decir, aumenta si otros tonos están presentes al mismo tiempo. La experiencia de enmascaramiento más familiar es aquella de no poder seguir una conversación en un lugar con mucho ruido de fondo. El nivel de enmascaramiento ML^* de un tono simple de frecuencia f en presencia de otro tono simple de características fijas (frecuencia de 415 Hz y varios niveles de intensidad IL) se muestra en la Fig. 3.15 (Egan, Hake, 1950). El nivel ML por encima del umbral normal de sensación al cual debe llegar el tono enmascarado para ser percibido (curva inferior en la Fig. 3.13), está

¹⁰ Su uso, obviamente, también contribuye a cambios en el timbre (Cap. 4).

¹¹ Muchos órganos de su tiempo tenían también un registro llamado «Zimbelstern», que hacía sonar pequeños platillos muy agudos o campanitas montadas sobre una estrella rotatoria en la parte superior de la caja del órgano; este registro se usaba para reforzar la sonoridad de un acorde final (sin interferir en modo alguno con su armonía).

* ML : masking level (nivel de enmascaramiento) (N. del T.).

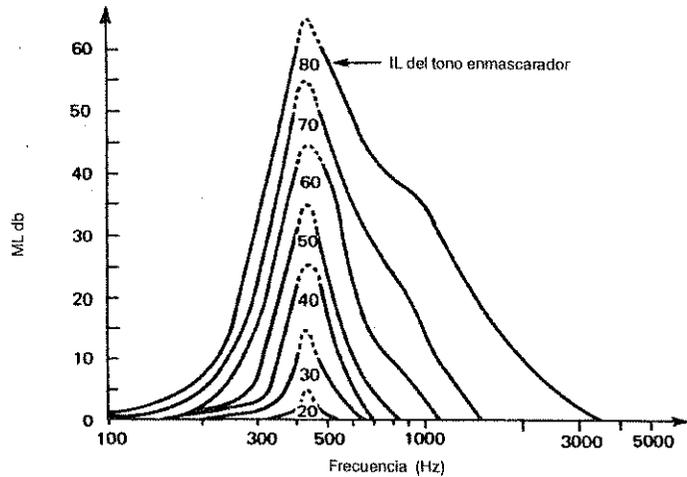


Figura 3.15 Nivel de enmascaramiento correspondiente a un tono puro de 415 Hz, para diferentes valores de nivel sonoro (Egan y Hake, 1950) del tono enmascarador. Reimpreso con permiso del *Journal of the Acoustical Society of America*.

graficado para distintos valores de IL del tono enmascarador. Las curvas en las regiones cercanas a f_0 deben extrapolarse a los valores deducidos de la DAP de sonoridad (pág. 101), por la presencia de fenómenos de batidos que nada tienen que ver con el enmascaramiento per se. A valores más altos de IL surgen complicaciones adicionales debido a la aparición de armónicos aurales de frecuencias $2f_0$, $3f_0$, etc. (Sec. 2.5). Obsérvese la asimetría de las curvas a IL más altos (causada por estos armónicos aurales): un tono de determinada frecuencia f_0 enmascara de manera más eficiente las frecuencias más altas que las más bajas.

El enmascaramiento es importante para la música polifónica, especialmente en la orquestación. Muchas veces en las partituras musicales, la participación de un instrumento dado, como ser un oboe o un fagot, puede pasar completamente inadvertida si ese instrumento toca al mismo tiempo que los metales tocan 'fortissimo'. Del mismo modo, en el órgano, la adición de juegos de flauta u otros registros débiles a un 'tutti' de principales, mixturas y lengüetas es totalmente irrelevante desde el punto de vista de la sonoridad.

Para concluir, debemos mencionar el efecto de la *duración* de un tono sobre la sensación de sonoridad (llamado *integración temporal*). Ante todo, existe un *umbral de tiempo*, una duración mínima que un tono simple debe tener para producir una sensación tonal, cualquiera que fuera. Esta duración mínima es de alrededor de los 10 - 15 ms, o por lo menos dos o

tres períodos de oscilación si la frecuencia está por debajo de los 50 Hz; los tonos que duran menos que esto se perciben como «clicks», no como «tonos». Los sonidos que duran más de 15 ms (o más de dos o tres períodos) pueden ser individualizados como tonos de determinada altura y sonoridad.¹² La sonoridad subjetiva, sin embargo, dependerá de la duración del tono (mientras que la altura no) (Plomp y Bouman, 1959). Para tonos de igual intensidad I , cuanto más corta la duración, tanto más baja será su sonoridad (por ej., Richards, 1977).

La Fig. 3.16 bosqueja la disminución relativa de sonoridad, o atenuación de sonoridad, en función de la duración del tono, para distintas frecuencias. Obsérvese que el valor asintótico de la respuesta es alcanzado antes para las frecuencias más altas; para duraciones de más de medio segundo, aproximadamente, la sonoridad alcanza un valor constante que depende sólo de la intensidad (relación dada en la Fig. 3.14). El enmascaramiento (Fig. 3.15) también tiene características que dependen del tiempo en el caso de tonos enmascaradores de corta duración; para detalles remitimos al lector a Zwicker y Fastl (1990). En general, se ve que para tonos de corta duración la sonoridad no está, a primera vista, relacionada con el flujo instantáneo de potencia (intensidad de la onda sonora), sino con la *energía* acústica total recibida por el oído (intensidad \times duración).

Efectivamente, hay indicios de que la sensación de sonoridad está en este caso relacionada con el número total de *impulsos nerviosos* que han sido generados por el tono corto (Zwislocki, 1969; véase próxima sección).

Cuando el tiempo de exposición al estímulo es largo, aparece un efecto conocido como *adaptación*; consiste en una *disminución* de la sonoridad subjetiva cuando un tono de intensidad constante es escuchado durante varios minutos¹³. Aunque existen notables diferencias entre individuos,

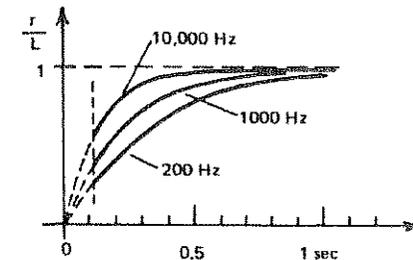


Figura 3.16 Sonoridad relativa de tonos puros de corta duración. r/L : razón de la sonoridad real (r) y la sonoridad (L) de un tono constante de igual frecuencia e igual amplitud.

¹² De manera general, hay un principio físico que dice que la frecuencia de una vibración no puede definirse con mayor precisión que el *inverso de la duración total* de la vibración.

¹³ Esto no debe ser confundido con la *fatiga* acústica, proceso psicológico por el cual nuestro cerebro es capaz de ignorar un sonido continuo entre otros de más relevancia.

experimentos en 'loudness matching' han revelado algunos principios generales (Scharf, 1983). La adaptación aumenta con la frecuencia para tonos simples del mismo SPL (tonos de frecuencias más altas «bajan de volumen» más rápido); para tonos de la misma frecuencia la adaptación disminuye cuando el SPL aumenta (cuando el valor de SPL supera los 40 db, aproximadamente, ya no hay más adaptación). En general, para el rango de frecuencias y SPL's usados en música, la sonoridad subjetiva se nivela al cabo de unos 100 segundos y luego permanece constante.

La atenuación de sonoridad en tonos de corta duración tiene consecuencias importantes para la ejecución musical. Si en el piano queremos tocar un pasaje «staccato» que suene con determinada 'fuerza' debemos bajar las teclas con más rudeza que si tuviéramos que tocar ese mismo pasaje «legato»¹⁴. Este efecto es mucho más pronunciado con tonos que no decaen en el tiempo, tales como los del órgano: el organista tiene un control considerable sobre la sonoridad subjetiva de una nota dada dándole la duración correcta. El fraseo en la ejecución organística es el arte de controlar la dinámica por medio de dar a las notas una duración apropiada (esto obviamente solo puede funcionar para tonos de corta duración; véase Fig. 3.16). Por otra parte, el fenómeno de la adaptación también se presenta en la música: ¡los trinos se han inventado para combatirla! En contraste, el uso de una «nota pedal» (una nota grave sonando constantemente) es prueba que la adaptación no es importante cuando las frecuencias son bajas. Dada la importancia de estos efectos temporales en la música, hace falta mucha más investigación psicoacústica en el tema de la atenuación y la adaptación.

3.5 El mecanismo de percepción de la sonoridad y procesos relacionados

¿Cuáles son los procesos físicos o neurales responsables de la diferencia entre la limitada escala subjetiva de sonoridad y el inmenso rango de intensidades de sonido detectables (tabla 3.1; relación [3.17])? En el caso de la percepción primaria de altura (Sec. 2.3) ya habíamos encontrado una compresión notable (sin mencionarla explícitamente): mientras que el rango de frecuencias audibles abarca desde los 20 hasta los 16.000 Hz aproximadamente, esto sólo representa nueve octavas de altura. En este caso la compresión se debe principalmente a las propiedades mecánicas de resonancia de la cóclea: la curva de la Fig. 2.8 representa una relación

¹⁴ Con los tonos de los instrumentos musicales surgen complicaciones: durante el ataque del tono, que puede durar varias décimas de segundo, un cambio en la intensidad y en el espectro ocurre en la misma fuente. También son importantes los efectos de reverberación (Sec. 4.7).

más o menos logarítmica entre la posición x de la región de resonancia en la membrana basilar y la frecuencia.

En lo que respecta al mecanismo de detección de sonoridad, la compresión es en parte neural y en parte mecánica. Cuando un sonido simple llega al oído, las neuronas primarias conectadas a las células sensoriales en la región de máxima amplitud de resonancia (es decir, fibras con similar frecuencia característica, pág. 65) aumentan su frecuencia de disparo por encima de su nivel de disparo espontáneo. Este aumento es una función monotónica de la amplitud del estímulo, pero no es una función lineal (por ej., Sachs y Abbas, 1974). Cuando la amplitud aumenta un factor de - digamos - 100 (un incremento de 40 db en el SPL, relación [3.16]), la frecuencia de disparo solo crece en un factor 3 o 4. Otro factor que contribuye a la compresión de sonoridad es el siguiente: a SPL's más altos, la frecuencia de disparo de una neurona inicial se *satura* a un nivel que es sólo unas pocas veces el nivel de disparo espontáneo. Cualquier aumento ulterior en la intensidad no alterará mayormente esta frecuencia de disparo; las neuronas simplemente no pueden transmitir a una frecuencia más alta que la determinada por su propio valor de saturación (determinado por el período refractario después de cada impulso). Las fibras nerviosas individuales del sistema auditivo de similar frecuencia característica, tienen muy distintos umbrales de disparo (el SPL necesario para acrecentar la frecuencia de disparo por encima del nivel espontáneo). El incremento en el estímulo requerido para alcanzar la saturación, también resulta variar entre 20 db y 40 db, o más, para diferentes fibras. A raíz de todo esto, el conjunto de neuronas que colectan información de un área determinada de la membrana basilar, cubren un amplio rango dinámico del sistema acústico a través de una apropiada «repartición de tareas». En un estudio sistemático de las características individuales de las neuronas del sistema auditivo, Liberman (1978) identificó tres grupos: fibras con alta frecuencia de disparo espontáneo (mayor a 20 impulsos por segundo) que tienen los umbrales más bajos; fibras con frecuencia espontánea de disparo muy baja (menor de 0,5 por segundo), con umbrales altos repartidos en un amplio rango (hasta 50 - 60 db); y un grupo con frecuencias de disparo espontáneo y umbrales intermedios. Cada grupo de fibras tiene características celulares distintas (diámetro y arquitectura del contacto sináptico con las células ciliadas internas). Lo interesante es que cada célula ciliada interna recibe fibras de los tres grupos. Esto da a cada célula ciliada individual la posibilidad de enviar su información en un amplio rango dinámico (y puede ser una razón más por la cual cada una de las células ciliadas esté en contacto con tantas fibras aferentes; véanse páginas 64 y 65). Según estos resultados, la intensidad del sonido está codificada tanto en la *frecuencia de disparo*, como en el *tipo* de fibra nerviosa auditiva que transmite la información.

Debemos añadir algo más. A medida que aumenta la intensidad de un

tono simple, la amplitud de la onda que se propaga a lo largo de la membrana basilar aumenta en todas partes, no solo en la región de resonancia máxima. Esto abre la posibilidad de que las neuronas que tienen una frecuencia característica *diferente* a la de la onda sonora (es decir, fibras en contacto con células ciliadas que no están en la zona de resonancia) también aumenten su frecuencia de disparo cuando sus umbrales (fuera de resonancia) son sobrepasados (véase Fig. 3.17). En resumen, un aumento de intensidad conduce a un aumento del número total de impulsos transmitidos, sea porque se ha incrementado la *frecuencia de disparo de cada neurona*, o bien el *número total de neuronas activadas*. Este efecto depende principalmente de la forma de la distribución de amplitud de oscilación a lo largo de la membrana, una propiedad puramente mecánica.

La relación entre sonoridad subjetiva y la totalidad de impulsos neurales explica cualitativamente las principales características de la suma de sonoridades (Sec. 3.4). Cuando se tienen dos tonos cuya separación en frecuencia es mayor que una banda crítica, el total de los impulsos nerviosos transmitidos es aproximadamente igual a la suma de los impulsos generados separadamente por cada tono; de aquí que la sonoridad total tienda a ser la suma de las *sonoridades* de cada uno de los tonos (relación [3.19]). Por otro lado, si los tonos tienen frecuencias dentro de una banda crítica, con zonas de resonancia superpuestas en la membrana basilar, el número total de impulsos estará determinado por la suma de las *intensidades* de los estímulos originales (relación [3.18]).

La dependencia de la sonoridad (y de los umbrales de enmascaramiento) de la duración del tono (Fig. 3.16) señala la existencia de un comienzo *gradual* del procesamiento de la señal acústica. Sólo después de varias décimas de segundo, el mecanismo alcanza un «estado estacionario» de procesamiento del sonido. Es importante volver a poner énfasis en que -afortunadamente para la música- durante esta etapa de transición inicial la sensación de altura es estable y única desde el comienzo.

Existe, sin embargo, una leve dependencia de la altura con la sonoridad, para tonos de frecuencia constante. Para frecuencias superiores a 2000 Hz, aproximadamente, la altura aumenta con la sonoridad y viceversa; por debajo de 1000 Hz ocurre lo contrario (por ej. Walliser, 1969). Este efecto es pequeño y varía mucho de persona a persona, necesiéndose grandes cambios en la intensidad para producir variaciones perceptibles en la altura¹⁵. Probablemente, este efecto es causado por la asimetría de la distribución de la excitación sobre la membrana basilar (ver Fig. 3.5), así como por la respuesta nerviosa no lineal debida a la saturación - un cambio en la intensidad produce un cambio del punto central de la

¹⁵ Experimentos con 'apareamiento de altura tonal' muestran que un aumento en el SPL de 40 db a 80 db incrementa la sensación de altura de un tono de 6000 Hz en un 7%, y disminuye la de un tono de 150 Hz en 3% (Walliser, 1969).

excitación (aun cuando la frecuencia quede constante), lo que a su vez conduce a un cambio en la sensación de altura primaria. Otra consecuencia de esta asimetría puede ser el efecto, pequeño pero no despreciable, que consiste en que la altura de un tono de frecuencia fija parece cambiar levemente cuando se superpone otro tono de diferente frecuencia (por ej., Walliser, 1969; Terhardt y Fastl, 1971). Este efecto puede tener consecuencias para la entonación musical¹⁶.

Aquí llegamos nuevamente a una cuestión, ya mencionada en la pág. 45, que tiene que ver con el mecanismo de detección de altura *primaria* o espectral: si un tono puro de intensidad y frecuencia determinadas pone a la membrana basilar en un estado de oscilación resonante, cubriendo un rango espacial finito Δx , y la sensación de altura primaria es codificada en términos de la posición x de las fibras nerviosas activadas, ¿cómo es que sólo se produce *una* única sensación de altura? Un mecanismo equivalente opera en el sistema visual, contribuyendo a mejorar la percepción de contrastes (Ratliff, 1972). En el sistema visual, este proceso de «agudización» es realizado por una red neural (en parte en la retina), cuya función es concentrar la actividad en un número limitado de neuronas rodeadas por un área de «quietud» neural o inhibición, mejorando de este modo la calidad del contraste. Hasta los años '70 se creía que un mecanismo análogo existía en el sistema auditivo, responsable de la agudización de la altura. Pero recientes estudios de la cóclea han revelado que las células pilosas externas poseen notables propiedades electromecánicas responsables de un proceso de realimentación no lineal que agudiza y amplifica la sintonización auditiva *antes* de que tenga lugar la conversión a señales nerviosas. De este proceso nos ocuparemos en la siguiente sección.

3.6 Música desde los oídos: emisiones otoacústicas y mecanismos cocleares

¿Música desde los oídos? Bueno... no tanto, ¡pero casi!

El estudio del mecanismo auditivo avanzó a través de tres épocas principales. La primera fue dominada por las ideas de von Helmholtz (1863) de que la membrana basilar funcionaba como un sistema mecánico analizador de espectros, capaz de responder al estímulo externo con la generación de ondas estacionarias (imagen de un «pequeño teclado resonante en el oído»).

¹⁶ Deberíamos señalar que también varía la altura percibida cuando cambia la presión del fluido coclear (por ej., las variaciones de altura percibidas durante un bostezo), o cuando cambia su composición química (por ej., inyecciones de droga en el fluido cerebroespinal). ¡Obsérvese que estos cambios de altura no pueden explicarse por medio de una teoría basada en claves temporales (Sec. 2.9)!

La segunda época, desde 1940 hasta la década del '70, es dominada por los resultados experimentales de von Békésy (1960), que mostraban que el sonido, al llegar al oído, genera una onda de propagación hidromecánica, confinada espacialmente, con máxima amplitud localizada en un área que depende de la frecuencia de la señal entrante (imagen de una «bandera que flamea», pág. 34). La época presente empezó en los '70, con un gran número de experimentos y estudios teóricos que demuestran que las ondas de von Békésy son localmente amplificadas por medio de un proceso electro-mecánico, en el cual las células ciliadas externas, a causa de su (hasta entonces insospechada) motilidad, funcionan tanto como elementos receptores, como de unidades de realimentación mecánica (para una reseña completa véase Dallos, 1992). Esta amplificación, ciclo por ciclo, funciona más eficientemente con señales de bajo nivel, lo que explica tanto la alta sensibilidad del oído, como su amplio rango dinámico de respuesta. Este proceso de realimentación puede producir una auto-oscilación, o resonar aun después que el estímulo externo haya cesado, y generar vibraciones cocleares en el dominio acústico que pueden ser captadas como tonos débiles por medio de micrófonos muy sensitivos ubicados dentro del canal del oído externo (Kemp, 1978).

Aunque estas *emisiones otoacústicas* no tienen consecuencias directas para la música, ponen en evidencia un proceso que puede explicar algunas de las sorprendentes, y en parte contradictorias, capacidades del sistema auditivo: detectar un sonido que desplaza los elementos receptores en sólo fracciones de nanómetro (10^{-9} m); ser sensitivo a un rango de intensidades que se extiende por un rango de mil millones a uno; responder a variaciones temporales del orden de unos pocos microsegundos; poseer una capacidad de resolución de frecuencia mucho mayor que la que puede proveer la membrana basilar en sí. Estos nuevos conocimientos ganados al nivel celular muestran a la cóclea como «un triunfo evolutivo de miniaturización... el aparato mecánico más complejo del cuerpo humano... un amplificador acústico y analizador de frecuencias compactados en el volumen de una bolita» (Hudspeth, 1985, 1989).

Las mediciones de von Békésy sobre el movimiento de la membrana basilar fueron realizadas en cócleas de animales muertos, conservadas en buen estado, revelando regiones anchas de resonancia en respuesta a tonos simples. Cuando fue posible registrar los impulsos de fibras nerviosas auditivas en animales vivos (véase Sec. 2.8), se notó que estas fibras estaban sintonizadas mucho más finamente¹⁷ a una frecuencia característica que la membrana basilar, indicando la presencia de un proceso de 'estrechamiento' o 'afinamiento', que en esa época se creyó

¹⁷ «Afinación fina» significa que la banda de frecuencias a la que responde una neurona es muy estrecha; «afinación gruesa» significaría que la respuesta ocurre para un rango amplio de frecuencias.

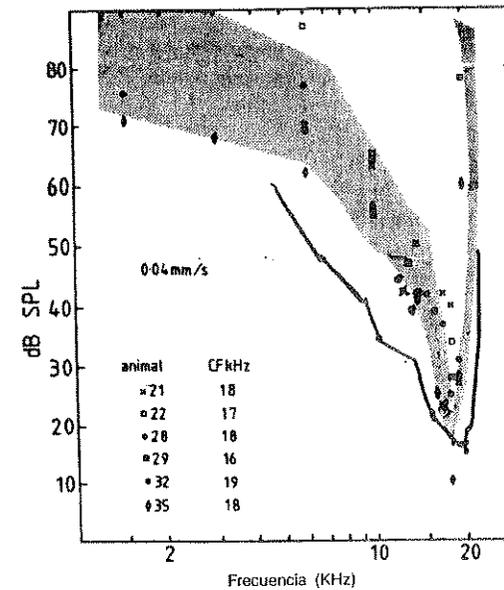


Figura 3.17 Comparación entre la membrana basilar (curva gruesa) y la sintonización neural. El área gris corresponde a las curvas de sintonización de diez nervios acústicos ligados a la misma región de la membrana basilar (con autorización, Johnstone et al., 1983).

ser obra del sistema nervioso (como ocurre en la visión). Mediciones más recientes, usando el efecto Mössbauer y técnicas de láser en animales vivos, revelaron una 'sintonización' mucho más 'fina' de la membrana basilar misma. La figura 3.17 compara la membrana basilar con la 'sintonización' nerviosa (Johnstone et al., 1983). La curva gruesa representa el SPL de una onda sonora incidente, de frecuencia dada, para la cual la membrana basilar responde con una *velocidad pico determinada* (0,04 mm/s en la figura) en un *punto determinado*, en el cual está colocada la fuente radiactiva (ésta curva se llama curva de sintonización de isovelocidad). El mínimo profundo a 18 - 19 KHz indica que en ese punto la membrana basilar tiene una respuesta máxima a señales de dicha frecuencia (solo se necesitan 20 db de SPL para producir esa velocidad pico). Una octava más baja, o sea a 9 KHz, se necesitan 15 db más en la señal para producir la misma velocidad de respuesta. Una curva de sintonización *neural* (el área gris define los límites de diez curvas de sintonización normales) representa, por otra parte, el *mínimo* SPL de una señal de frecuencia determinada al cual responde una fibra dada (por ej., incrementando su frecuencia de disparo por encima del umbral). Existe mucho más similitud entre la 'sintonización' neural y la de la membrana basilar, que lo previamente sospechado en base a las antiguas mediciones

con cócleas de animales muertos. El uso de microelectrodos ha mostrado que las curvas de sintonización de las células pilosas internas son muy similares a aquellas de las fibras nerviosas aferentes; por lo tanto, el proceso de afinamiento en cócleas vivas debe ocurrir dentro de la cóclea, probablemente en el espacio subreticular (Fig. 2.7[a]).

Con nuevas técnicas Mössbauer y de láser muy perfeccionadas, es posible hoy día realizar mediciones mucho más minuciosas del movimiento de la membrana basilar en animales vivos. (por ej., Johnstone et al., 1986; Ruggiero y Rich, 1991). Los resultados muestran que, en condiciones reales, las curvas de sintonización de la membrana basilar son muy similares a aquellas de las células ciliadas internas o de las fibras nerviosas asociadas al mismo lugar de la membrana. Lo que es más, las mediciones evidencian que la agudeza de la curva de sintonización de la membrana basilar aumenta considerablemente a medida que el SPL del estímulo *disminuye* hacia el umbral; esto indica la operación de un mecanismo de amplificación no lineal capaz de robustecer, en el caso de señales débiles, más de 10.000 veces la oscilación de la membrana basilar en la zona de resonancia (el mecanismo es «no lineal» porque no incrementa la sensibilidad en la misma proporción para todas las frecuencias y amplitudes).

Estudios técnicos muestran que la energía suministrada al mecanismo de amplificación no puede proceder de la fuente sonora original (por ej. de Boer, 1983). La sensibilidad de la membrana basilar, su fina selectividad a la frecuencia y sus propiedades de respuesta no lineal deben provenir de la acción de un mecanismo de realimentación de las células pilosas externas. Un hecho convincente es que estas células, bajo estimulación eléctrica, son capaces de *cambiar de forma* - principalmente su longitud - a frecuencias de audio (Kachar et al., 1986). Es concebible que estos cambios de forma afecten la deformación local de la partición coclear inducida por la onda incidente, por medio de una realimentación de energía que haga que la onda resultante sea amplificadora (Dallios, 1992). La motilidad de las células ciliadas externas se produce por la acción concertada de moléculas proteicas contractiles individuales que se encuentran en la membrana de la célula; la velocidad de su acción es tan increíblemente alta que se pueden excluir mecanismos electromecánicos similares a los encontrados en las fibras musculares normales. Mientras está demostrado que el mecanismo de realimentación coclear es disparado localmente por el mecanismo de activación de las células ciliadas externas, mismas, existe la posibilidad que este proceso pueda ser modulado por órdenes procedentes del sistema nervioso central, por vía de las fibras eferentes del fascículo olivo-coclear (pág. 77). Finalmente, dada la saturación en la respuesta de las células ciliadas externas para altos niveles de estimulación, la acción de realimentación está limitada a niveles bajos del estímulo externo, dentro de los 40 db por encima del umbral.

En resumen, el cuadro que emerge de este proceso de realimentación no lineal que amplifica la respuesta mecánica de la cóclea (Dallios, 1992; Allen y Neely, 1992) es, para sonidos *simples*, el siguiente: 1) La oscilación sinusoidal del «piston» en la ventana redonda (Fig. 2.6 [b]) genera una onda de propagación en la membrana basilar, la cual alcanza su amplitud máxima en una posición que depende de la frecuencia. 2) La oscilación local es detectada por las ciliadas de las células ciliadas. Este proceso hace que se dispersen señales eléctricas en cada célula ciliada externa, causando la contracción de moléculas motoras en su membrana. Este proceso puede ser influenciado por órdenes procedentes del cerebro que llegan por las fibras eferentes del complejo olivo-coclear. 3) El resultado colectivo de las respuestas de las moléculas motoras es un cambio en la forma de la célula (probablemente su largo), cambio sincronizado con la oscilación producida por el estímulo externo. 4) Esta reacción mecánica en fase refuerza la oscilación local de la membrana basilar (en forma hidrodinámica o por medio de un «raqueteo» [«rattling»] sobre la membrana reticular.) 5) Si se alcanza un cierto límite superior, la reacción celular se nivela, se satura; de lo contrario... ¡volver al paso 2!

Pregunta: ¿Por qué actúa este proceso sólo en la región basilar de máxima resonancia para la frecuencia del estímulo? Después de todo, alejándonos del lugar de máxima amplitud de la onda, seguimos encontrando oscilación, si bien con amplitudes cada vez más pequeñas (Fig. 3.5). En otras palabras ¿por qué no se amplifica el patrón vibratorio *entero*, a lo largo de toda la membrana basilar? ¡Podría estar operando un segundo mecanismo de sintonización! Efectivamente, existen indicios experimentales de que las *células mismas* están afinadas a la frecuencia de resonancia *local*; de ahí que ellas respondan de manera eficiente sólo cuando la frecuencia de la onda que pasa corresponde a la frecuencia de resonancia *local*; esto ocurre *sólo* alrededor de la máxima resonancia para la frecuencia del estímulo (pág. 89 y Fig. 3.5).¹⁸

Otra cuestión difícil de desentrañar se relaciona con los efectos de ruido térmico. Una cilia oscila en su base (como sobre un pivote) y cuando se mueve en una dirección específica produce una respuesta eléctrica excitatoria en la célula en que está anclada (e inhibitoria si se mueve en la dirección contraria, mientras que las desviaciones ortogonales son ineficaces - ¡esto explica los disparos en fase de una fibra nerviosa auditiva, mostrada en la Fig. 2.22!). Lo sorprendente es que el umbral de la

¹⁸ Investigaciones electrofisiológicas muestran que las frecuencias a las cuales las células ciliadas son más sensibles están en relación inversa a la longitud de sus paquetes de cilia, y ocurre que esta longitud aumenta varias veces a lo largo de la membrana basilar, desde la base (extremo de frecuencias altas) hasta el ápex (extremo de frecuencias bajas). Existen también algunos indicios de que una afinación ulterior puede ocurrir por acción de una resonancia eléctrica en la membrana de la célula ciliada (Hudspeth, 1985). Todo esto se encontró en cócleas de vertebrados inferiores; todavía no queda claro en qué medida estos resultados valen también para mamíferos superiores y seres humanos.

sensación auditiva ocurre para desvíos de las cilias de sólo 0,003 grados! Puede demostrarse que desplazamientos de tal pequeñez deberían ocurrir aun en ausencia de *cualquier* estimulación acústica, debido al así llamado movimiento Browniano en el fluido endolinfático (pequeños elementos de fluido moviéndose a causa de las fluctuaciones del equilibrio térmico). ¡Pero el ruido resultante es mitigado por el hecho que la cilias están ligadas entre sí por filamentos, de espesor muscular, que las hacen mover en conjunto y en forma concertada! Con todo, el límite último de la audición está determinado por el punto en el cual los más débiles sonidos son ahogados por el ruido térmico de los componentes del oído (Hudspeth, 1989).

¿Pueden todos los fascinantes procesos arriba mencionados ocurrir en una escala de tiempo del orden de fracciones de microsegundo, tal como lo requiere el increíble poder de resolución temporal del sistema acústico? En otras palabras, ¿puede la maquinaria electromecánica de la cóclea ser *mejor* que la maquinaria nerviosa (en términos de eficacia temporal) ? La generación de impulsos eléctricos en una célula ciliada está basada en el intercambio de iones (calcio y potasio) con el medio circundante. Hay experimentos que demuestran la existencia de poros en la membrana a través de los cuales los iones pasan cuando las cilias se tuercen en una determinada dirección (Hudspeth, 1985). El pulso de voltaje generado por la corriente eléctrica llevada por los iones es responsable de la respuesta integral de la célula; para el caso de una célula ciliada externa esto incluye la contracción de las moléculas motoras, lo cual, efectivamente, puede acontecer en el margen de un microsegundo a partir de la aplicación del estímulo. Queda ahora aclarado que para satisfacer las muy exigentes demandas de rapidez del sistema auditivo, las operaciones de detección y amplificación de señal deben ser manejadas por procesos que ocurran *dentro* de una célula: ¡el aparato auditivo no puede darse el lujo del «paso tranquilo» del sistema nervioso, que trabaja en una escala de tiempo del orden de los milisegundos o más!

Para concluir, volvamos a la «música desde los oídos». Como dijimos antes, las emisiones otoacústicas son sonidos muy débiles que pueden ser detectados en el canal cerrado del oído; su nivel está generalmente muy por debajo del umbral de la sensación auditiva. Hay dos tipos de emisiones: emisiones espontáneas y emisiones evocadas por estímulos sonoros externos. La figura 3.18 muestra el espectro de una emisión otoacústica espontánea (Zwicker y Fastl, 1990) (cero SPL corresponde al umbral de audición). Las emisiones espontáneas varían mucho de individuo a individuo. Alrededor del cincuenta por ciento de las personas con audición normal tienen una o varias emisiones; la mayoría de los animales las tienen también. La ausencia de las mismas *no* indica anomalía en la audición (el «zumbido» de oídos en general no tiene relación con estas emisiones). El rango de sus frecuencias se extiende de 400 a 4.000 Hz,

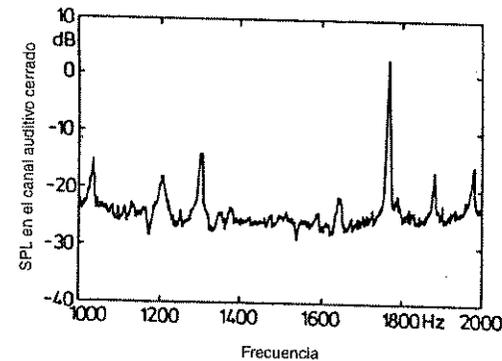


Figura 3.18 Ejemplo de un espectro de frecuencia de emisiones otoacústicas (Zwicker y Fastl, 1990), detectadas por un micrófono muy sensible insertado herméticamente en el canal del oído externo. Cero SPL corresponde al umbral de audición para la frecuencia dada.

aproximadamente, aunque la mayor parte de ellas ocurre entre los 1.000 y los 2.000 Hz; para una persona dada, su intensidad puede variar considerablemente en el curso de un día, e incluso ser intermitente, pero las frecuencias a las que ocurren son muy estables para cada cóclea. Las emisiones otoacústicas evocadas son más difíciles de medir; requieren un transmisor de sonido en el canal auditivo, y las más exigentes condiciones de análisis de la señal detectada por el micrófono. Pueden dividirse en dos clases: emisiones que ocurren simultáneamente con un tono de prueba continuo, y respuestas demoradas a cortos impulsos sonoros. No hay duda de que las emisiones otoacústicas son un subproducto de los procesos antes mencionados de realimentación no lineal, controlados por las células ciliadas externas; *por qué* ocurren cuándo ocurren, y por qué aparecen a las frecuencias indicadas, aún no se sabe. Lo único que podemos observar es que circuitos de realimentación tan delicados, al ser activados, tienen en general una capacidad de autoexcitación o de resonancia post-estímulo. Dado que las emisiones otoacústicas parecen no tener ningún rol en la música (¡salvo que se reprodujeran en forma muy amplificadas!) remitimos al lector interesado en obtener más información a la bibliografía técnica (por ej., Zwicker y Fastl, 1990; Gelfand, 1990, y referencias ahí dadas).

4

Generación de sonidos musicales, tonos compuestos y percepción del timbre

«... no es sorprendente que haya sido posible posar un hombre en la Luna antes de que la acústica de un instrumento tradicional como el piano fuera totalmente explicada».

A. Askenfelt, en *The Acoustics of the Piano*, publicación de la Royal Swedish Academy of Music, 1990.

En los dos capítulos precedentes hemos analizado dos atributos principales de un tono, la altura y la sonoridad, principalmente sobre la base de sonidos simples, de una sola frecuencia. Estos no son, ¡por suerte!, los tonos que intervienen en la música. La música está formada por tonos *compuestos*, cada uno de los cuales consiste de una superposición de tonos puros, los cuales, debido a sus relaciones de frecuencia, aparecen como un único ente perceptual. Así surge un tercer atributo tonal fundamental: la cualidad del tono, o timbre, relacionado con el tipo de mezcla de sonidos puros, o *componentes armónicos*, que forman el tono compuesto (Sec. 1.2).

La mayor parte de los instrumentos musicales genera ondas sonoras por medio de cuerdas o columnas de aire en vibración. En el capítulo 1 los denominamos elementos primarios de vibración. La energía necesaria para sostener su vibración es suministrada por un mecanismo de excitación, en tanto que en muchos instrumentos la energía acústica final irradiada está controlada por un resonador. La habitación o sala de concierto en la cual se está tocando el instrumento puede considerarse como una «extensión» natural del instrumento mismo, cumpliendo una función sustancial en la modelación del sonido que llega a los oídos del oyente.

En este capítulo discutiremos cómo tonos musicales reales son generados en los instrumentos musicales, cómo están formados por superposiciones de tonos simples, cómo interactúan con el entorno en salas de concierto y otros recintos, y cómo todo esto conduce a la percepción del timbre y al reconocimiento del instrumento. El capítulo termina con una breve reseña de las funciones cognitivas del cerebro relevantes a la percepción de tonos compuestos.

4.1 Ondas estacionarias en una cuerda

Consideremos el caso de una cuerda bajo tensión, fija en los puntos P y Q (Fig. 4.1), de longitud L y masa por unidad de longitud d . La fuerza de tensión T puede ser modificada a gusto, por ejemplo, cambiando la masa m del cuerpo suspendido, tal como se muestra en la figura. Ahora punteamos o percutimos la cuerda en un punto determinado. Dos ondas elásticas transversales (dos impulsos) se propagarán hacia la derecha y hacia la izquierda, alejándose de la región de la perturbación inicial con una velocidad dada por la relación (3.3). Cuando alcancen los puntos fijos P y Q , estos impulsos de onda serán así reflejados: un impulso positivo «hacia arriba» volverá como un impulso negativo «hacia abajo», y viceversa. Después de cierto tiempo (extremadamente corto, en virtud de la alta velocidad de las ondas en una cuerda tensa) habrá ondas propagándose en ambas direcciones a lo largo de la cuerda. En otras palabras, tendremos energía elástica en forma de ondas «atrapadas» entre P y Q , dos puntos que siempre permanecen en reposo. Si no hubiera pérdidas, esta situación se mantendría invariable y la cuerda continuaría vibrando indefinidamente. Sin embargo, la fricción y las pérdidas en P y Q irán disipando gradualmente la energía almacenada, y las ondas decaerán¹. Por ahora ignoraremos este efecto.

En vista de lo expuesto en la Sec. 3.3, nos damos cuenta que el proceso descrito arriba se asemeja mucho a la situación que se presenta con una onda estacionaria. En realidad, puede demostrarse matemáticamente que *las ondas estacionarias son el único modo posible de vibración estable para una cuerda con extremos fijos, y en el que los puntos de anclaje P y Q asumen el rol de nodos*.

Esto tiene una consecuencia muy importante. Entre todas las formas de ondas estacionarias imaginables, sólo son posibles aquellas para las cuales

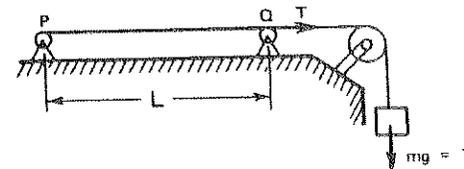


Figura 4.1 Instrumento (sonómetro) para el estudio de ondas estacionarias en una cuerda de longitud L , bajo tensión controlable.

¹Es precisamente esta pérdida de energía por unidad de tiempo a través de los puntos fijos (principalmente en el puente), la que se transforma en potencia sonora en el cuerpo de resonancia de un instrumento de cuerda.

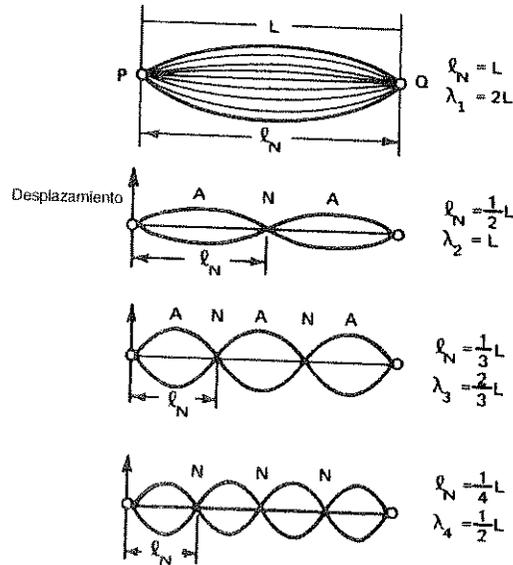


Figura 4.2 Modos de vibración en una cuerda con extremos fijos (ondas estacionarias).

ocurren nodos en P y Q . En otras palabras, únicamente son permitidas aquellas ondas sinusoidales que «entran» un número entero de veces entre P y Q (Fig. 4.2), es decir, para las cuales la longitud L de la cuerda es un múltiplo entero de la distancia entre nodos l_N , dada por la relación (3.13). Teniendo en cuenta esta relación, obtenemos la condición $L = nl_N = n\lambda/2$, donde n es cualquier número entero 1, 2, 3, ... Esto nos dice que las únicas longitudes de onda permitidas (Fig. 4.2) son

$$\lambda_n = \frac{2L}{n} \quad n = 1, 2, 3, \dots \quad (4.1)$$

Usando la relación (3.8), encontramos que las únicas frecuencias de vibración posibles de una cuerda son las siguientes:

$$f_n = \frac{1}{\lambda_n} \sqrt{\frac{T}{d}} = \frac{n}{2L} \sqrt{\frac{T}{d}} = nf_1 \quad (4.2)$$

La frecuencia más baja posible f_1 se obtiene para $n = 1$:

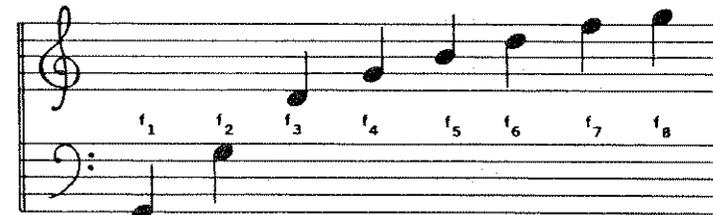
$$f_1 = \frac{1}{2L} \sqrt{\frac{T}{d}} \quad (4.3)$$

Esta es la llamada *frecuencia fundamental* de la cuerda vibrante. Obsérvese en la ecuación (4.2) que la frecuencia de cualquier otra vibración posible

es un múltiplo entero de la frecuencia fundamental. Estos son los llamados armónicos superiores de f_1 (Sec. 2.7). Nótese en particular que el primer armónico ($n = 1$) es idéntico a la frecuencia fundamental; el segundo armónico f_2 es la octava superior de f_1 ; el tercer armónico es la duodécima (una quinta arriba de la octava); el cuarto armónico es la décimoquinta (doble octava); etc. (Fig. 4.3). Los armónicos superiores son llamados también tonos *parciales*².

La relación (4.3) nos dice que la frecuencia fundamental de oscilación de una cuerda es proporcional a la raíz cuadrada de la tensión, e inversamente proporcional a su longitud y a la raíz cuadrada de su masa por unidad de longitud. Esto explica muchas propiedades características de las cuerdas de piano: en la parte superior del teclado, las cuerdas son cada vez más cortas (frecuencia fundamental f_1 cada vez más alta); si tenemos que afinar una cuerda determinada un poco más agudo, es necesario aumentar la tensión (f_1 más alta) y recíprocamente; en la región de alturas graves, con el fin de ahorrar espacio y maximizar la potencia de salida, en lugar de acrecentar la longitud de una cuerda se incrementa su masa por unidad de longitud d (f_1 más baja) rodeando la cuerda con un rollo de alambre adicional. En el violín, que sólo tiene cuatro cuerdas de casi igual longitud, cada una debe tener diferente tensión y/o masa para que cada una pueda exhibir una altura propia diferente. Para variar la frecuencia fundamental f_1 de una cuerda determinada, se cambia su longitud vibrante L presionando la cuerda contra la tastiera, introduciendo de este modo un nodo en el punto de contacto.

La aparición natural de frecuencias discretas relacionadas en forma simple con una frecuencia fundamental, determinada por las condiciones del sistema, con todas las otras frecuencias «prohibidas», recibe el nombre de «cuantización» y tiene un rol importantísimo en toda la física. Las



Primeros ocho armónicos de $f_1 = 98$ Hz (Sol)

Figura 4.3 Primeros ocho armónicos de $f_1 = 98$ Hz (Sol).

²Más precisamente, los *parciales* son los componentes de frecuencias más altas de una vibración compleja, *sin importar* si sus frecuencias son o no múltiplos enteros de la frecuencia fundamental.

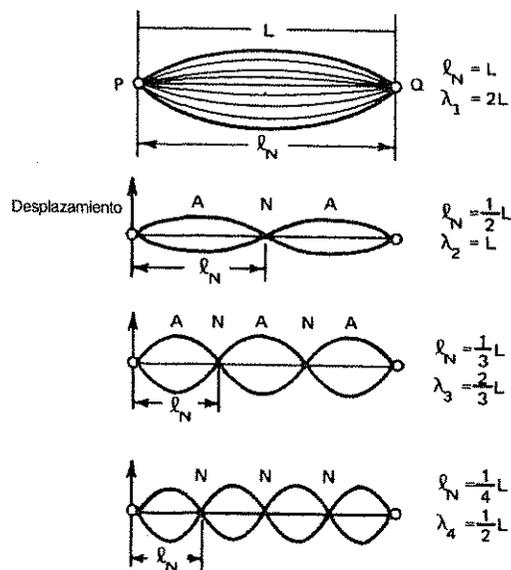


Figura 4.2 Modos de vibración en una cuerda con extremos fijos (ondas estacionarias).

ocurren nodos en P y Q . En otras palabras, únicamente son permitidas aquellas ondas sinusoidales que «entran» un número entero de veces entre P y Q (Fig. 4.2), es decir, para las cuales la longitud L de la cuerda es un múltiplo entero de la distancia entre nodos l_N , dada por la relación (3.13). Teniendo en cuenta esta relación, obtenemos la condición $L = nl_N = n\lambda/2$, donde n es cualquier número entero 1, 2, 3, ... Esto nos dice que las únicas longitudes de onda permitidas (Fig. 4.2) son

$$\lambda_n = \frac{2L}{n} \quad n = 1, 2, 3, \dots \quad (4.1)$$

Usando la relación (3.8), encontramos que las únicas frecuencias de vibración posibles de una cuerda son las siguientes:

$$f_n = \frac{1}{\lambda_n} \sqrt{\frac{T}{d}} = \frac{n}{2L} \sqrt{\frac{T}{d}} = nf_1 \quad (4.2)$$

La frecuencia más baja posible f_1 se obtiene para $n = 1$:

$$f_1 = \frac{1}{2L} \sqrt{\frac{T}{d}} \quad (4.3)$$

Esta es la llamada *frecuencia fundamental* de la cuerda vibrante. Obsérvese en la ecuación (4.2) que la frecuencia de cualquier otra vibración posible

es un múltiplo entero de la frecuencia fundamental. Estos son los llamados armónicos superiores de f_1 (Sec. 2.7). Nótese en particular que el primer armónico ($n = 1$) es idéntico a la frecuencia fundamental; el segundo armónico f_2 es la octava superior de f_1 ; el tercer armónico es la duodécima (una quinta arriba de la octava); el cuarto armónico es la décimoquinta (doble octava); etc. (Fig. 4.3). Los armónicos superiores son llamados también tonos *parciales*².

La relación (4.3) nos dice que la frecuencia fundamental de oscilación de una cuerda es proporcional a la raíz cuadrada de la tensión, e inversamente proporcional a su longitud y a la raíz cuadrada de su masa por unidad de longitud. Esto explica muchas propiedades características de las cuerdas de piano: en la parte superior del teclado, las cuerdas son cada vez más cortas (frecuencia fundamental f_1 cada vez más alta); si tenemos que afinar una cuerda determinada un poco más agudo, es necesario aumentar la tensión (f_1 más alta) y recíprocamente; en la región de alturas graves, con el fin de ahorrar espacio y maximizar la potencia de salida, en lugar de acrecentar la longitud de una cuerda se incrementa su masa por unidad de longitud d (f_1 más baja) rodeando la cuerda con un rollo de alambre adicional. En el violín, que sólo tiene cuatro cuerdas de casi igual longitud, cada una debe tener diferente tensión y/o masa para que cada una pueda exhibir una altura propia diferente. Para variar la frecuencia fundamental f_1 de una cuerda determinada, se cambia su longitud vibrante L presionando la cuerda contra la tastiera, introduciendo de este modo un nodo en el punto de contacto.

La aparición natural de frecuencias discretas relacionadas en forma simple con una frecuencia fundamental, determinada por las condiciones del sistema, con todas las otras frecuencias «prohibidas», recibe el nombre de «cuantización» y tiene un rol importantísimo en toda la física. Las

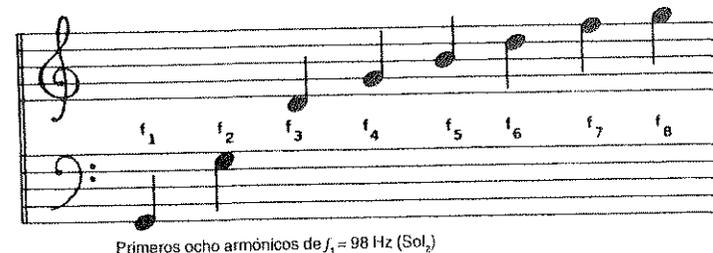


Figura 4.3 Primeros ocho armónicos de $f_1 = 98$ Hz (Sol_2).

² Más precisamente, los *parciales* son los componentes de frecuencias más altas de una vibración compleja, *sin importar* si sus frecuencias son o no múltiplos enteros de la frecuencia fundamental.

diferentes formas discretas, permitidas, de vibración de un sistema físico, se llaman *modos* de vibración de ese sistema. La fundamental, la octava, la duodécima, etc., representan el primer, segundo, tercer, etc., modo de vibración de una cuerda bajo tensión. Las frecuencias que corresponden a todos los modos posibles de vibración de una cuerda están dadas por la relación (4.2), en la que solo aparecen cantidades que dependen de la cuerda; *los modos de vibración son así una característica permanente del sistema físico particular considerado*. Dados los modos posibles de vibración, ¿en cuáles de ellos vibrará la cuerda *realmente*? Esto estará determinado por la forma en que las vibraciones sean iniciadas, es decir, por el mecanismo primario de excitación. A causa de la propiedad de superposición lineal de las ondas, *muchos modos de vibración pueden ocurrir simultáneamente sin molestarse entre sí*. En esta sección centraremos nuestra atención sobre cómo una cuerda *puede vibrar*; en la próxima sección discutiremos la cuestión de cómo realmente *vibrará* en un caso dado.

Consideremos el caso de una cuerda vibrante en la cual sólo *un* modo de vibración está excitado. Esto puede ser fácilmente realizado en un laboratorio, pasando una corriente eléctrica alterna de determinada frecuencia por una cuerda metálica extendida entre los polos de un imán permanente. Las fuerzas magnéticas sobre la corriente en la cuerda producirán una vibración transversal a la frecuencia de la corriente. Cuando esta frecuencia se acerca a uno de los armónicos de la cuerda, aparece una onda estacionaria de gran amplitud. Pueden observarse a simple vista los nodos y los vientres (como en la Fig. 4.2), y oírse claramente el sonido producido (siempre y cuando la cuerda esté montada sobre una caja resonante). El uso de un estroboscopio (fuente luminosa que pulsa a una frecuencia determinada, controlable) da la posibilidad de «congelar» la forma de la cuerda, o de observarla en «cámara lenta».

Un experimento más accesible, y muy conocido, puede hacerse con un piano. Baje lentamente una tecla en la región grave, por ejemplo el Sol_2 (Fig. 4.3), de manera que no produzca ningún sonido (el martillo no golpea) pero que la sordina quede separada de esa cuerda. Luego toque con fuerza y staccato la tecla una octava arriba (Sol_3). Una vez que este sonido haya desaparecido, usted podrá escuchar claramente la cuerda Sol_2 vibrando una octava más alta: ¡ha sido excitada (por resonancia) en su segundo modo armónico, Sol_3 ! Repitamos el experimento tocando la duodécima (Re_4), mientras mantenemos baja la tecla Sol_2 : usted podrá escuchar la cuerda Sol_2 vibrando en Re_4 . Continúe con Sol_4 , Si_4 y así sucesivamente. Como otra prueba, toque La_3 o Fa_3 mientras mantiene baja la tecla Sol_2 - no habrá ningún efecto: la cuerda Sol_2 queda en silencio. La razón es que La_3 o Fa_3 no son armónicos superiores de Sol_2 , y la cuerda Sol_2 simplemente no puede sostener vibraciones estables a esas frecuencias.

La relación (4.2) es en realidad sólo aproximada, especialmente para

los modos de vibración de orden alto. La razón es que la velocidad de una onda transversal en una cuerda *depende* levemente de la frecuencia (o de la longitud de onda) de la onda (esto se llama *dispersión*), y las expresiones (3.3) y (3.8) no son enteramente correctas. En realidad, la velocidad V_T de la onda es un poco mayor que el valor dado por la relación (3.3). Esta desviación aumenta a medida que aumenta la distorsión de la cuerda; es decir, se hace más importante para longitudes de onda pequeñas y amplitudes mayores. El resultado es que las frecuencias de los modos de vibración más altos en una cuerda de piano son un poco más agudos que los valores dados por la ecuación (4.2).³ En general, cuando la frecuencia de un modo de vibración de un sistema no es un múltiplo entero de la frecuencia fundamental, el modo se designa *inarmónico*. Cuerpos sólidos en vibración como, por ejemplo, barras de xilófono, campanas o carillones, tienen muchos modos de vibración inarmónicos, cuyas frecuencias no son en absoluto múltiplos enteros de la frecuencia fundamental. Por razones de simplicidad, en lo que sigue supondremos que, en general, los parciales de una cuerda en vibración siempre coinciden con los armónicos superiores, y que la relación (4.2) es exacta. Por ello, usaremos de manera indistinta los términos «armónicos superiores», «modos», o «parciales», aunque desde el punto de vista físico ellos representan conceptos distintos en el caso de las cuerdas reales.

4.2 Generación de ondas estacionarias complejas en instrumentos de cuerda

Existen dos formas fundamentales de excitar una vibración en una cuerda tensa: 1) un suministro de energía dado una única vez por medio de la acción de un golpe (piano) o de una pulsación (clavicémbalo, guitarra); y 2) un suministro de energía continuo por medio de una acción de frotamiento (instrumentos de la familia del violín). En ambos casos el resultado es una *superposición de muchos modos vibratorios activados simultáneamente*. En otras palabras, los sonidos musicales generados en forma natural en las cuerdas contienen muchas frecuencias diferentes al mismo tiempo: aquellas que corresponden a los armónicos del sistema en vibración. La figura 4.4 muestra cómo esto puede ocurrir en la práctica: sumando el primer y, digamos, el tercer armónico, se obtiene como resultado una superposición que en un instante dado puede tener la forma de la figura. Cada modo se comporta independientemente, y la forma de la cuerda en cualquier instante estará dada por la superposición (suma)

³ Esto no afecta mayormente la altura resultante del tono compuesto (véase Apéndice II), pero sí influencia la afinación del piano en los registros agudo y grave (cuando la afinación es hecha por octavas).

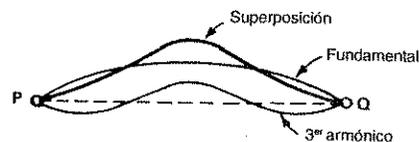


Figura 4.4 Superposición de dos ondas estacionarias (fundamental y tercer armónico).

de los desplazamientos individuales, dictados por cada componente separadamente. Es posible utilizar las condiciones iniciales del experimento mencionado anteriormente, en que teníamos una cuerda metálica con corriente eléctrica en un campo magnético, pero esta vez usando la salida combinada de *dos* generadores sinusoidales de voltaje cuyas frecuencias correspondan respectivamente a la de dos armónicos de la cuerda. Empleando un estroboscopio es posible visualizar con claridad la forma instantánea de la cuerda mientras vibra con los dos modos al mismo tiempo. La proporción relativa con que cada armónico interviene en la vibración resultante, determina (Secs. 1.2 y 4.8) el carácter particular, la cualidad o timbre del tono generado. La altura del tono compuesto de la cuerda está determinada por la frecuencia fundamental (4.3), como se ha anticipado en la Sec. 2.7.

Un simple experimento con el piano muestra de manera convincente que una cuerda efectivamente *puede* vibrar en más de un modo a la vez. Baje lentamente, y mantenga apretada, una tecla determinada (por ejemplo *Sol*₂) (Fig. 4.3) de manera de separar el apagador de la cuerda. Luego toque con fuerza y staccato *simultáneamente* *Re*₄, *Sol*₄, *Si*₄. Después que este sonido desaparece, es posible escuchar con claridad la cuerda *Sol*₂ vibrar en los tres modos, simultáneamente - ¡tenemos una sola cuerda sonando en una tríada de *Sol Mayor* completa! Lo que ocurre es que han sido excitados tres modos (armónicos tercero, cuarto y quinto) con amplitudes aproximadamente iguales (por resonancia). Un experimento más drástico es el siguiente: mantenga la tecla *Sol*₂ baja - y golpee con su antebrazo derecho todas las teclas, negras y blancas, de dos o más octavas arriba del *Sol*₃ - después que haya decaído el estampido inicial, se escucha la cuerda *Sol*₂ vibrar bellamente en un acorde de séptima de dominante *Sol*₃, *Re*₄, *Sol*₄, *Si*₄, *Re*₅, *Fa*₅, *Sol*₅, ... (Fig. 4.3). Cualquier otra de las notas tocadas no pudo excitar una vibración estable en la cuerda *Sol*₂.

Mientras que el experimento de arriba muestra que una cuerda de piano determinada *puede* vibrar simultáneamente en diferentes modos al mismo tiempo, el siguiente experimento muestra que, cuando se toca normalmente, una cuerda de piano realmente *vibra* en muchos modos armónicos a la vez. Elija de nuevo una nota baja, digamos *Sol*₂, pero esta vez baje lentamente la tecla *Sol*₃ sin hacerla sonar y manténgala baja. Luego toque fuerte y staccato *Sol*₂. La cuerda *Sol*₃ empezará a vibrar en

su propio modo fundamental de vibración (es decir, *Sol*₃). La razón es que este modo ha sido excitado (por resonancia) por el *segundo* armónico de la cuerda vibrante *Sol*₂. Si en lugar de *Sol*₂ se hubiera tocado *La*₂, la cuerda *Sol*₃ hubiera permanecido en silencio. Repita el experimento varias veces, bajando silenciosamente y en forma sucesiva las teclas *Re*₄, *Sol*₄, *Si*₄, *Re*₅, ... etc. Cada una de ellas será excitada por el correspondiente modo armónico superior de la cuerda *Sol*₂.⁴

Muchos modos de vibración aparecen juntos cuando una cuerda es puesta en vibración. ¿Qué es lo que determina *cuáles*, y *cuánto* de cada uno? Esto, en principio, está dado por la forma específica en que la cuerda es puesta en vibración, es decir, por el mecanismo primario de excitación. De cómo y dónde golpeemos, punteemos o frotamos la cuerda, depende que obtengamos diferentes mezclas de armónicos y, por ende, diferentes cualidades del sonido resultante. Podemos explicar esto basándonos en los siguientes ejemplos. Supongamos que le damos a una cuerda la forma inicial mostrada en la Fig. 4.5 (a) (esto, sin embargo, sería bastante difícil de lograr en la práctica). Dado que la forma se ajusta más o menos al modo fundamental (Fig. 4.2), la cuerda, cuando sea soltada, efectivamente empezará a vibrar en ese modo simple. Si la forma inicial fuera la mostrada en la Fig. 4.5(b), la cuerda, al ser soltada, vibraría en el tercer modo (Fig. 4.2). ¿Pero qué ocurrirá cuando la forma inicial sea aquella (más realista) mostrada en la Fig. 4.6, y que se asemeja a la forma inicial cuando punteamos la cuerda en el punto medio *A* entre *P* y *Q*? Para averiguarlo, superpongamos - es decir, sumemos linealmente - los casos de las figuras 4.5(a) y 4.5(b). Lo que obtenemos es la forma mostrada en la Fig. 4.7(a),

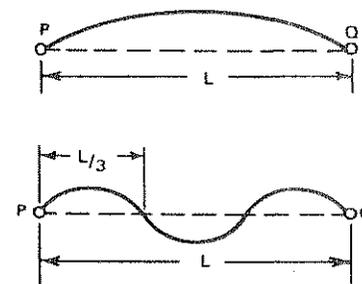


Figura 4.5 Forma inicial (deformación) de una cuerda requerida para que vibre simultáneamente en el modo fundamental (a) y en el tercer armónico (b).

⁴El principal objetivo del mecanismo del *pedal* del piano se basa precisamente en este fenómeno: el apretar el pedal levanta *todos* los apagadores, dejando las cuerdas libres para vibrar por resonancia. Cuando se toca una nota determinada, todas aquellas cuerdas que pertenecen a la serie de armónicos de esa nota serán inducidas a vibrar.

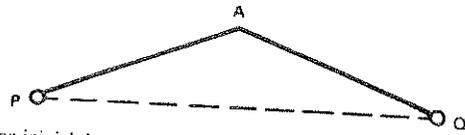


Figura 4.6 Forma inicial de una cuerda cuando se puntea en el punto A.

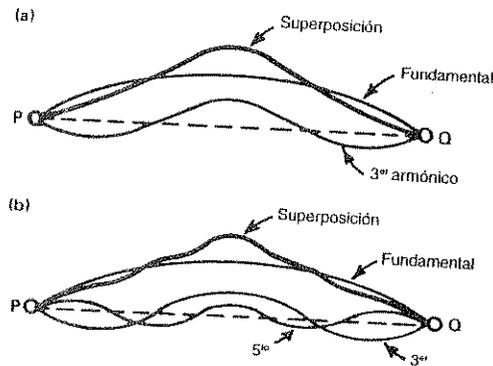


Figura 4.7 Superposición de dos (a) y tres (b) armónicos, elegidos como primera aproximación de la forma triangular mostrada en la Fig. 4.6.

que concuerda bastante bien con la forma inicial de la cuerda punteada de la figura 4.6. Por lo tanto, podemos anticipar que el modo fundamental y por lo menos el tercer armónico deberán estar presentes simultáneamente en la vibración de una cuerda punteada en el punto medio. Podemos mejorar en gran medida la aproximación a la forma de la figura 4.6, sumando más armónicos superiores en proporciones adecuadas (Fig. 4.7 [b]). Se pueden limar los serpeos que quedan, agregando más y más armónicos superiores en proporciones adecuadas hasta que la forma deseada sea reproducida casi exactamente.

En todo esto no hay conjetura alguna: puede demostrarse matemáticamente que *cualquier forma inicial arbitraria de una cuerda puede ser reproducida con un grado arbitrario de precisión por medio de una determinada superposición de formas geométricas que correspondan a los modos de vibración armónicos de la cuerda* (ondas estacionarias). Es esta superposición «matemática» de formas geométricas, en particular, la proporción de sus amplitudes y fases, la que define la vibración de la cuerda cuando ésta se suelta desde su posición inicial. Expresado en otros términos, cada una de las ondas estacionarias componentes, que al superponerse dan a la cuerda su forma inicial (por ej., Fig. 4.7(b)), procederá a vibrar, cuando se suelte la cuerda, en su propio modo, con su propia frecuencia y amplitud. A medida que el tiempo transcurre, la forma

instantánea de la cuerda cambia periódicamente de una manera complicada, pero cada vez que pase un período fundamental $\tau = 1/f_1$, todos los modos componentes se encontrarán en la misma relación que al comienzo, y la cuerda tendrá la misma forma que en el instante inicial. Es muy importante señalar aquí que la configuración inicial de la cuerda no solo determina las amplitudes de los modos de vibración armónicos, sino también sus fases ('timings' relativos). El punto en el cual la cuerda es punteada determinará la particular proporción de armónicos superiores, es decir, el timbre del sonido emitido (Sec. 1.2). Si punteamos en el centro, tendremos la situación mostrada en la Fig. 4.7(b), y solo aparecerán armónicos *impares*. Por otra parte, cuanto más cerca punteamos de los extremos fijos, tanto más rica será la proporción de armónicos superiores. En general, todos aquellos armónicos que tengan un nodo en el punto donde se puntea serán suprimidos (por ej., todos los armónicos pares en el ejemplo de la Fig. 4.6), mientras que aquellos que tengan un vientre serán reforzados. Esto, por ejemplo, es explotado al máximo por el arpista para controlar el timbre de su instrumento.

Cuando una cuerda es puesta en vibración por la acción de puntear, estamos en una situación en la que el mecanismo primario de excitación por medio de la deformación de la cuerda, entrega cierta cantidad de *energía potencial* al sistema. Después de soltar la cuerda, la energía inicial se convierte periódicamente en energía cinética vibratoria (Sec. 3.1). Por otra parte, cuando una cuerda es puesta en vibración por la acción de percudir, una cierta cantidad de *energía cinética* es aportada por el mecanismo que percute (por ej., el martillo de un piano), poniendo en movimiento los puntos de la cuerda inicialmente no deformada. Esta energía inicial es luego convertida periódicamente en energía potencial de deformación. Puede demostrarse matemáticamente que *conociendo las velocidades iniciales de los puntos de la cuerda golpeada, es posible deducir la superposición de armónicos resultante*. De este modo, una cuerda golpeada en el punto medio oscilará principalmente con la frecuencia fundamental, más una mezcla de armónicos impares con intensidades decrecientes. Cuanto más cerca de los puntos extremos P y Q se golpee una cuerda, tanto más rico en armónicos superiores será el tono producido. Como ocurre con una cuerda punteada, serán excluidos los armónicos cuyos nodos están en el punto donde se percute, o cerca de él; aquellos que tengan un vientre allí serán reforzados. En situaciones reales de cuerdas de piano percutidas con el martillo, estudios teóricos y cuidadosas mediciones (Hall y Askenfelt, 1988) han establecido que la *duración* del contacto martillo-cuerda tiene gran influencia en la mezcla de modos armónicos superiores: cuanto más larga sea la duración del contacto, tanto más pobre en armónicos superiores resultará la vibración de la cuerda (porque quedarán excluidos los modos de períodos más cortos que la duración del contacto).

66

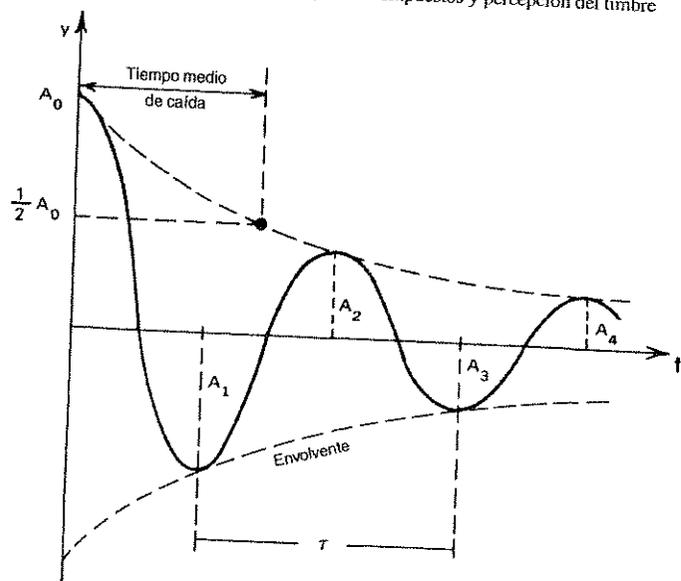


Figura 4.8 Gráfico de una oscilación armónica amortiguada.

Cuando una cuerda es puesta en vibración punteándola o percutiéndola, observamos que la vibración decae en forma bastante rápida. Esto se debe a la acción de fuerzas disipativas: la fricción elástica dentro de la cuerda misma y, más que nada, fuerzas que ponen en leve movimiento vibratorio todo aquello que está sujetando la cuerda en sus extremos. Solo una parte de esta energía disipada es convertida en una onda sonora. Una cuerda que vibra libremente, montada sobre una estructura rígida y pesada, produce un sonido apenas audible: la mayor parte de la energía vibratoria desaparece en forma de energía de fricción (calor). La conversión en energía de onda sonora puede incrementarse mucho montando la cuerda sobre una plancha de madera de propiedades elásticas especiales, llamada *resonador* (tabla armónica del piano, cuerpo del violín). En ese caso, se permite que los puntos extremos de la cuerda vibren mínimamente (tan poco que, comparados con el resto de las vibraciones de la cuerda, estos puntos extremos técnicamente siguen funcionando como nodos), y la energía de la cuerda puede ser gradualmente convertida en energía vibratoria de la plancha. Debido a la superficie - en general relativamente grande - de esta plancha, esta energía es convertida de manera mucho más eficiente en energía de onda sonora. El sonido resultante es mucho más fuerte que en el caso de una cuerda montada rígidamente; pero *decae mucho más rápido* a causa de la *razón*, considerablemente mayor, a la

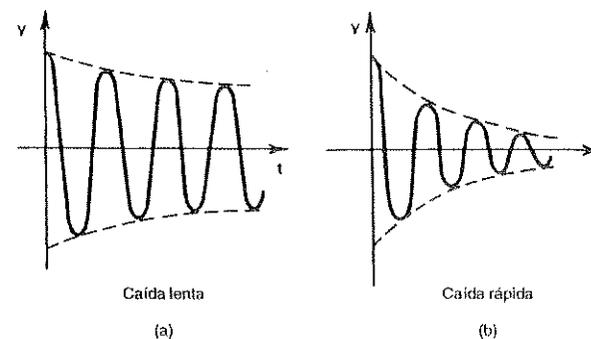


Figura 4.9 Caídas lenta y rápida de una oscilación amortiguada.

cual se gasta la energía disponible de la cuerda (véase potencia, Sec. 3.1).

Examinemos con más detalle el proceso de decaimiento de una vibración. Para simplificar, consideremos una cuerda que ha sido puesta a vibrar libremente sólo en su modo fundamental. Centremos nuestra atención en la amplitud gradualmente decreciente de la oscilación de la cuerda, digamos, en un vientre. Las mediciones muestran que, para una cuerda dada, las oscilaciones amortiguadas con amplitudes más grandes decaen más rápidamente que aquellas que tienen menor amplitud. El movimiento resultante se muestra en la Fig. 4.8. Obsérvese la pendiente de la curva de la envolvente, la cual es menos pronunciada a medida que decrece la amplitud. Esto se llama *caída exponencial* de la amplitud. Lo importante (¡afortunadamente para la música!) es que la frecuencia de una oscilación amortiguada lentamente se mantiene constante.

Esta es más o menos la forma en que se comporta una cuerda que vibra libremente en un solo modo después de haber sido punteada o percutida. Si está montada sobre una plancha rígida, la pérdida de energía será relativamente pequeña y lo mismo ocurrirá con el amortiguamiento de la amplitud (Fig. 4.9 [a]). Si, en cambio, está montada sobre una tabla armónica, pasará energía a una razón mayor por poner a la plancha y al aire circundante en estado de oscilación. Las oscilaciones decaerán entonces más rápidamente (Fig. 4.9 [b]).

Una cantidad característica es el así llamado *tiempo medio de caída*. Este es el intervalo de tiempo después del cual la amplitud de las oscilaciones ha sido reducida a la mitad de su valor inicial (Fig. 4.8). Lo notable es que en una caída exponencial, este tiempo medio es siempre el mismo a lo largo de toda la caída: tarda el mismo tiempo reducir a la mitad la amplitud, no importa cuál sea el valor actual de esta última. El tiempo medio de caída es, por lo tanto, una constante característica de una oscilación amortiguada. El tiempo medio típico de caída de una cuerda de piano es de alrededor de 0,4s.

Cuando una cuerda vibra en varios modos al mismo tiempo, la situación es más compleja. Si bien *cada modo* decaerá exponencialmente, el tiempo medio de cada será distinto para cada uno. El tono compuesto resultante no solamente decrecerá en sonoridad, sino también su timbre cambiará en forma gradual. En las cuerdas de piano, los modos de vibración de frecuencias más altas decaen algo más rápido que los armónicos más bajos; en una campana que vibra, los parciales más bajos continúan sonando mucho tiempo después que los más altos han decaído.

Ha habido una larga disputa entre pianistas y físicos con respecto a lo que se llama «toque» en la ejecución pianística. Los pianistas prestan gran atención *al modo* en que aprietan una tecla de piano y argumentan que esto influye al sonido producido mucho más allá de determinar su sonoridad. El físico responde que, en vista de que el martillo está en vuelo libre totalmente desligado del ejecutante durante la última porción de su movimiento, el tono resultante solamente puede depender de *un* único parámetro: la *velocidad* final con que el martillo golpea a la cuerda. Por lo tanto, en el caso de un tono *aislado*, el toque pianístico no es otra cosa que sonoridad, con un timbre irrevocablemente acoplado a esa sonoridad⁵, y al decaimiento subsiguiente. El físico dice: todo lo que un ejecutante realmente puede hacer, es controlar la velocidad final del martillo; la cualidad del tono no puede ser cambiada independientemente de la sonoridad, y el buen o mal «toque» no puede existir para tonos aislados. El toque «real» cuando se ejecuta una pieza musical está relacionado con otros efectos psicoacústicos, tales como el sutil control de la duración de cada tono, pequeñas variaciones de sonoridad de un tono a otro, resaltar la melodía por encima del acompañamiento, diferencias en la sonoridad y distribución temporal de las notas de un acorde, incluso el componente percusivo dado por el ruido de las teclas cuando éstas golpean el freno (Askenfelt y Jansson, 1990). Hay esperanza, sin embargo, para los pianistas que participan en la «disputa sobre el toque»: mediciones recientes (Askenfelt y Jansson, 1990) han revelado que el movimiento detallado del martillo volando libremente hacia la cuerda como un *cuerpo elástico oscilante y que rota*, puede ser influenciado por distintos tipos de toque

⁵ Hay una razón física un tanto complicada para este acoplamiento entre sonoridad y timbre. Como mencionamos antes (pág. 129), la duración del contacto martillo-cuerda influye en la proporción relativa de los modos armónicos superiores; un contacto más largo conduce a menos armónicos superiores. La duración del contacto a su vez depende de la rigidez del fieltro en la cabeza del martillo: un martillo más blando queda más tiempo en contacto con la cuerda que un martillo más duro (para igual velocidad de impacto). Pero hay un hecho notable (Hall y Askenfelt, 1988): la *rigidez* efectiva de un martillo depende de la velocidad del impacto con la cual el martillo golpea la cuerda. Velocidades de impacto más altas conducen a una mayor rigidez efectiva, y viceversa (esto se llama comportamiento no lineal de la rigidez). Como consecuencia, golpear con mayor fuerza una tecla de piano no solo incrementa la amplitud de la oscilación de la cuerda (tono más sonoro) sino que también acorta el tiempo de contacto martillo-cuerda, incrementando *automáticamente* la proporción de armónicos superiores (timbre más brillante).

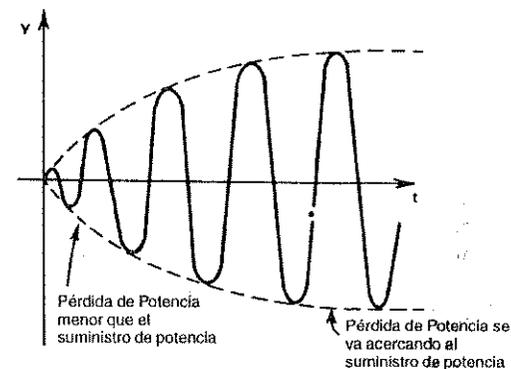


Figura 4.10 Formación de una oscilación armónica a partir de un suministro constante de potencia.

(para ser más precisos, por diferentes aceleraciones del martillo antes de su escape, aceleraciones controladas por el ejecutante). Esto podría conducir a movimientos de roce contra la cuerda durante el contacto que dependerían del toque; pero todavía no ha sido demostrado que este efecto influya en forma mensurable en la excitación de la cuerda.

¿Qué podemos hacer para evitar el amortiguamiento de la vibración de una cuerda? Obviamente, debemos compensar la pérdida de energía, suministrando de algún modo energía extra a nuestro sistema vibratorio, a un monto por unidad de tiempo que sea igual al de la potencia disipada. Si la potencia que se suministra *excede* la cantidad de energía que se disipa, la amplitud crecerá gradualmente. Pero este crecimiento no seguirá indefinidamente: si el suministro de potencia se mantiene constante, la disipación de potencia aumentará a medida que aumenta la amplitud, y eventualmente se llegará a un régimen en el cual la potencia disipada es igual a la potencia suministrada (Fig. 4.10). Esto ocurre en instrumentos con tonos de sonoridad constante (cuerda de violín frotada, flauta, tubo de órgano, etc.). En tal caso cada armónico se forma independientemente, como si hubiera un mecanismo de suministro de potencia para cada modo. Cuanto más grande sea este suministro de potencia, tanto mayor será, por supuesto, el nivel de intensidad final.

El mecanismo de frotamiento es un buen ejemplo de cómo las oscilaciones de una cuerda pueden ser mantenidas en un régimen constante. El problema físico es matemáticamente complicado, y solo puede encararse después de hacer varias simplificaciones (Friedlander, 1953; Keller, 1953). Aquí sólo podemos dar una descripción cualitativa de la teoría. La interacción entre el arco y la cuerda se produce debido a fuerzas de fricción. En general, distinguimos dos tipos de interacción por fricción. La primera se llama *fricción estática*, en la que no hay *ningún*

68

desplazamiento entre los cuerpos interactuantes. Esto ocurre cuando la cuerda «está pegada» al arco, moviéndose a la misma velocidad que éste (o, en términos más familiares, cuando uno trata de mover una mesa pesada, mientras ésta permanezca «pegada» al piso). El segundo tipo es la llamada fricción *dinámica*, que aparece cuando los dos cuerpos que interactúan (sus superficies de contacto) se *deslizan* uno contra el otro. Esto acontece cuando la cuerda «salta para atrás» y se mueve en dirección opuesta al arco (y ocurre cuando uno continúa empujando la mesa *después* que ésta haya empezado a moverse). La fricción dinámica es más débil que la fricción estática; los dos mecanismos están controlados por la fuerza perpendicular a la superficie de contacto, que un cuerpo ejerce sobre el otro. En el caso de una cuerda frotada, esta fuerza se llama *presión de arco*, expresión que suena horrible a un físico, porque *no* es una presión, sino simplemente una fuerza.

En el Apéndice I discutimos con mayor detalle un caso ideal. Las principales conclusiones desde el punto de vista físico son las siguientes: 1) *la amplitud de la vibración de una cuerda frotada (sonoridad del tono) está controlada exclusivamente por la velocidad del arco; pero para mantener constante el tipo o la naturaleza del movimiento de la cuerda (timbre del tono), la presión de arco debe ser mantenida proporcional a la velocidad del arco*. Esto es bien conocido por los ejecutantes de instrumentos de cuerda, quienes incrementan al mismo tiempo tanto la velocidad *b* del arco como la presión de arco *P*, para lograr un aumento de la sonoridad *sin* producir un cambio de timbre; para lograr un aumento de sonoridad *con* cambio de timbre deben incrementar *b* y disminuir *P*. 2) *Una cuerda frotada tiene siempre una forma instantánea que está formada por segmentos rectilíneos*; este resultado fue verificado experimentalmente hace ya mucho tiempo. Un estudio del balance de energía en el mecanismo de frotamiento nos muestra que la mayor parte de la energía suministrada a la cuerda por el arco durante las etapas de «adherencia» es disipada en forma de calor de fricción (trabajo de la fuerza de fricción dinámica) en las etapas en que la cuerda se «zafa» del arco. ¡Solo una pequeña fracción es convertida en energía sonora!⁶

Como en el caso de una cuerda punteada o golpeada, la particular mezcla de modos armónicos de vibración dependerá de la posición del punto de frotamiento. Frotar cerca del puente (*sul ponticello*) realza los armónicos superiores y hará al tono más «brillante»; frotar más cerca del punto medio de la cuerda (*sul tasto*), reduce la intensidad de los armónicos superiores considerablemente, y el sonido es «más blando».

En el comentario previo hemos supuesto tácitamente que el arco se desplazaba en forma perfectamente perpendicular a la cuerda. Si hay un pequeño componente de movimiento paralelo, pueden excitarse modos

de vibración *longitudinales* de la cuerda. Su frecuencia es mucho más alta que la frecuencia fundamental de los modos vibratorios transversales; estas oscilaciones longitudinales son las responsables de los chillidos que se escuchan cuando toca un principiante.

4.3 Espectros sonoros y resonancia

Cuando una cuerda vibra simultáneamente en una serie de modos diferentes, las ondas sonoras generadas también son complejas. Cada componente armónico de la vibración original de la cuerda contribuye independientemente a la onda sonora resultante, con una frecuencia igual a la del modo correspondiente, y una intensidad y fase dadas por los procesos de transformación que intervienen. El resultado es una superposición de ondas sonoras formando una onda compleja, de frecuencia fundamental f_1 (igual a la frecuencia fundamental del elemento vibrante [4.3]) y con una serie de armónicos superiores de frecuencias $2f_1, 3f_1, 4f_1, \dots$, etc. La vibración resultante es periódica, repitiéndose con un período igual a $\tau_1 = 1/f_1$. En otras palabras, la frecuencia fundamental f_1 también representa la frecuencia de repetición de la vibración compleja resultante (Sec. 2.7). La forma de la curva resultante depende de *qué* armónicos están presentes, y de *cuánto* haya de cada uno de ellos (o sea, de sus amplitudes relativas), así como de sus '*timings*' relativos (es decir, de sus fases relativas).

Aquí llegamos a un teorema que ha tenido un tremendo impacto en todas las ramas de la física, incluso la física de la música. El teorema dice que *cualquier vibración periódica, por complicada que sea, puede ser representada como una superposición de vibraciones armónicas puras*, cuya frecuencia fundamental está dada por la frecuencia de repetición de la vibración periódica. Pero esto no es todo: ¡este teorema también da las «recetas» matemáticas para que se puedan calcular las amplitudes y fases de los componentes armónicos superiores! Se llama teorema de Fourier, en honor a un famoso matemático francés del siglo diecinueve. La determinación de los componentes armónicos de un movimiento periódico complejo dado se llama *análisis de Fourier*; la determinación del movimiento periódico complejo resultante a partir de un conjunto de armónicos dado se denomina *síntesis de Fourier*. De manera similar, dado un tono compuesto, el proceso por el cual hallamos los componentes armónicos se llama *análisis de frecuencia*. Recíprocamente, dado un grupo de armónicos, la operación de mezclarlos para formar un tono compuesto recibe el nombre de *síntesis sonora*.

Examinemos un ejemplo de análisis de Fourier. Desde luego, no podemos describir aquí las operaciones matemáticas necesarias para obtener los resultados numéricos; aceptemos que existen y

⁶Para un estudio detallado sobre cuerdas frotadas, véase Schelleng (1973).

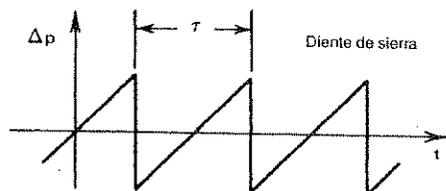


Figura 4.11 Oscilación de presión «diente de sierra».

concentrémonos en su interpretación física. Tomemos el movimiento periódico graficado en la Fig. 4.11, que corresponde a una onda «diente de sierra». τ es el período, $f_1=1/\tau$ es la frecuencia de repetición o frecuencia fundamental. Este tipo de vibración puede generarse electrónicamente. En cierta medida, representa la forma idealizada del movimiento de una cuerda frotada. La figura 4.12 muestra cómo este movimiento periódico puede construirse como una serie de vibraciones armónicas puras. Desde luego, todavía quedan muchos serpenteos en la curva que representa la suma. Pero esto se debe a que nosotros, para mayor claridad, paramos de sumar armónicos después del sexto. Si sumamos más y más armónicos superiores (con amplitudes y fases obtenidas por el análisis de Fourier), eliminaremos estas ondulaciones, aproximándonos cada vez más a la forma de diente de sierra. Obsérvese atentamente las amplitudes y fases relativas de los armónicos componentes, y cómo las porciones positivas y negativas se suman para dar la curva resultante. Con un poco de intuición y un ojo entrenado es posible predecir cualitativamente los principales componentes armónicos de un movimiento periódico de casi cualquier tipo.

Ahora nos proponemos encontrar un modo de caracterizar, desde el punto de vista físico, un tono compuesto dado. En principio, debemos especificar tres series distintas de cantidades: las frecuencias de los armónicos componentes, las amplitudes de la variación de presión o intensidades de los componentes, y sus fases o 'timings' relativos (por ej., Fig. 2.5). En la práctica, sin embargo, es habitual especificar solamente la frecuencia fundamental f_1 y las intensidades de los armónicos componentes, porque, primero, se sabe que todas las frecuencias superiores son múltiplos enteros de la fundamental, y segundo, las fases de los componentes, particularmente las de la primera docena (los más importantes), solo tienen un rol secundario en la percepción del timbre (Sec. 4.8).

La secuencia de los valores de intensidad I_1, I_2, I_3, \dots de los componentes armónicos de un tono compuesto se llama *espectro de potencia* del tono. Dos tonos compuestos de la misma altura y sonoridad, pero de distintos espectros, suenan diferentes, con diferente timbre. La diferencia en el espectro provee claves importantes al cerebro para distinguir tonos que proceden de distintos instrumentos. Sin embargo, también son necesarias otras claves, en especial el ataque y la caída del tono, para poder identificar

un instrumento dado (Sec. 4.9). El hecho de que toda una multiplicidad de parámetros físicos (I_1, I_2, I_3, \dots) esté relacionada con la sensación de timbre, indica que este último es una magnitud psicofísica multidimensional.

Los espectros sonoros pueden representarse gráficamente, trazando para la frecuencia de cada armónico (eje horizontal) la intensidad con la cual ese armónico interviene (eje vertical) (por ej., Fig. 4.16). Con frecuencia, en lugar de usar intensidades para representar un espectro dado se emplean los valores de IL (3.15) o de SPL (3.16) de cada armónico. También se utilizan valores de intensidad o de IL relativos a aquél de la fundamental, o relativos a la intensidad total $I = I_1 + I_2 + I_3 + \dots$. Existen muchos libros en los cuales se reproducen los espectros de tonos de instrumentos musicales reales (por ej., Culver, 1956). Es necesario, sin embargo, poner un toque de atención con relación a esto. Desde el punto de vista psicofísico, la representación (Fourier) convencional del espectro de un tono no tiene mayor sentido más allá del sexto o séptimo armónico, porque en ese rango los componentes vecinos empiezan a caer dentro de una misma banda crítica. Dado que ésta es la unidad de recolección e integración de información fundamental del oído (Sec. 2.4), el sistema

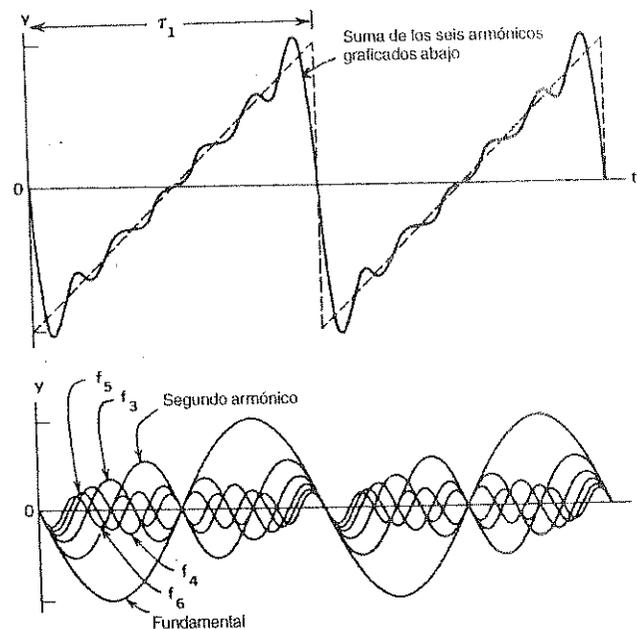


Figura 4.12 Análisis de Fourier (hasta el sexto armónico) de una onda diente de sierra.

70

auditivo no es capaz de distinguir separadamente las intensidades de estos armónicos más altos (caso 1] en pág. 105). Una representación más significativa, en el sentido psicofísico, del espectro de un tono está dada por los valores integrados de intensidad *por banda crítica* (intervalos de frecuencia de aproximadamente $\frac{1}{3}$ de octava).

Solamente los tonos constantes en el tiempo pueden ser «descompuestos» en una superposición de armónicos de frecuencias discretas, múltiplos enteros de la fundamental. Cuando un patrón vibratorio *cambia* en el tiempo, esto ya no es más posible. No obstante, se puede aplicar una suerte de «versión expandida» del análisis de Fourier. Puede demostrarse matemáticamente que un tono dependiente del tiempo tiene un *espectro continuo*, en el cual se presentan todas las frecuencias, con una determinada intensidad en cada intervalo infinitesimal de frecuencia. Si el tono cambia *lentamente* en el tiempo, las frecuencias discretas (aquellas de los armónicos) todavía seguirán siendo las más intensas (picos espectrales), pero si el cambio es notable de ciclo a ciclo, el carácter discreto desaparecerá, y el espectro tenderá a convertirse en una curva continua que cubre la totalidad del rango de frecuencias (incluso si el tono fue originalmente puro). Este fenómeno nos lleva a otra importante observación con respecto a los equipos de «alta fidelidad» (véase también pág. 50). Dijimos en la sección 1.2, y volveremos a esto en la sección 4.8, que los *transientes*, es decir, las rápidas variaciones en el tiempo del patrón vibratorio de un tono, tienen un rol muy importante en la percepción del timbre (cualidad del sonido). Por lo tanto, para reproducir correctamente los transientes de un tono, los sistemas de grabación y reproducción deben evitar la distorsión del espectro de un tono *a lo largo de todo el rango de frecuencias*. Nuestro oído no necesita los componentes de frecuencias mucho más allá de los 5.000 Hz, de un tono constante, ¡pero el sistema de reproducción sí los necesita para darnos una versión correcta de las porciones del tono que cambian rápidamente!

El espectro sonoro de un instrumento de cuerda no es igual al de las vibraciones de las cuerdas. La razón de esto radica en la eficiencia, dependiente de la frecuencia, del *resonador* (tabla armónica del piano, cuerpo del violín), cuya función principal es extraer energía de la cuerda vibrante y convertirla más eficientemente en potencia de onda sonora. Como lo mencionamos antes, las vibraciones de la cuerda son convertidas en vibraciones del resonador mediante un proceso en el cual se permite a los puntos extremos de la cuerda (especialmente el que está sobre el puente) vibrar en una medida muy pequeña. Esta vibración residual es tan pequeña que, desde el punto de vista de la cuerda, estos puntos siguen actuando como nodos vibratorios. Sin embargo, estas vibraciones involucran *una considerable transferencia de energía*.⁷ La explicación de esto radica en

⁷La función de la *sordina*, cuando se aplica en el puente de un instrumento de cuerda, es

la definición misma de trabajo (Sec. 3.1): aunque los desplazamientos de los puntos extremos de la cuerda son extremadamente pequeños, las *fuerzas* aplicadas sobre ellos son grandes (del orden de la tensión de la cuerda), de modo que el *producto* fuerza por desplazamiento (trabajo) puede ser apreciable. Debido a que un resonador típico tiene una superficie relativamente amplia, la conversión de la energía de sus vibraciones a energía de ondas sonoras es muy eficiente; miles o incluso millones de veces más eficiente que la conversión directa de la energía de la cuerda vibrante a sonido.

Tal como una cuerda de extremos fijos, la compleja estructura elástica de una tabla armónica de piano o la del cuerpo de un violín tiene modos de oscilación preferidos. En este caso, sin embargo, no hay más una simple relación de múltiplos enteros entre las frecuencias asociadas, como en la relación (4.2). Lo que es más, hay tantos modos con frecuencias prácticamente superpuestas, que lo que se obtiene es un 'continuum' de frecuencias de vibración preferidas, en vez de una serie de valores discretos.⁸ Veamos brevemente cómo surgen estos modos de vibración. Para ello, en lugar de considerar el cuerpo de un violín entero, examinemos como vibra sólo la tapa o el fondo del instrumento. Para averiguar los modos de vibración posibles es necesario excitar la tapa con un vibrador mecánico sinusoidal, de una sola frecuencia, en un punto determinado (por ejemplo, en el punto donde está en contacto con el puente). Las ondas elásticas se propagan en dos dimensiones alejándose del punto de excitación y se reflejan en los bordes de la tapa. Los únicos modos de vibración estables serán ondas estacionarias compatibles con las condiciones particulares del contorno de la tapa. Es muy difícil visualizar el tratamiento matemático de este proceso. En el laboratorio, sin embargo, es posible hacer visibles las vibraciones a través de una moderna técnica de láser llamada holografía (Reinicke y Cremer, 1970). El modo de oscilación más simple (llamado «modo en anillo» [«ring mode»]) es uno en el cual la región central de la tapa se mueve sinusoidalmente para arriba y para abajo, con el borde actuando como línea nodal. Los modos de anillo de las tapas del cuerpo de un violín determinan la altura del «tono de golpeteo» [«tap tone»], que es el tono que surge al dar palmaditas al cuerpo del instrumento. Las figuras 4.13 y 4.14 (Jansson et al., 1970) muestran hologramas de cuatro modos de vibración sucesivos de la tapa (con oídos y alma, pero sin mango) y del fondo respectivamente. Cada una de las curvas oscuras representa un contorno de igual amplitud de deformación instantánea. La diferencia en amplitud entre franjas vecinas es de alrededor de 2×10^{-5} cm. En el instrumento *armado*, los modos de

disminuir esta transferencia de energía para los componentes de frecuencias más altas, alterando de este modo la cualidad del tono resultante.

⁸Una cuerda *real*, de espesor apreciable, tampoco tiene modos discretos, angostos, de vibración.

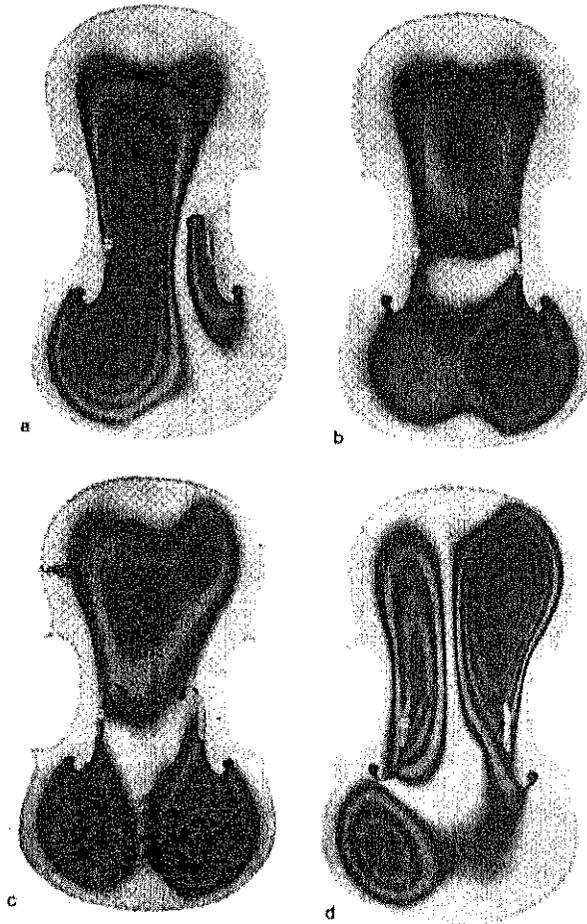


Figura 4.13 Hologramas que muestran los primeros cuatro modos de vibración de la tapa de un violín (con oídos y alma, sin mango) Cada una de las curvas oscuras representa un contorno de igual amplitud de deformación. (a) 540 Hz; (b) 775 Hz; (c) 800 Hz; (d) 980 Hz. Reproducido con permiso de Jansson et al., 1970.

vibración de la tapa (Fig. 4.13) prácticamente no cambian, pero nuevos modos de vibración aparecen en el rango de frecuencias bajas.

La respuesta vibratoria de un resonador a una determinada señal de amplitud fija (sea de un vibrador mecánico o de una cuerda vibrando montada sobre ese resonador) depende mucho de la frecuencia de las oscilaciones primarias. Por esa razón, una tabla armónica reacciona de

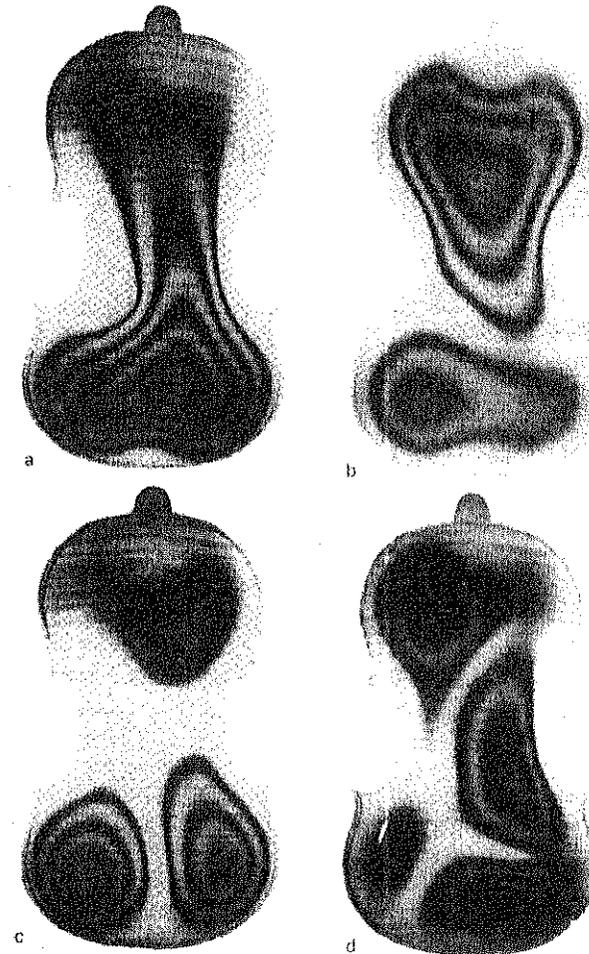


Figura 4.14 Lo mismo que la Fig. 4.13, pero para el fondo del violín. (a) 740 Hz; (b) 820 Hz; (c) 960 Hz; (d) 1110 Hz. Reproducido con permiso de Jansson et al., 1970.

manera distinta a vibraciones de distintas frecuencias. Algunas frecuencias serán realzadas de modo preferencial, mientras que otras no serán amplificadas en absoluto. Una frecuencia para la cual la conversión de energía es particularmente eficiente se llama *frecuencia de resonancia* del resonador. Un resonador puede tener varias frecuencias de resonancia; ellas pueden estar bien definidas (resonancia angosta) o esparcidas en una ancha gama de frecuencias. El gráfico que se obtiene trazando la

señal de salida (por ejemplo, midiendo la intensidad de la onda sonora que emerge) como función de la frecuencia de una vibración sinusoidal de amplitud constante que sirve de estímulo, se llama *curva de resonancia*, o curva de respuesta.

Habitualmente, la intensidad I de la señal de salida se representa en relación a alguna señal de referencia dada I_{ref} , y se expresa en *decibeles* (db) (Sec. 3.4):

$$R = 10 \log \frac{I}{I_{ref}} \quad (\text{en decibeles}), \quad (4.4)$$

donde R es el valor de la función de respuesta. El gráfico de la dependencia de R con la frecuencia da la curva de resonancia arriba mencionada. La Fig. 4.15 es un ejemplo, que corresponde a la tapa de un violín (Hutchins y Fielding, 1968). El primer pico pronunciado en la Fig. 4.15 corresponde al tono de golpeteo [‘tap tone’]⁹ La curva de respuesta de la caja de un violín *ensamblado* revela un primer pico de resonancia en el rango 280 a 300 Hz (¡cerca de la cuerda *Re!*) que corresponde al primer modo de vibración del *aire* encerrado en el cuerpo del violín. La siguiente resonancia, que se encuentra en general una quinta arriba de la resonancia del aire, se llama *resonancia principal de la madera*. Más allá de los 1.000Hz, los múltiples picos de resonancia de un violín completo, son mucho menos pronunciados que aquellos de la tapa o del fondo vibrando aisladamente, como surge de la Fig. 4.15. La curva de respuesta de una tabla armónica de piano es aún más complicada; esta complicación, sin embargo, asegura una amplificación relativamente pareja a lo largo de todo el rango de frecuencias.

La figura 4.15 representa la curva de respuesta de un resonador a una vibración simple, *armónica*, de una sola frecuencia f . ¿Qué ocurre cuando el resonador es excitado con una cuerda que vibra con un espectro complejo de armónicos, de frecuencias $f_1, f_2 = 2f_1, f_3 = 3f_1, \dots$, etc., e intensidades I_1, I_2, I_3, \dots ? Cada armónico será transformado independientemente, según el valor de la función de respuesta R para su propia frecuencia. El timbre del sonido resultante estará por lo tanto determinado por dos factores: el espectro de la vibración original de la cuerda y la curva de respuesta del resonador.

Para dar un ejemplo, consideremos el espectro hipotético de una cuerda en vibración, de la figura 4.16. Esta cuerda está montada sobre una hipotética tabla de resonancia, cuya curva de respuesta también se muestra en la figura. El espectro sonoro final estará dado por el gráfico de la derecha (valores relativos). La fundamental se reduce considerablemente, mientras que el quinto armónico aparece realizado por encima de todos los demás.

⁹ La posición (en frecuencia) y la forma de este particular pico de resonancia son de capital importancia para la calidad de un instrumento de cuerda (Hutchins y Fielding, 1968). Véase también págs. 170 y 171.

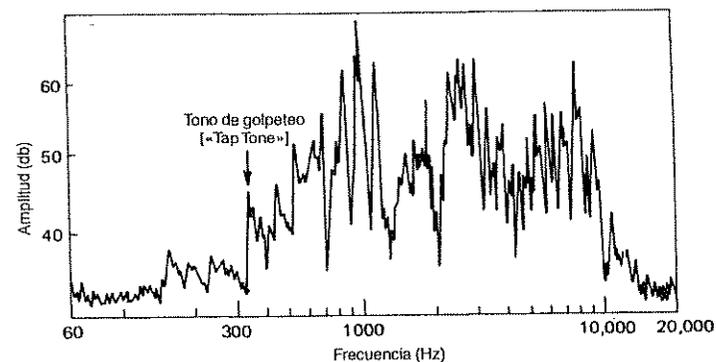


Figura 4.15 Curva de resonancia de un plato de violín (Hutchins y Fielding, 1968). Con permiso de *Physics Today*.

Según el ejemplo de la figura 4.16, se extraería más potencia del quinto armónico que de cualquier otro. Si la cuerda hubiese sido originalmente punteada o percutida, este armónico decaería más rápido que los otros, porque su reserva de energía se agotaría con mayor rapidez. Esto lleva a un cambio del espectro del tono (de su timbre), cambio que depende del tiempo, a medida que el sonido se extingue. Si, por el contrario, se frotara la cuerda, la pérdida de energía de cada modo sería automáticamente compensada por el mecanismo de frotamiento, y el tono resultante tendría un timbre que se mantendría constante a través del tiempo.

Finalmente, llegamos a un punto muy importante para la música. La curva de respuesta de un resonador es una característica invariable de un instrumento musical. Si, por ejemplo, el resonador de un instrumento tiene una región de resonancia alrededor de los 1.000 Hz, reforzará todos los armónicos superiores cuyas frecuencias caigan cerca de los 1.000 Hz, sin importar *qué* nota sea la que esté sonando (siempre que, por supuesto, la frecuencia fundamental de esa nota esté debajo de los 1.000Hz), ni cuál

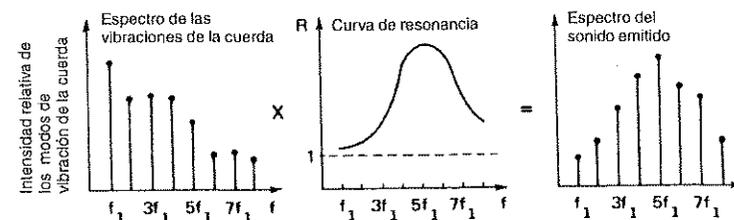


Figura 4.16 Efecto de una tabla armónica (de resonancia) con una hipotética curva de resonancia sobre el espectro de una vibración compleja de una cuerda.

sea el espectro de la vibración original de la cuerda. Una región de resonancia ancha, que refuerza los armónicos superiores que caen en un rango de frecuencias fijo, se llama *formante*. Un instrumento musical (su resonador) puede tener varios formantes. Se cree que los formantes, es decir, el reforzamiento de armónicos en bandas de frecuencia fijas, características del instrumento, son usados por el sistema auditivo como uno de los «rasgos» más importantes en el proceso de *identificación* de un instrumento musical (Sec. 4.9). Una de las razones que apoyan esta hipótesis es el hecho de que los formantes son la única característica común a la mayor parte de los sonidos de un instrumento determinado, mientras que el espectro de cada una de las notas individuales puede variar mucho de una a otra.

4.4 Ondas estacionarias longitudinales en una columna de aire ideal

Consideremos un cilindro largo, muy angosto, abierto en ambos extremos (Fig. 4.17). El aire en su interior puede ser considerado como un medio elástico unidimensional (Sec. 3.2) a través del cual pueden propagarse ondas longitudinales. En cualquier punto *dentro* del cilindro, la presión podrá momentáneamente aumentar, disminuir u oscilar considerablemente con relación a la presión atmosférica normal externa; las paredes rígidas y la inercia de la columna de aire mantienen el necesario equilibrio de fuerzas (3.1) que se originan a causa de las diferencias de presión. Pero en los extremos abiertos *P* y *Q* no podrán ocurrir variaciones de presión grandes, ni siquiera en un intervalo de tiempo muy breve, porque no hay nada en esos puntos para equilibrar las diferencias de presión que se generan. Estos puntos funcionan por lo tanto como nodos de *presión*, y cualquier onda sonora que cause una perturbación en el interior del tubo y que se propague a lo largo de éste, será reflejada en cualquiera de los extremos abiertos. De aquí que nos hallemos ante una situación análoga a la de una cuerda vibrante, caso tratado en la Sec. 4.1: las ondas sonoras generadas en el tubo quedan atrapadas dentro del mismo y *los únicos modos estables de vibración posibles son ondas estacionarias longitudinales con nodos de presión en los extremos abiertos P y Q* (Fig. 4.17). Obsérvese que, de acuerdo con lo discutido en la sección 3.3, pág. 98, los extremos abiertos serán *vientres de desplazamiento*, es decir, puntos con máxima amplitud de vibración.

La columna de aire abierta no necesariamente tiene que estar definida físicamente de la manera mostrada en la Fig. 4.17. Por ejemplo, en la Fig. 4.18, tenemos una columna de aire abierta, comprendida entre los puntos *P* y *Q* del tubo. Efectivamente, dado que tenemos agujeros en los puntos *P* y *Q*, la presión del aire en esos lugares debe quedar constante e igual a la presión externa. *P* y *Q* funcionan de este modo como los extremos

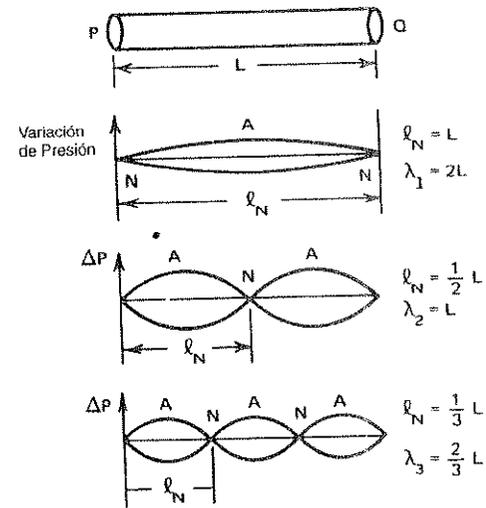


Figura 4.17 Modos de ondas estacionarias (variaciones de presión) en un tubo idealmente cilíndrico, abierto en ambos extremos.

abiertos de la columna de aire encerrada. La Fig. 4.18 corresponde al caso de una flauta idealizada, donde *P* es la embocadura y *Q* el primer orificio destapado.

En un tubo abierto real de diámetro finito, *los nodos de presión no ocurren exactamente en el extremo abierto sino a una pequeña distancia afuera («corrección del extremo» [end correction], pág. 156)*. Las relaciones que se dan abajo son sólo aproximaciones.

A partir de la Fig. 4.17 y de la relación (3.6) obtenemos las frecuencias de los modos de vibración de un tubo cilíndrico abierto:

$$f_n = \frac{n}{2L} 20,1 \sqrt{t_A} = n f_1 \quad n = 1, 2, 3, \dots \quad (4.5)$$

f_1 es la frecuencia fundamental

$$f_1 = \frac{10,05}{L} \sqrt{t_A} \quad (4.6)$$

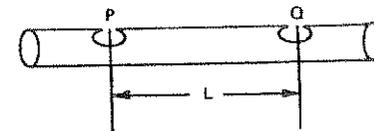


Figura 4.18 Tubo elemental con dos orificios.

Recuérdese que t_A es la temperatura absoluta del aire en el tubo, dada por la ecuación (3.5). En las ecuaciones (4.5) y (4.6) L debe expresarse en metros. Considerando que la longitud de onda λ_1 del tono fundamental está relacionada con la longitud L del tubo, teniéndose que $\lambda_1 = 2L$ (Fig. 4.17), y examinando la Fig. 3.8, se puede obtener una idea de las longitudes típicas de los tubos labiales abiertos de órgano, y de flautas y flautas dulces como función de la frecuencia. Un incremento en la frecuencia (altura) requiere una disminución en el largo. La relación (4.6) muestra también el efecto de la temperatura del aire sobre la altura fundamental de una columna de aire cilíndrica en vibración. Un incremento en la temperatura produce un aumento en la frecuencia (tono más agudo). Es así como las flautas y los tubos de órgano deben afinarse de acuerdo con la temperatura de la sala. Afortunadamente, la frecuencia fundamental (4.6) está controlada por la temperatura absoluta t_A , que aparece bajo una raíz cuadrada. Ambos hechos hacen que la influencia de las variaciones de temperatura sobre la altura sea pequeña, pero suficiente como para ser motivo de preocupación, como bien lo saben flautistas y organistas.

Veamos ahora el caso de un cilindro estrecho tapado en un extremo (Fig. 4.19). Notamos que, mientras en el extremo abierto P la presión debe permanecer constante e igual a la del aire exterior (nodo de presión), en el extremo cerrado Q la presión interna puede aumentar o disminuir sin restricción. Efectivamente, en Q aparece un *vientre de presión*. Esto es más fácil de entender cuando se visualiza el movimiento vibratorio real de los puntos del medio. Es evidente que debe haber un *nodo de vibración* para todos los elementos componentes de aire cercanos a Q : la tapa del tubo evita que esos elementos puedan oscilar longitudinalmente. Según lo explicado en la sección 3.3, tal nodo de vibración corresponde a un vientre de presión.

La figura 4.19 muestra cómo los modos vibratorios de ondas estacionarias «encajan» dentro de un tubo tapado de manera tal que siempre aparezca un nodo de presión en el extremo abierto y un vientre de presión en el extremo cerrado. Para la frecuencia fundamental, obtenemos la relación

$$f_1 = \frac{1}{4L} 20,1 \sqrt{t_A} = \frac{5,03}{L} \sqrt{t_A} \quad (4.7)$$

(L está en metros; t_A es la temperatura absoluta [3.5]). Esto es exactamente la mitad de la frecuencia fundamental (4.6) de un tubo abierto de la misma longitud. En otras palabras, un tubo cilíndrico tapado ideal suena una octava más abajo que un tubo similar abierto en ambos extremos.

Con respecto a los modos más altos de vibración de un tubo cilíndrico tapado, un examen de la Fig. 4.19 (y la conversión de longitud de onda a frecuencia) revela que sólo son posibles múltiplos *impares* de la frecuencia fundamental f_1 (4.7):

$$f_1, f_3 = 3f_1, f_5 = 5f_1, \dots \quad (4.8)$$

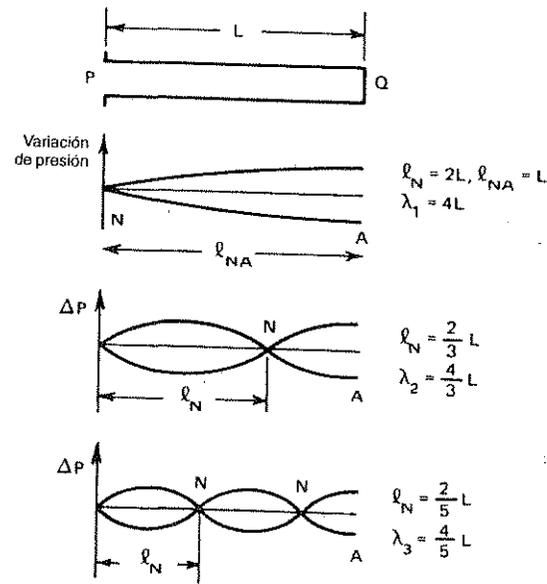


Figura 4.19 Modos de ondas estacionarias en un tubo idealmente cilíndrico, cerrado en un extremo.

Las frecuencias $2f_1, 4f_1, 6f_1, \dots$ están prohibidas - sus modos de vibración no pueden ser sostenidos de manera estable en un tubo cilíndrico tapado ideal. En otras palabras, *los armónicos de un tubo tapado son los armónicos impares de su fundamental*.

El clarinete es el ejemplo más familiar de un instrumento que se comporta de manera muy similar a un tubo cilíndrico tapado. La embocadura con la caña funciona como extremo tapado, la campana o el primer orificio destapado del tubo como extremo abierto. Por ello, la altura fundamental de una nota tocada en el clarinete está una octava más baja que la nota que corresponde a una columna de aire del mismo largo, tocada en una flauta.

Los órganos incluyen varios juegos de tubos tapados. Una de las razones es el ahorro de dinero y de espacio: los tubos abiertos graves son muy largos (según la relación [4.6], un tubo abierto de altura C_1 tiene un largo de 5,3m. El mismo tubo, tapado en un extremo, sólo necesita tener un

largo de 2,65m.) Desde luego, en esto hay algo más allá del mero costo: un tubo cerrado produce un sonido de muy diferente timbre al emitido por un tubo abierto de igual frecuencia fundamental (menos brillante, más oscuro).

El último caso que trataremos aquí es el de un tubo cónico (muy estrecho), tapado en el extremo P (Fig. 4.20). La determinación de los modos de vibración requiere un análisis matemático complejo. Los resultados pueden sintetizarse simplemente: un tubo idealizado estrecho, cónico y cerrado en la punta tiene los mismos modos de vibración que un tubo abierto, del mismo largo. En otras palabras, valen las relaciones (4.5) y (4.6). Un cono (estrecho) *truncado* (Fig. 4.21), cerrado en el extremo P , tiene una serie de modos de vibración que no guardan relación de números enteros; si bien en el rango de frecuencias más bajas (cerca de la fundamental) ellos corresponden bastante bien a los modos de vibración de un tubo abierto de la misma longitud L , para las frecuencias más altas ellos se aproximan a los de un tubo cilíndrico cerrado, de largo L . Dicho de otra manera, los modos de vibración son *inarmónicos*.



Figura 4.20 Tubo cónico.

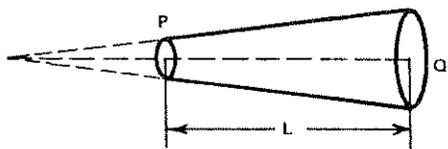


Figura 4.21 Tubo cónico truncado, cerrado en P.

4.5 Generación de vibraciones estacionarias complejas en instrumentos de viento

Los resultados de la sección precedente son idealizaciones que sólo valen

para columnas de aire hipotéticas (cilindros, conos), de diámetros muy pequeños con respecto a sus longitudes L . Este no es el caso, sin embargo, de los instrumentos musicales y tubos de órgano reales. Además, en estos instrumentos las columnas de aire son cilíndricas o cónicas sólo en una porción de su longitud, cobrando formas más complejas cerca de la embocadura y del extremo abierto (orificios destapados, campanas, etc.).

Con el fin de analizar cómo se comportan desde el punto de vista físico los instrumentos de viento reales debemos inspeccionar con mayor detalle todos los fenómenos involucrados. Consideremos primero el *mecanismo de excitación*. En el caso de los instrumentos de viento, no existe ningún equivalente a la acción de «puntear» o «percutir» una cuerda. La razón es que las oscilaciones de una columna de aire vibrando libremente decaen de manera prácticamente instantánea. Esto puede verificarse fácilmente, acercando la oreja a uno de los extremos de un tubo abierto (cuanto más largo mejor), y dando una palmada en el otro extremo o en la pared del tubo. Se podrá escuchar un sonido seco con altura igual a la frecuencia fundamental del tubo, pero decayendo en una fracción de segundo. Para mantener la vibración de una columna de aire, es necesaria la acción de un mecanismo primario de excitación equivalente al frotamiento de una cuerda, que suministre energía continuamente a una razón determinada.

Hay dos tipos de tales mecanismos. El primero consiste en una corriente de aire muy veloz que sopla contra un borde rígido, afilado E (Fig. 4.22), ubicado a una determinada distancia d de la abertura S . Este sistema es aerodinámicamente inestable: la corriente de aire se mueve alternativamente hacia uno u otro lado del borde, rompiéndose en remolinos de aire o «torbellinos». A medida que la velocidad de la corriente aumenta, también aumenta la frecuencia a la cual se forman los remolinos. Dado que ellos representan una perturbación periódica del aire del ambiente, cuando la frecuencia de generación de remolinos cae dentro del rango audible, se generan ondas sonoras. El sonido así generado se llama *tono de bisel* [edge tone].¹⁰ El mecanismo del «tono de bisel» es el proceso de excitación primario para todos los instrumentos de viento de la familia de la flauta y para los tubos labiales del órgano. Las oscilaciones de la corriente de aire son en general muy complejas; para intensidades del flujo muy bajas, aquéllas se hacen casi sinusoidales. La frecuencia fundamental de un «tono de bisel» libre (en ausencia de resonador) depende de la velocidad de la corriente de aire v y de la distancia al borde d (Fig. 4.22). En el rango de frecuencias bajas es proporcional a la razón v/d , es decir, aumenta al aumentar v y al disminuir d .

¹⁰ Los remolinos pueden formarse aun en ausencia de un borde afilado, siempre que la abertura S sea suficientemente pequeña y la velocidad v suficientemente alta. Esto representa la física básica del *silbido*, en que el tamaño de la abertura (apertura de labios) y la velocidad de la corriente de aire (presión del soplo) determinan la frecuencia fundamental.

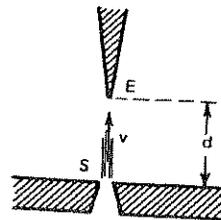


Figura 4.22 Generación de un tono de bisel.

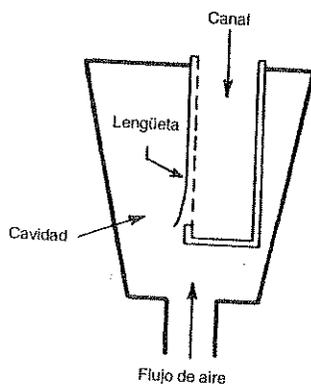


Figura 4.23 Mecanismo en un tubo de lengüeta de órgano.

El otro mecanismo de excitación de importancia para la música es la *lengüeta*, fina lámina de caña, plástico o metal, colocada frente a una abertura de forma similar y tamaño algo más pequeño (Fig. 4.23). Cuando entra aire en la cavidad soplando de abajo (o sea, a medida que la presión aumenta), el exceso de aire escapa hacia el canal a través del pequeño espacio que hay entre la lengüeta levemente levantada y la abertura. Durante este proceso la lengüeta es atraída hacia la abertura¹¹ y, en consecuencia, el flujo será eventualmente interrumpido. La propia elasticidad de la lengüeta abre la hendidura otra vez y todo el proceso empieza de nuevo. En otras palabras, la lengüeta empieza a oscilar, cerrando (parcial o totalmente) y abriendo alternativamente la abertura. El aire entra en soplos periódicos al canal, originando un sonido llamado *tono de lengüeta* [reed tone]. La frecuencia fundamental de un 'tono de lengüeta' libre depende tanto de las propiedades elásticas de la lengüeta

¹¹ Por la diferencia en la presión *dinámica* (no presión estática) entre ambos lados de la lengüeta - ¡el mismo principio que mantiene en vuelo a una aeronave!

como del exceso de presión en la cavidad (presión de soplo). En general, el movimiento vibratorio de una lengüeta libre es complejo, excepto para amplitudes muy pequeñas, para las cuales es casi sinusoidal. Algunos instrumentos (oboe, fagot) tienen *lengüetas dobles*, que se baten una contra la otra. También los labios del ejecutante de un instrumento de la familia de los metales pueden ser considerados como un sistema de doble lengüeta (de mucha masa).

El 'tono de bisel' y el 'tono de lengüeta' son raramente usados de manera aislada ('tono de bisel' y 'tono de lengüeta' libres). En las maderas ellos sirven meramente de mecanismo primario de excitación, abasteciendo de energía a la columna de aire del tubo. En esos casos no sólo el espectro, sino también la frecuencia de las vibraciones de la corriente de aire o de la lengüeta están controlados por la columna de aire, por medio de un mecanismo de retroalimentación (no lineal). Esto ocurre por acción de las ondas sonoras en el tubo: el primer pulso de compresión se propaga a lo largo del tubo, se refleja en el otro extremo (abierto o cerrado) y vuelve hacia la boquilla (como pulso de rarefacción en los tubos abiertos, como pulso de compresión en los cerrados). Esto causa una variación de presión que en el caso de las maderas sobrepasa a todas las otras fuerzas (aerodinámicas o elásticas), controlando de este modo el movimiento de la corriente de aire o de la lengüeta. La altura resultante es diferente (habitualmente mucho más baja) de aquella producida por el mecanismo del 'tono de bisel' o del 'tono de lengüeta' libres, en ausencia del tubo. Esto es muy distinto al caso de una cuerda montada sobre una tabla armónica, que prácticamente no afecta la frecuencia de la cuerda. En el caso de los instrumentos de la familia de los cobres, la masa de los labios del ejecutante es tan grande que la retroalimentación del tubo puede influir, pero no sobreponerse, a la vibración de los labios; esta última debe ser controlada por el ejecutante mismo a través de la tensión de sus labios. Hay unos pocos instrumentos de lengüetas abiertas (acordeón, armónica, órgano de lengüeta).

El proceso de formación de un tono en las maderas es muy complicado y todavía no completamente investigado. Sin embargo, es de capital importancia para la música. En muchos instrumentos, los armónicos superiores se forman antes que la fundamental; a veces, esto puede realizarse artificialmente, dándole un característico efecto de «pifiado» al tono resultante.

Para comprender la *parte estacionaria* en la generación de sonido en las maderas, bronces y tubos de órgano, es necesario analizar las propiedades resonantes de sus columnas de aire, y el acoplamiento de éstas al mecanismo primario de excitación (corriente de aire, lengüeta o labios). Con ese fin, partamos de los siguientes hechos verificados experimentalmente: 1) El mecanismo primario de excitación sostiene una oscilación periódica compleja, de cierta frecuencia fundamental, y

acompañada de una serie de armónicos que constituyen un determinado espectro. 2) La frecuencia fundamental y el espectro de las oscilaciones primarias están controlados (influidos en el caso de los instrumentos de bronce) por las propiedades de resonancia de la columna de aire; la amplitud total de las oscilaciones está determinada por el suministro primario de energía (flujo total de corriente de aire, presión de soplo). 3) El espectro de las oscilaciones de presión fuera del instrumento (onda sonora generada) está relacionado al espectro interno por una transformación gobernada por la forma y distribución detalladas de los orificios en la superficie del tubo y/o la forma de la campana (para un tratamiento más detallado, véase, por ejemplo, Benade, 1976).

Para explorar las propiedades de resonancia, es decir, la curva de resonancia de una columna de aire en un instrumento de viento determinado, se debe conducir experimentos similares al experimento con la cuerda metálica puesta en oscilación por medio de una corriente eléctrica alterna (Sec. 4.1). Para ello, se reemplaza el mecanismo natural de excitación por un generador mecánico de oscilaciones (por ej., un miniparlante) y se miden, con un pequeño micrófono, las amplitudes de las oscilaciones de presión en la embocadura (donde, en estas condiciones, se formará un vientre de presión). La curva de resonancia se obtiene graficando las amplitudes de las oscilaciones de presión en función de la frecuencia, para una amplitud constante de oscilación del productor [driver]. Las amplitudes medidas se expresan habitualmente en decibeles (R en expresión (4.4)), referidos a algún nivel standard. Curvas de este tipo también se llaman *gráficos de impedancia de entrada*.

La figura 4.24 muestra curvas de resonancia típicas para columnas de aire del tipo del clarinete y del oboe (Benade, 1971) (sin boquillas, campanas ni orificios abiertos). Es importante observar que los picos de resonancia obtenidos de esta manera corresponden a los modos de vibración de una columna de aire *cerrada* en el extremo del excitador primario, es decir, al caso real de una lengüeta actuando como mecanismo de excitación (vientre de presión, nodo de vibración, en el lugar de la lengüeta). Para hallar la curva de resonancia de la misma columna de aire para el caso en que ésta sea excitada por medio de una corriente de aire (flautas traverseras, flautas dulces, tubos labiales de órgano), basta trazar los negativos, $-R$, de los valores obtenidos en la medición anterior.¹² Los picos de resonancia se convierten en valles y los valles se hacen picos (simplemente dése vuelta a la figura 4.24). Lo que justifica esta operación es, básicamente, el hecho que en la boquilla, el vientre de presión (que aparece cuando se usa una lengüeta) es reemplazado por un vientre de vibración, o sea, por un nodo de presión, en el caso de una flauta.

¹² Sólo si R está expresada en decibeles.

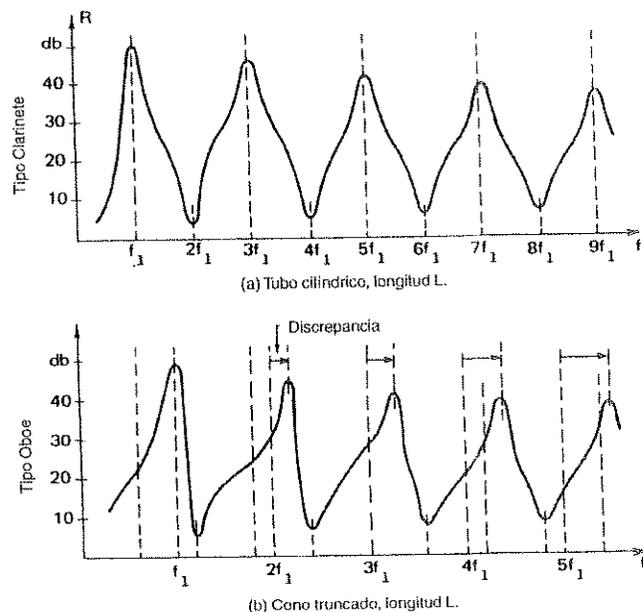


Figura 4.24 Curvas típicas de resonancia (desde Benade, 1971) de columnas de aire tipo clarinete (a) (cilíndricas) y tipo oboe (b) (cónicas) (sin boquillas ni campanas, con orificios tapados).

Obsérvese en la figura 4.24 que los picos de resonancia no son en absoluto «afilados»; los modos de oscilación posibles no corresponden, por lo tanto, a frecuencias únicas, discretas, como sería el caso en columnas de aire infinitamente angostas (Sec. 4.4). Aún más, en el caso del cono truncado (b), los picos de resonancia son asimétricos e inarmónicos (véase la discrepancia con respecto a los múltiplos armónicos $2f_1$, $3f_1$, etc.). Hacia las frecuencias más altas, los picos de resonancia del cono truncado se asemejan a aquellos del cilindro. Por otra parte, si el cono no fuera truncado, todos los picos de resonancia caerían cerca de los pozos de la curva (a) (armónicos pares del cilindro) con muy poca inarmonicidad.¹³

Analícemos, desde un punto de vista cualitativo, cómo la curva de resonancia controla al mecanismo primario de excitación; por ejemplo, una lengüeta. Para intensidades muy bajas (pequeñas amplitudes de la lengüeta) su movimiento es casi sinusoidal; en principio, podría excitarse la frecuencia de cualquier pico de resonancia (Fig. 4.24). En la práctica, sin embargo, se observa que a un nivel de intensidad muy bajo

¹³ En instrumentos reales, las curvas de resonancia del tipo del clarinete, curva (a), también revelan una discrepancia con relación a la armonicidad (Backus, 1974).

('pianissimo') solo es excitada la frecuencia que corresponde al pico de resonancia *más pronunciado* (alto). En general, este es el pico que tiene la frecuencia de resonancia más baja; el sonido resultante corresponde al «registro bajo» del instrumento.

A medida que se incrementa la amplitud de oscilación de la lengüeta (aumentando la presión del soplo), el carácter no lineal de la retroalimentación de la columna de aire destruye la vibración puramente sinusoidal de la lengüeta; van apareciendo armónicos superiores con cada vez mayor intensidad, y el sonido resultante adquiere más brillo. Al mismo tiempo, la frecuencia fundamental se reacomoda, en el caso en que los picos de resonancia superiores sean algo inarmónicos. La regla que gobierna esta reacomodación de altura es la siguiente: la frecuencia fundamental se ubica en una posición tal que *maximice el promedio ponderado de altitud de todos los valores de resonancia* R_1, R_2, R_3, \dots ¹⁴ correspondientes a los armónicos $f_1, 2f_1, 3f_1, \dots$ (Benade, 1971). Si, por ejemplo, los picos de resonancia en las frecuencias altas se desvían de la armonicidad como en la Fig. 4.24(b), la altura del tono deberá hacerse más aguda a medida que su intensidad aumente, con el fin de acomodar los cada vez más importantes armónicos superiores, tan cerca como sea posible a un pico de resonancia. Debido a este efecto, un cono truncado no se puede usar como instrumento de viento, a menos que se haga algo para reducir al mínimo la inarmonicidad.

Una situación interesante surge en el caso de una curva de resonancia del tipo del clarinete (Fig. 4.24[a]). Los picos de resonancia solamente corresponden a los múltiplos impares de la fundamental (véase también Sec. 4.4). Por ello, todos los armónicos pares presentes en la oscilación de la lengüeta serán notablemente atenuados. Partiendo en el registro grave con un 'piano-pianissimo' (excitación de una sola frecuencia en el pico de resonancia correspondiente a la fundamental) e incrementando paulatinamente la presión del soplo, se tenderá, en principio, a producir el segundo armónico, pero la energía de este último será eficientemente disipada a causa del valle en el lugar de la frecuencia correspondiente (Fig. 4.24 [a]). Para un mismo incremento en la presión del soplo, el aumento resultante en la sonoridad (y en el «brillo») será entonces considerablemente menor que en el caso de una columna de aire del tipo del oboe (curva [b]), donde el segundo armónico puede crecer sin obstáculos. Esta es la razón por la cual las transiciones del 'ppp' al 'pp' son más fáciles de manejar en el clarinete que en el oboe o en el saxofón.

Finalmente, otro hecho que vale la pena observar en la Fig. 4.24 es la ubicación casi idéntica de los *valles* en ambas curvas. Cuando estos valles son convertidos en picos, trazando los valores -R en lugar de R, para obtener las curvas de resonancia cuando estas columnas de aire son

excitadas con una corriente de aire (boquilla abierta), se obtiene una serie de armónicos prácticamente idéntica para ambos casos. Esto quiere decir que para instrumentos del tipo de la flauta, pueden usarse, casi indistintamente, tanto cilindros como conos truncados.

Hasta ahora hemos considerado ondas estacionarias en las cuales la frecuencia fundamental está determinada por el primer pico de la curva de resonancia (el de la frecuencia más baja), produciendo los tonos del registro grave de un instrumento de la familia de las maderas. En el «registro medio», la frecuencia fundamental cae cerca del segundo pico de resonancia. En los instrumentos de lengüeta, esto se consigue reduciendo el tamaño del primer pico de resonancia por debajo de aquel del segundo y apartando su posición de la serie de armónicos. La llave conocida como portavoz cumple esta función. En la flauta, esta transición, o «sobresoplo», se logra por medio de un cambio (un incremento) de la velocidad del aire del soplo contra el borde biselado. Obsérvese que en el primer 'sobresoplo' de un tubo del tipo del clarinete, excitado con lengüeta, la altura salta al tercer armónico o duodécima (segundo pico de la Fig. 4.24 [a]), mientras que tubos de lengüeta cónicos (y todas las flautas) tienen su primer sobresoplo a la octava (segundo pico en la Fig. 4.24 [b]).

Otro sobresoplo conduce al «registro agudo» de las maderas, con frecuencias fundamentales dadas por el tercer y/o cuarto pico de resonancia. Para lograr esto en un instrumento de lengüeta, los dos primeros picos de resonancia deben ser rebajados y cambiados en frecuencia para destruir su relación armónica.

Los *tubos de órgano* funcionan esencialmente sobre los mismos principios de la flauta (tubos labiales, abiertos y tapados) o de las maderas de lengüeta (tubos de lengüeta). La diferencia principal radica en que, como hay un tubo para cada nota de un juego dado, las llaves y el 'sobresoplo' no son necesarios. Los tubos del órgano son siempre activados en el registro grave (con algunas pocas excepciones en órganos «románticos»). Las curvas de resonancia de tubos labiales abiertos tienen picos en frecuencias ubicadas cerca de los múltiplos enteros de la frecuencia fundamental, con una leve inarmonicidad que depende de la razón $r = \text{diámetro/longitud}$. Las curvas de resonancia de tubos tapados se asemejan a aquella del gráfico superior de la Fig. 4.24, con máximos que corresponden a los múltiplos impares de la fundamental. Cuanto más grande sea el valor de r , tanto mayor será la inarmonicidad de las resonancias más altas. Como resultado, habrá un cambio en la frecuencia fundamental del tono resultante, más una creciente atenuación de los armónicos superiores (los cuales se alejan cada vez más de los picos de resonancia, inarmónicos). Como consecuencia de esto, *el sonido de tubos de órgano anchos es menos rico en armónicos superiores* (sonido cantante, «aflautado»). Tubos angostos (r pequeño) tienen picos de resonancia muy cercanos a los múltiplos enteros de la frecuencia fundamental, lo

¹⁴ Ponderados con los valores de intensidad espectrales correspondientes, I_1, I_2, I_3, \dots

que lleva a una eficiente excitación de los armónicos más altos (sonido brillante, tipo cuerda). La frecuencia fundamental está levemente desplazada con respecto al valor dado por la relación (4.6) (tubo abierto) o por la relación (4.7) (tubo tapado). Estas relaciones, sin embargo, pueden seguir siendo usadas si sumamos al largo L del tubo un valor de corrección, $0,3 \times$ diámetro, por cada terminación abierta («corrección del extremo»). Los tubos de órgano de lengüeta cubren un rango que abarca desde tipos donde la vibración de la lengüeta está muy controlada por la retroalimentación proveniente de la columna de aire (por ej., juegos de la familia de las trompetas) a tipos en que la vibración de la lengüeta es prácticamente autónoma (familia de juegos de regalía).

4.6 Espectros sonoros de instrumentos de viento

La resonancia característica de la columna de aire y el mecanismo de excitación colaboran para determinar el espectro de potencia y la intensidad de la onda estacionaria en el tubo de un instrumento. En la Fig. 4.24 vimos dos curvas de resonancia hipotéticas; los instrumentos reales, sin embargo, muestran una conducta más complicada debido a la peculiar forma de la boquilla (y de la embocadura), a la forma y distribución de los orificios abiertos, al efecto de la campana, y en el caso de las flautas, al efecto de la velocidad del aire sobre el ancho y la posición de los picos de resonancia (Benade, 1971). Aquí solo podemos resumir brevemente los efectos más importantes. Los *orificios*, además, desde luego, de determinar la longitud efectiva de la columna de aire y, de aquí, la posición absoluta de los picos de resonancia, son en parte responsables de un *corte* de los picos de resonancia por encima de los 1500 - 2000Hz. Este corte tiene un importante efecto sobre el timbre (atenuación de los armónicos superiores) y sobre el control dinámico de los sonidos '*forte*' en los instrumentos de madera, especialmente en los registros medio y agudo. En el oboe, la cavidad de la lengüeta y el angostamiento de la pieza a la cual la lengüeta va adosada contribuyen a reducir la inarmonicidad de las resonancias de un cono truncado.

Los bronceos merecen una especial atención en esta sección. Como ya se ha señalado, el mecanismo de realimentación es menos eficiente en la determinación de la frecuencia fundamental, y el ejecutante debe ajustar la frecuencia del zumbido de sus labios a la frecuencia que quiere obtener para lograr la altura correcta. En un instrumento de bronce, los armónicos superiores están creados por las propiedades de resonancia oscilantes de la boquilla causadas por el abrir y cerrar alternado de los labios (Backus y Hundley, 1971), más que por el movimiento no sinusoidal de estos últimos controlado por realimentación. La boquilla, embocadura cónica,

tubo principal cilíndrico, y campana en conjunto producen una curva de resonancia característica muy diferente de aquella de una madera típica. La Fig. 4.25 es un ejemplo (Benade, 1971). Obsérvese lo pronunciado de la frecuencia de corte (determinada principalmente por la campana) y el gran tamaño de la joroba de los picos y los valles en el rango de las frecuencias medias (originado principalmente por la forma de la boquilla). Esta joroba tiene un rol clave en la formación del timbre de los instrumentos de bronce. Finalmente, se observa que el primer pico de resonancia está *por debajo* de la frecuencia fundamental (marcada con la flecha) correspondiente al resto de los picos. Nótese también la asimetría característica de los picos en el rango de las frecuencias bajas (similar a aquellas del cono truncado, Fig. 4.24 [b]), en comparación con la región de las frecuencias altas (donde su forma se invierte).

Los bronceos no tienen orificios para alterar la longitud efectiva de sus columnas de aire; los cambios de altura son realizados principalmente

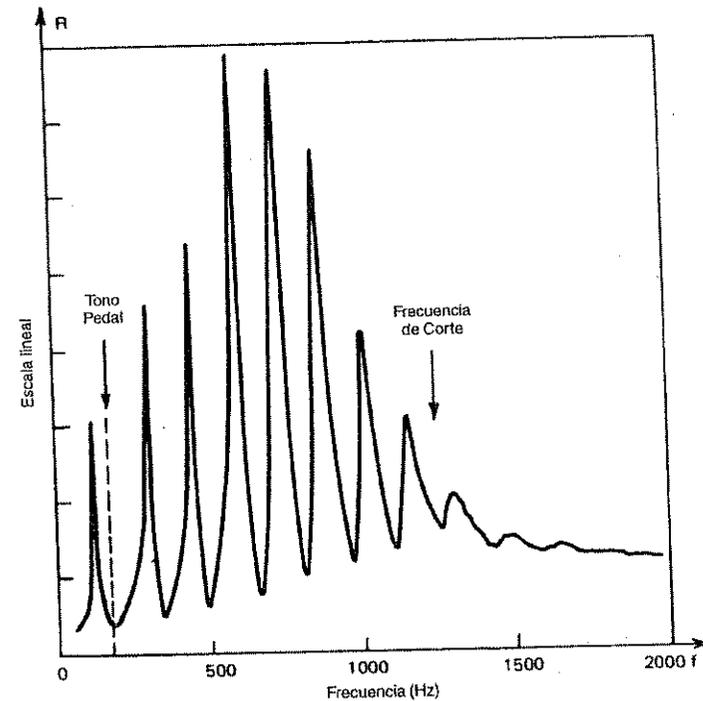


Figura 4.25 Curva de resonancia de una trompeta (Benade, 1971) (escala lineal). Con autorización del Profesor A. Benade, Case Western Reserve University, Cleveland, Ohio.

por 'sobresoplo', es decir, haciendo saltar la frecuencia fundamental de un pico a otro. Esto se logra ajustando apropiadamente la tensión de los labios. En la trompeta se puede alcanzar hasta el octavo modo y en el corno francés hasta el modo dieciséis. Para obtener notas entre picos de resonancia se dispone de un sistema de válvulas que ofrece una limitada selección entre longitudes de tubo levemente diferentes. En el trombón se puede efectuar un cambio continuo de la longitud del tubo (y por ende, la altura) a través de la vara. El pico de resonancia más bajo está desafinado con respecto al resto de la serie, casi armónica, de los picos, y no puede ser usado. Lo que se hace, más bien, es ajustar la frecuencia fundamental de la vibración de los labios al valor de la fundamental *ausente* (flecha en la Fig. 4.25) que pertenece a las frecuencias del pico segundo, tercero, etc. Esto lleva a la así llamada «*nota pedal*» de un instrumento de bronce (usada solo en el trombón). Esta nota puede tocarse únicamente a niveles bajos de sonoridad.

La composición espectral de las ondas sonoras de un instrumento de viento es diferente de aquella de las vibraciones estacionarias sostenidas en su propia columna de aire. La campana y/o los orificios abiertos son los principales responsables de esta transformación espectral. En esta transformación, la composición espectral por encima de la frecuencia de corte queda prácticamente sin alteración, mientras que los armónicos más bajos tienden a ser atenuados. En otras palabras, los espectros sonoros de instrumentos de madera y de bronce son en general más ricos en armónicos superiores que las vibraciones producidas en el interior del instrumento (Benade, 1973, 1976).

Los espectros de algunos instrumentos de viento presentan formantes, es decir, características que son independientes de la frecuencia fundamental del tono (Sec. 4.3). El fagot y el corno inglés son ejemplos, con reforzamientos espectrales (no muy bien definidos) alrededor de 450 y 1100 Hz, respectivamente. Estos formantes están causados por el espectro característico de la excitación por *dobles lengüetas*; no están determinados por las propiedades de resonancia del tubo del instrumento. Debemos mencionar aquí a la *voz humana* que, aunque no sea un tópico explícito de este libro, representa el ejemplo más notable de un «instrumento de viento» en el cual los formantes tienen un rol crucial: son la característica determinante de todos los *sonidos vocales*. En la voz humana, los formantes están principalmente determinados por las propiedades de resonancia de la cavidad nasofaríngea (Flanagan, 1972). La forma de esta cavidad determina cual de los dos rangos principales de frecuencia de las vibraciones de las cuerdas vocales van a ser reforzadas. Estas, a su vez, determinan si el sonido emitido es «a», «e», «i», «o», «u».

4.7 Atrapamiento y absorción de ondas sonoras en ambientes cerrados

Los instrumentos musicales se tocan habitualmente en habitaciones, salas, auditorios e iglesias. El sonido que percibe un oyente en estas condiciones no es en absoluto idéntico al sonido emitido por el instrumento. Por esta razón, el recinto donde se toca un instrumento puede ser considerado como una extensión natural de ese instrumento, con la diferencia de que, mientras un instrumento determinado tiene ciertas propiedades acústicas inmutables, aquellas del recinto pueden variar mucho de un lugar a otro, o de un momento al otro. El estudio de la acústica de las salas es tan importante para la música, como lo es la física de los instrumentos musicales.

Para estudiar el efecto de un recinto sobre una fuente sonora ubicada en algún lugar dentro del mismo, consideremos un instrumento musical en la posición S y un oyente en la posición L (que puede coincidir con S si el oyente es el ejecutante), en una habitación con paredes perfectamente reflectoras (Fig. 4.26). El instrumento empieza a tocar una nota determinada en el instante $t = 0$, manteniendo en todo momento su intensidad constante. Supongamos que el sonido es emitido por igual en todas las direcciones. A medida que las ondas sonoras se propagan alejándose de S , el oyente recibirá la primera señal después de un corto intervalo de tiempo SL/V ; este intervalo es el que tarda el sonido en llegar, propagándose en forma *directa*, desde S hasta L (por ejemplo, si $SL = 10\text{m}$, $V = 334\text{m/s}$. (relación [3.6]), el tiempo de llegada directa es 0,03s; para el ejecutante es prácticamente cero). Como veremos en la sección 5.1, el sonido directo tiene un rol fundamental en el proceso

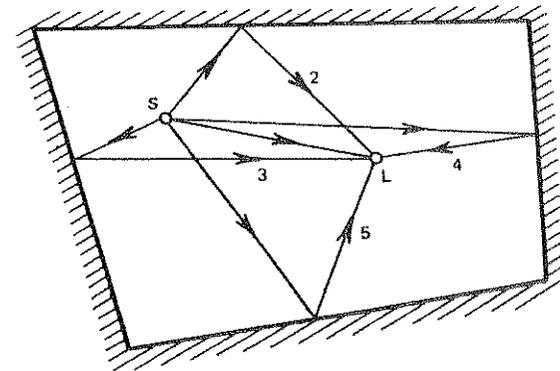


Figura 4.26 Diferentes caminos de propagación de una onda sonora, desde una fuente S a un oyente L .

perceptivo (*efecto de precedencia*). Inmediatamente después, pasarán por el punto L , en rápida sucesión, ondas reflejadas (trayectorias 2, 3, 4, 5, etc.). Cabe señalar que en la figura no se han tenido en cuenta las reflexiones en el piso y en el techo. Las primeras reflexiones, cuando son muy pronunciadas y están claramente separadas entre sí, se llaman *ecos*. Este proceso continúa con reflexiones secundarias, terciarias y múltiples (no mostradas en la figura). A medida que pasa el tiempo y el instrumento sigue sonando, la energía acústica que atraviesa L se irá reforzando. Si no hubiera absorción alguna, las ondas sonoras «llenarían» la habitación, y la energía acústica emitida por el instrumento se acumularía y quedaría atrapada en el recinto; la sonoridad, de este modo, se reforzaría gradualmente en cada punto del interior¹⁵. En un caso real, por supuesto, hay absorción cada vez que se refleja una onda. Por ello, la intensidad de la onda sonora no crece indefinidamente, sino que se nivela cuando la potencia disipada en los procesos de absorción es igual a la potencia emitida por la fuente (una situación similar al reforzamiento de una vibración en una cuerda frotada, Fig. 4.10). Este nivel de equilibrio de intensidad I_m del sonido difuso es mucho más alto que aquel del sonido directo (excepto en las cercanías de la fuente sonora).

Cuando la fuente sonora se calla, ocurre un proceso inverso: primero, desaparece el sonido directo, luego desaparecen las primeras reflexiones, luego las segundas, etc. La figura 4.27 muestra esquemáticamente cómo se comporta la intensidad sonora en un punto determinado dentro de un recinto típico. La caída del sonido, después que la fuente se ha callado, se llama *reverberación*, y representa un efecto de mayor importancia para la acústica

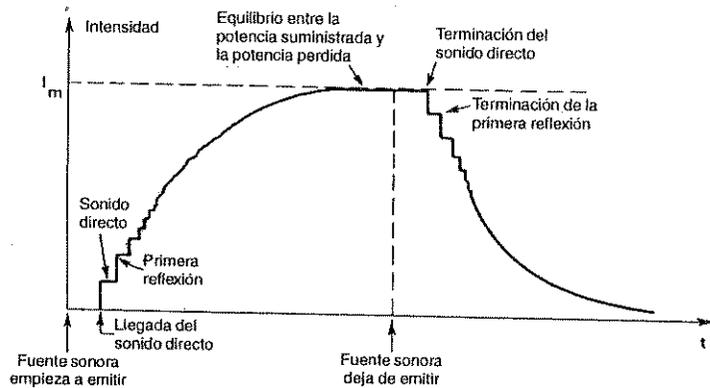


Figura 4.27 Crecimiento y caída típicas de la intensidad de un tono en una sala (escala lineal).

¹⁵ Para casi cualquier forma de recinto y posición de la fuente dentro de él, puede haber regiones prácticamente inaccesibles para las ondas sonoras emitidas desde S (puntos ciegos), o puede haber regiones hacia las cuales las ondas están focalizadas (por ejemplo, los puntos focales en recintos elípticos).

de salas. Esta caída es aproximadamente exponencial (por ej., véase Fig. 4.8); con algo de arbitrariedad se define como *tiempo de reverberación* al intervalo que tarda el nivel sonoro en decaer en 60db. Según la tabla 3.2, esto representa una disminución de la intensidad por un factor de un millón. Tiempos de reverberación deseables en buenas salas de concierto de tamaño mediano son del orden de 1,5 a 2 segundos. Tiempos más largos borrarían demasiado sucesiones típicas de sonidos; tiempos más cortos harían que la música sonara «seca» y tediosa (véase Sec. 4.8).

Examinemos algunas relaciones matemáticas simples que aparecen en el estudio de la acústica de salas. Imaginemos un recinto de paredes perfectamente reflectoras, sin absorción alguna, pero en las que se haya abierto un agujero de área A . En cuanto se haya alcanzado la intensidad máxima I_m (Fig. 4.27), la energía acústica estará escapando a través del agujero a razón dada por el producto $I_m A$ ¹⁶. Dado que esto corresponde al estado estacionario, en el cual la potencia P aportada por el instrumento iguala la razón a la cual se pierde energía, se puede establecer que $P = I_m A$, o

$$I_m = \frac{P}{A} \quad (4.9)$$

En los casos reales, por supuesto, no tenemos paredes perfectamente reflectoras con agujeros en ellas. Sin embargo, podemos *imaginar* una pared real con absorción, como hecha de un material perfectamente reflector pero con agujeros, estos últimos representando una fracción a de su superficie total. a se llama *coeficiente de absorción* del material de la pared. Una superficie de S metros cuadrados, de coeficiente de absorción a , tiene las mismas propiedades de absorción que una pared perfectamente reflectora con un agujero de área $A = Sa$. Los coeficientes de absorción dependen de la frecuencia del sonido (habitualmente son más altos para frecuencias más altas), y tienen valores cuyo rango va desde 0,01 (mármol, un reflector casi perfecto) hasta 0,9 (paneles acústicos). Teniendo en cuenta todo esto, podemos expresar la relación (4.9) en términos de las superficies reales de las paredes S_1, S_2, \dots con sus correspondientes coeficientes de absorción a_1, a_2, \dots :

$$I_m = \frac{P}{S_1 a_1 + S_2 a_2 + \dots} \quad (4.10)$$

Esta relación se puede usar para estimar las dimensiones necesarias de un auditorio, para lograr valores deseados de I_m para un determinado instrumento de potencia P , y una determinada distribución del material absorbente en las paredes.

El tiempo de reverberación τ_r resulta ser proporcional al volumen V de la sala e inversamente proporcional al área absorbente de las paredes,

¹⁶ Se supone que I_m representa aquí el flujo difuso, *omnidireccional* de energía sonora.

$A = S_1 a_1 + S_2 a_2 + \dots$ Experimentos muestran que, aproximadamente,

$$\tau_r = 0,16 \frac{V}{S_1 a_1 + S_2 a_2 + \dots} \quad (4.11)$$

donde V está dado en metros cúbicos, S en metros cuadrados y τ_r en segundos. Dado que los coeficientes de absorción en general aumentan con la frecuencia del sonido, τ_r decrecerá para alturas más agudas: las notas graves reverberan más tiempo que las agudas.

Uno de los problemas en la acústica de salas es que la *audiencia* influye considerablemente, aumentando las propiedades de absorción de esos ámbitos. Esto debe ser tenido en cuenta cuando se diseña un auditorio. Con el fin de minimizar los efectos de una audiencia de magnitud y distribución espacial imprevisibles, es aconsejable usar un tipo de butacas cuyo coeficiente de absorción sea independiente del hecho de estar ocupadas o no. El efecto de la audiencia sobre la absorción es especialmente grande en recintos con tiempos de reverberación muy largos, tales como iglesias y catedrales. Pocos ejecutantes están tan expuestos a cambios imprevisibles del entorno acústico (y hostigados por esos cambios) como un organista.

La distribución espacial, formación y decaimiento de un tono, así como la dependencia de los coeficientes de absorción de la frecuencia, influyen decididamente sobre las características físicas de los sonidos emitidos por un instrumento en entornos reales, y por ende, sobre la percepción de la música por parte del oyente. La dependencia temporal de los sonidos se ve muy afectada: se alteran las características transientes y, por ejemplo, una nota 'staccato', se alarga en el tiempo, dependiendo de las propiedades de reverberación de la sala. El espectro del tono también resulta afectado, dado que los coeficientes de absorción son dependientes de la frecuencia. Finalmente, teniendo en cuenta que las fases de las ondas que pasan por un punto determinado en un recinto reverberante están mezcladas en forma aleatoria, puede demostrarse que el SPL de cada armónico componente también fluctuará en forma aleatoria, imponiendo un límite a la capacidad del oyente de reconocer timbres en ambientes cerrados (Plomp y Steeneken, 1973).

Existen otros efectos de segundo orden, habitualmente ignorados, relacionados con un fenómeno ondulatorio conocido como *difracción*. Cuando una onda sonora incide sobre un obstáculo (por ejemplo, un pilar en una iglesia, o una persona sentada delante del oyente), pueden ocurrir tres situaciones: 1) Que la longitud de onda de la onda sonora sea mucho menor que el tamaño (diámetro) del obstáculo (por ej., un sonido agudo) (Fig. 4.28[a]). En este caso se formará una «sombra» acústica detrás del obstáculo y se tendrá una reflexión normal en el lado frontal. 2) Si el obstáculo y la longitud de onda son más o menos de la misma magnitud,

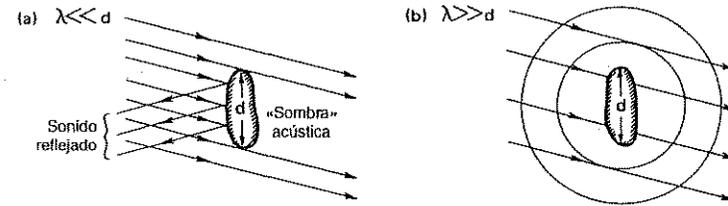


Figura 4.28 Interacción de una onda sonora con un obstáculo. (a): longitudes de onda pequeñas (reflexión y/o absorción); (b) longitudes de onda grandes (difracción).

se dará una situación más complicada, en la cual el objeto mismo actuará como un re-emisor sonoro, irradiando en todas las direcciones (no mostrado en la Fig. 4.28). 3) Si la longitud de onda es mucho mayor que el obstáculo (Fig. 4.28[b]) (por ej., sonidos graves), éste no afectará la onda sonora en absoluto, la cual se propagará prácticamente sin perturbaciones. La disposición regular de obstáculos (por ej., la de los asientos o de los oyentes en un auditorio) puede dar lugar a ciertas zonas de interferencia para determinadas longitudes de onda y direcciones de propagación. Finalmente, ciertas configuraciones del recinto, frecuencias y posiciones de la fuente pueden reforzar ondas estacionarias y conducir a la formación de molestos nodos y vientres dentro del recinto (Sec. 3.3). Para detalles científicos y técnicos de la acústica de salas, véase Ando (1985).

La difracción y el efecto de precedencia tienen un rol importante en la reproducción electroacústica de sonido. Una compensación electrónica de efectos de difracción de las ondas sonoras en la cabeza del oyente es conveniente cuando las señales estereofónicas son emitidas desde dos parlantes (Damasko, 1971). Para evitar que los auriculares den la sensación de una imagen sonora «dentro de la cabeza», es muy importante una correcta reproducción estereofónica del «sonido directo» (es decir, del efecto de precedencia) (véanse también págs. 53 y 71).

4.8 Percepción de la altura y el timbre de tonos musicales

Si bien se han realizado numerosas investigaciones sobre la percepción de altura y sonoridad de tonos simples (Secs. 2.3, 2.9, 3.4, y 3.5), mucho queda por hacer en el estudio de la percepción de tonos compuestos (por ej., véase Yost y Watson, 1987). Que el timbre de un tono pueda modificarse reforzando ciertos armónicos es algo conocido desde hace siglos. En realidad, la primera *síntesis de sonido* genuina ya fue realizada por los constructores de órganos de tubos en los siglos XIII y XIV. Los órganos de aquellos tiempos no tenían juegos múltiples; mas bien, cada tecla hacía sonar un número fijo de tubos llamados «Blockwerk», compuesto por uno o varios tubos afinados a la altura fundamental de la nota escrita, más

una serie de tubos afinados a la octava, la duodécima, la décimoquinta, etc., siguiendo la serie de armónicos superiores (excluyendo la séptima). La particular combinación de sonoridades elegidas para cada tubo componente determinaba la particular cualidad del sonido (timbre) del instrumento. Más tarde, aparecieron los registros múltiples: ellos permitieron al organista activar o desactivar selectivamente las diversas hileras de tubos correspondientes a los armónicos superiores del Blockwerk, y elegir, de este modo, entre varias opciones, el timbre del sonido del órgano (y alterar la sonoridad - Sec. 3.4). Fue sólo uno o dos siglos más tarde cuando se agregaron nuevos juegos independientes, en la forma de hileras de tubos de timbres diferentes entre sí.

La *síntesis sonora* es pues un asunto de antigua data. Sin embargo, el *análisis* del sonido, es decir la individualización de los armónicos superiores que aparecen simultáneamente en un tono producido naturalmente, no figura mencionada explícitamente en la literatura hasta 1636, cuando el notable físico, matemático y músico francés, Père Mersenne, publicó el primer estudio de análisis (cualitativo) de los armónicos superiores que aparecen en un tono compuesto.

Dos cuestiones fundamentales surgen con respecto a la percepción de tonos compuestos: 1) ¿Por qué un tono compuesto, formado de una superposición de diferentes frecuencias, da lugar a una *única* sensación de altura? 2) ¿Qué es lo que nos permite distinguir un espectro sonoro de otro, aun cuando la altura y la sonoridad sean la misma? Aunque hemos respondido, en parte, a la primera de estas preguntas (Sec. 2.9), es útil volver a examinar desde un comienzo el proceso de percepción de una onda sonora compleja que arriba al tímpano. El tímpano se moverá periódicamente para adentro y para afuera con un patrón vibratorio dictado por el patrón de vibración complejo de la onda. Este movimiento es transmitido mecánicamente por la cadena de huesecillos a la membrana de la ventana oval, que reproduce casi exactamente el mismo patrón complejo de vibración. Ni el tímpano ni la cadena de huesos «saben» que la vibración que están transmitiendo es producto de una superposición de diferentes armónicos. Este análisis recién se hace en el paso siguiente.

La vibración compleja de la ventana oval dispara ondas de propagación en el fluido coclear. Esta es la etapa en la cual tiene lugar la separación en diferentes componentes de frecuencia. Como vimos en las secciones 2.3 y 3.2, la región de resonancia para una componente dada (región de la membrana basilar donde una onda causa máxima excitación) está ubicada en una posición que depende de la frecuencia. Por lo tanto, un tono compuesto originará múltiples regiones de resonancia (Fig. 2.25), una para cada armónico, cuyas posiciones pueden obtenerse usando la Fig. 2.8. En vista de la relación casi logarítmica entre x y f , las regiones de resonancia se irán amontonando cada vez más a medida que progresamos en la serie de los armónicos (Fig. 2.25). Dado que cada región de resonancia

se extiende sobre cierta longitud (Sec. 2.4), regiones de resonancia vecinas empezarán a superponerse. En realidad, más allá del séptimo armónico, todas las regiones de resonancia están superpuestas, 'amontonadas' dentro de una banda crítica, lo que hace que no tenga demasiado sentido considerarlas individualmente.¹⁷ Cada una de las regiones de resonancia de la membrana basilar oscila con su propia frecuencia resonante (pág. 89), y con una fase que está relacionada (pero que no es igual) a la fase del armónico correspondiente (componente de Fourier) de la vibración original del tímpano. Las superposiciones, desde luego, complican considerablemente este cuadro.

Vemos que un solo tono compuesto suscita una situación extremadamente complicada en la cóclea. ¿Por qué, entonces, percibimos este tono como *una* entidad de características bien definidas en cuanto a altura, sonoridad y timbre? Como lo explicamos en la Sec. 2.9, esto puede deberse a la acción de un *proceso central de reconocimiento del patrón espacial*. El rasgo característico que se reconoce en este proceso - común a todos los sonidos periódicos, sin importar cuáles fueren su frecuencia fundamental y su espectro de Fourier - es *la relación invariante de distancias entre las regiones de máxima resonancia* sobre la membrana basilar. La sensación de altura debe ser considerada como la «señal de salida» de este proceso de reconocimiento¹⁸. Los componentes del tono compuesto más importantes para la activación del procesador central de altura, responsable de este proceso de reconocimiento, son los primeros seis a ocho armónicos. El mecanismo de reconocimiento del procesador de altura (el «ajuste de moldes» [«template fitting»] o la «autocorrelación espacial», descritos en la sección 2.9) puede funcionar aun si parte de la información de entrada está ausente (por ej., supresión de la fundamental). En tales casos es posible que se cometan errores de equiparamiento [error matching], o se produzcan sensaciones de altura ambiguas o múltiples (Sec. 2.7). Una discusión más detallada de cómo puede funcionar en términos neurales un proceso de reconocimiento se encuentra en la Sec. 4.10 y en el *Apéndice II*.

Debemos insistir una vez más que un análisis de la distribución temporal de los impulsos nerviosos (Sec. 2.9) también puede participar en la determinación de la altura subjetiva de un tono compuesto. En realidad, casi nos hemos olvidado del otro notable rasgo invariante de un patrón de excitación producido por un tono compuesto (Sec. 2.7): *pares de componentes consecutivos de una serie armónica producen una frecuencia de repetición común* (Fig. 2.19). Como se mencionó en la Sec. 2.8, un

¹⁷ Experimentos demostraron que para los instrumentos musicales comunes, debido a esta superposición (discriminación de frecuencia, Sec. 2.4), los armónicos más allá del séptimo no pueden ser discriminados auditivamente en forma separada (Plomp, 1964).

¹⁸ Recomendamos muy especialmente al lector que vuelva a leer la Sec. 2.9.

proceso basado en claves temporales debería funcionar mejor en el rango de los armónicos superiores, donde las regiones de resonancia vecinas están superpuestas, dando a las fibras nerviosas pertinentes la posibilidad de transmitir información sobre el patrón de vibración que resulta de la superposición (por ej., un patrón como el que se muestra en la Fig. 2.20). Por otra parte, como se indicó en la pág. 74, experimentos psicoacústicos realizados con tonos de frecuencia baja y muy corta duración también parecen favorecer la acción de un proceso basado en claves temporales en la percepción de altura: solo bastan dos o tres ciclos para que surja una clara sensación de altura, hecho que no puede ser explicado por medio de un mecanismo de análisis espectral.

La percepción de la altura de un tono compuesto es solo uno de varios resultados de una secuencia de tareas neurales. Es el mecanismo psicofísico que transforma el patrón de actividad periférico suscitado por un tono compuesto en otro patrón, de manera tal que todos los estímulos con las mismas regiones de resonancia, sin importar sus amplitudes, son representados de manera similar, produciendo la misma sensación de altura. Una analogía óptica sería el reconocimiento de una letra determinada, independientemente de su orientación, tamaño, color, o tipo. Cada vez que nuestro tímpano entra en una vibración periódica, aunque sea compleja, el hecho de que percibamos una altura nos dice que nuestro cerebro recibió el mensaje: «¡ajá!, un tono periódico» (en contraste con el ruido, o con la sensación de alturas múltiples o ambiguas al percibir un sonido inarmónico).

Nuestra reacción subjetiva a los tonos compuestos depende mucho del contexto del cual forman parte (Houtsma y Goldstein, 1972). El modo en que los tonos compuestos son procesados en el cerebro está influenciado fuertemente por la ejecución de otras tareas musicalmente «significativas», tales como la identificación de la fuente sonora, es decir del instrumento, y el reconocimiento de melodías y armonías. Esto incluso rige para la percepción de alturas: experimentos con componentes espectrales generados electrónicamente muestran de manera convincente que la percepción de la altura de estos sonidos compuestos se ve muy facilitada por (e incluso a veces requiere) la presentación de estos tonos de prueba en forma de una melodía¹⁹ (Houtsma y Goldstein, 1972). Los tonos compuestos sintetizados electrónicamente, fuera de un contexto musical, pueden conducir a sensaciones de altura múltiples o ambiguas (Sec. 2.7, pág. 59).

¹⁹ Un experimento notable para demostrar esta dependencia del contexto para el efecto de seguimiento de la fundamental se puede hacer en el órgano. Ejecute una pieza (por ej., el coral del *Orgelbüchlein* de Bach «Wenn wir in höchsten Nöten sein») tocando la melodía usando una combinación tipo cometa $8' + 4' + 2\frac{2}{3}' + 2' + 1\frac{3}{5}' + 1\frac{1}{3}' + 1'$ acompañada en *todo* momento con un suave $8' + 4'$ y un $16' + 8'$ respectivamente. Pida a una audiencia musicalmente entrenada que siga con atención las alturas de la melodía, pero advirtiéndole que habrá cambios de timbre. Después de tocar los primeros cinco a seis compases, repita la pieza, pero eliminando esta vez el juego de $8'$ de la melodía. Vuelva a repetir, eliminando

Al escuchar un tono compuesto, *nuestro sistema auditivo presta más atención al resultado originado en el mecanismo central de procesamiento de altura* (el cual produce una sensación de altura única) *que a la altura primaria de cada uno de los componentes armónicos*. Si queremos «escuchar por separado» (discriminar) los primeros seis o siete armónicos de un tono compuesto que suena de manera constante, debemos enviar una orden de «apagado» (inhibición) al mecanismo central de altura subjetiva, y focalizar nuestra atención sobre la información que suministra el más primitivo mecanismo de altura primaria o espectral, determinada por la posición espacial de las regiones individuales activadas en la membrana basilar (Sec. 2.3). Este proceso de inhibición y refocalización insume un tiempo considerable - mucho mayor que lo que tarda el mecanismo central de procesamiento tonal (Sec. 3.5). Esta es la razón por la cual los armónicos superiores no pueden ser «escuchados por separado» en tonos de corta duración o de decaimiento muy rápido.²⁰ Es importante observar que, en vista de la distribución asimétrica de la actividad a lo largo de la membrana basilar (por ej., Fig. 3.5), mediciones de apareamiento de alturas de los parciales de un tono compuesto *siempre conducen a intervalos levemente alargados*, por ejemplo, entre el primer y segundo armónico (octava alargada), y así sucesivamente (Terhardt, 1971). Este cambio se debe a la influencia perturbadora de todos los restantes armónicos sobre el armónico cuya altura primaria está siendo equiparada. El efecto es pequeño (menor de 2 - 3%), pero puede ser musicalmente importante (Secs. 5.4 y 5.5).

Desde luego, existen otros aspectos clave del estímulo auditivo primario (ignorados por el procesador de alturas) que llevan a resultados perceptivos en otras etapas del proceso de reconocimiento de sonidos. Al escuchar un tono compuesto, percibimos *sonoridad* (relacionada con el número total por segundo de impulsos nerviosos generados, Sec. 3.5) y *calidad del sonido o timbre*. Aquí debemos hacer una clara distinción entre la sensación estática, que aparece al escuchar un tono compuesto que no varía, de frecuencia fundamental, intensidad y espectro constante, de la situación más real, dinámica, al escuchar un sonido con características transientes dentro de un contexto musical significativo. Analicemos primero el caso

el 4'; luego el 2', y por último el 1'. Al final, muestre a su audiencia lo que quedó en la voz superior (sólo los juegos auxiliares de $2\frac{2}{3}' + 1\frac{3}{5}' + 1\frac{1}{3}'$) y señale que la altura de la nota escrita estuvo ausente en cada nota de la melodía (en cualquiera de sus octavas) - ¡les costará creerlo! Una repetición del experimento, sin embargo, probablemente falle ¡porque la audiencia reorientará sus estrategias de procesamiento de altura!

²⁰ El hecho de que el séptimo armónico sea una disonancia ha preocupado a los músicos durante mucho tiempo. Esta preocupación es, sin embargo, infundada. El séptimo armónico es extremadamente difícil de distinguir auditivamente por separado, incluso en tonos compuestos constantes, generados electrónicamente (Plomp, 1964). En efecto, algunos órganos modernos grandes tienen un juego auxiliar $1\frac{1}{7}'$ que suena el séptimo armónico de la nota escrita, el cual da un timbre muy particular cuando se usa juiciosamente con otros juegos.

estático. Experimentos psicoacústicos realizados con tonos compuestos, generados electrónicamente, que no varían en el tiempo, de igual altura y sonoridad pero con diferentes espectros y relaciones de fase entre armónicos, demuestran que la sensación de timbre está determinada principalmente por el espectro de potencia (Plomp, 1970, 1976). Los cambios de fase, si bien claramente perceptibles, - especialmente cuando ocurren entre los componentes de baja frecuencia -, solo desempeñan un papel secundario. La sensación *estática de timbre aparece como el correlato perceptivo de la distribución de la actividad producida a lo largo de la membrana basilar, siempre y cuando exista la correcta relación (armónica) de distancias entre los máximos* para unir todo en la sensación de un «único tono». Dividiendo el rango de frecuencias audibles en bandas de alrededor de un tercio de octava cada una (lo que corresponde aproximadamente a una banda crítica, Sec. 2.4), y midiendo la intensidad o flujo de energía sonora que para un tono compuesto está contenida en cada banda, fue posible definir «índices de disimilitud» cuantitativos para los sonidos (estáticos) de varios instrumentos musicales, que guardan buena correlación con los juicios de similitud y disimilitud tímbrica determinada psicofísicamente (Plomp y Steeneken, 1971). Es importante señalar que la sensación de timbre está controlada por la distribución *absoluta* de la energía sonora en las bandas críticas, no por valores de intensidad relativos a aquel de la fundamental (Grey y Gordon, 1978). Esto puede verificarse fácilmente escuchando una grabación en cinta reproducida a una velocidad que no sea la correcta. Esta operación no cambia los espectros de potencia relativa -simplemente desplaza a todas las frecuencias en un factor constante; aun así, se percibe un claro cambio de timbre en todos los instrumentos.

La sensación de timbre estática es una magnitud psicológica «multidimensional» relacionada no con uno, sino con todo un conjunto de parámetros del estímulo acústico original (el conjunto de intensidades en todas las bandas críticas)²¹ Esta es la principal razón por la cual descripciones semánticas de timbre son más difíciles de hacer que aquellas referidas a las magnitudes psicofísicas «unidimensionales» altura y sonoridad. Con excepción de algunos términos generales que van desde 'pobre' (pocos armónicos superiores), 'nasal' (principalmente armónicos impares), a 'brillante' (muchos armónicos superiores reforzados), la mayoría de las calificaciones dadas por los músicos invocan una comparación con tonos instrumentales reales ('aflautado', 'tipo cuerda', 'tipo bronce', 'tipo lengüeta', 'organístico', etc.).

²¹ Aunque hay alrededor de quince bandas críticas en el rango de frecuencias musicalmente relevantes, cuyas intensidades deberían ser especificadas con el fin de determinar el espectro, un estudio de identificación de vocales (Klein, Plomp, y Pols, 1970) indica que bastan solo *cuatro* parámetros de intensidad independientes (siendo cada uno una combinación lineal específica de las intensidades presentes en todas las bandas críticas) para especificar un tono compuesto.

Los aspectos dinámicos de la percepción del timbre están íntimamente relacionados con el proceso de identificación de la fuente sonora, es decir, con la identificación cognitiva del instrumento musical (Sec. 4.9). Los atributos correspondientes de las señales incidentes son mucho más difíciles de cuantificar (por ej., Iverson y Krumhansl, 1993); la envolvente de amplitud del tono, particularmente en el ataque y la caída, juega un papel fundamental. La similitud y disimilitud del timbre dependen más de los atributos gruesos, y más estáticos, del tono compuesto, tales como la frecuencia centroide (que no cambia en gran medida durante la duración del tono). La identificación cognitiva, en cambio, depende principalmente de atributos característicos del ataque del tono, atributos que contienen información crucial sobre el instrumento en cuestión, tal como el tipo de mecanismo de excitación (punteado, percutido, frotado, soplado). Evidentemente, la identificación ocurre a un nivel neural mucho más alto de funciones cognitivas del cerebro, que la sensación estática de la cualidad tonal o timbre (véase la próxima sección).

4.9 Identificación de sonidos musicales

Una propiedad característica de todos los procesos de reconocimiento de patrones («pattern recognition») es el «descarte selectivo de información». Esto se relaciona con el principio del 'mínimo esfuerzo - máxima eficiencia' con que opera el sistema nervioso. Para ser capaz de extraer del increíble volumen de información sensorial que traen los estímulos, aquella información que es *significativa* (y de este modo ser capaz de identificar los objetos y sus interrelaciones causales), el sistema debe contar con una serie de «filtros» que le ayuden a separar aspectos relevantes de los irrelevantes. Estos filtros deben estar «sintonizados» en rasgos o patrones invariantes de los estímulos considerados importantes (por traspasso de información genética, o como resultado de un proceso de aprendizaje). En la percepción sonora, el primer rasgo, el más primitivo, que tiende a ser reconocido por el sistema nervioso, es probablemente la intensidad, con la *sonoridad* como correlato perceptivo. Este correlato no depende de ningún rasgo o detalle fino de la estructura del estímulo, sino del flujo total de energía acústica (Sec. 3.5). El siguiente rasgo a ser considerado es la periodicidad de un sonido (representada por la distribución espacial de los máximos de resonancia o por la distribución temporal de los impulsos nerviosos). En este proceso de reconocimiento se extrae la sensación subjetiva de *altura*. El tercer nivel de refinamiento es la consideración del espectro de potencia del tono, que lleva a la sensación (estática) de *timbre*.

La percepción del timbre es, sin embargo, sólo la primera etapa en la operación de *reconocimiento de la fuente sonora*, que en música es la

identificación del instrumento. Desde este punto de vista, la percepción de la cualidad tonal es el mecanismo por medio del cual se extrae información de la señal auditiva adecuada para: 1) *almacenamiento* en la memoria con un *adecuado rótulo de identificación*, y 2) *comparación* con información previamente almacenada e identificada. La primera operación involucra *aprendizaje* o condicionamiento. A un niño que aprende a reconocer un instrumento musical determinado se le hace escuchar sucesivas veces una melodía tocada por ese instrumento, mientras se le dice: «esto es un clarinete». Su cerebro extrae información apropiada de la sucesión de estímulos auditivos, rotula esta información con el título de «clarinete» y la almacena en la memoria. La segunda operación representa la respuesta condicionada a un patrón aprendido: cuando el niño escucha un clarinete después de la experiencia de aprendizaje, su cerebro compara la información extraída de la señal que llega (es decir el timbre) con claves previamente almacenadas, y si logra una equiparación dirá: «un clarinete» (para más detalle, véase Sec. 4.10). Por otra parte, si escuchamos un sonido «nuevo», por ejemplo una serie de tonos hechos con un sintetizador electrónico, nuestro sistema de extracción de información alimentará al mecanismo de equiparación, el cual tratará desesperadamente de comparar la información que llega con aquella previamente almacenada. Si en este proceso no hay apareamiento, eventualmente se abrirá un nuevo «archivo» de almacenamiento para esta nueva cualidad sonora, ahora identificada. Si la equiparación se cumple solo parcialmente, reaccionamos con juicios del tipo «casi como un clarinete» o «un trombón que ladra».

Los músicos protestarán y dirán que la sensación del timbre involucra mucho más que el mero encontrar claves para descubrir «qué instrumento está tocando». Por ejemplo: ¿qué es lo que hace que un instrumento suene mucho más bellamente que otro del mismo tipo? Primero deberíamos señalar que esto está relacionado, obviamente, con un mayor grado de sofisticación del mecanismo de identificación arriba mencionado; podemos aprender a extraer de los patrones de vibración sonoros información cada vez más refinada, lo que nos capacita para distinguir entre distintas muestras de instrumentos de la misma clase. El porqué algunos patrones vibratorios nos parecen ser más bellos que otros es algo que, en realidad, no se sabe. Se han realizado muchas investigaciones para descubrir cuáles son las características que hacen de un violín Stradivarius un gran instrumento (por ej. Saunders, 1946). Muchas de estas características están relacionadas con aspectos dinámicos, y la mayoría de ellas parece estar más relacionada con la mayor o menor facilidad con la que el *ejecutante* puede controlar el «color» del tono deseado (espectro y transientes) que con un efecto «pasivo» sobre el oyente. Por ejemplo, un aspecto significativo de los sonidos de violín parece estar relacionado con el efecto de los picos de resonancia

estrechamente espaciados (por ej., Fig. 4.15) sobre la sonoridad y el timbre, cuando la frecuencia fundamental del tono es modulada por el ejecutante por medio de un *vibrato* (Matthews y Kohut, 1973). En tales circunstancias, las frecuencias de los componentes armónicos periódicamente barren esos picos de resonancia. Como resultado, la amplificación de cada componente variará periódicamente, y lo mismo ocurrirá con la sonoridad y el timbre del tono. Según sea la particular microestructura de la curva de resonancia de su instrumento, *el ejecutante de un instrumento de cuerda puede producir con su vibrato cambios extremadamente finos en la sonoridad y el timbre*. En gran medida, la impresión sobre el oyente se basa también en su experiencia: ¡desde el comienzo de nuestra educación musical *se nos ha dicho* qué tonos deben ser rotulados como bellos y cuáles no! Y nuestro cerebro ha robustecido los patrones de identificación correspondientes, basándose en un procesamiento de información tremendamente refinado, que le permite reconocer aspectos pequeñísimos, prácticamente inmensurables, de la estructura de la señal acústica recibida.

4.10 Procesos cognitivos relevantes a la percepción de tonos individuales

Terminaremos este capítulo con una breve reseña de los procesos cerebrales de más alto nivel involucrados en el reconocimiento de los sonidos musicales. Comenzaremos tratando algunos conceptos básicos, tales como vida, información y evolución de la función cerebral en los seres humanos (Roederer, 1978, 1979).

Los sistemas animados, desde los microorganismos hasta los primates, son «islas» de materia organizada que evolucionan persistentemente hacia un orden cada vez mayor. Para ser capaz de generar orden, un sistema biológico debe mantenerse a sí mismo invariante durante un cierto intervalo finito de tiempo, a pesar de interacciones con un entorno cambiante. Debe operar de un modo orientado hacia el futuro, es decir, mantener su funcionalidad siguiendo cursos de acción que se adapten favorablemente a los cambios del entorno, y ser capaz de efectuar cambios en este último que le sean de su propio beneficio. Esta es la diferencia más significativa entre los sistemas animados y los abióticos. Está caracterizada por el hecho que las interacciones de los sistemas vivientes con el entorno y con otros sistemas vivientes están basadas en procesos que involucran reconocimiento de patrones («pattern recognition») y procesamiento de información.

El reconocimiento de patrones requiere la generación de idénticas respuestas a diferentes constelaciones complejas de estímulos [input] que contengan un cierto rasgo común, espacial o temporal (el patrón que está siendo reconocido), independientemente del lugar, tamaño, orientación u orden particular en que ese rasgo aparece (por ej. una disposición espacial

particular de los radicales nucleótidos en la molécula de DNA ejerciendo un efecto específico sobre su entorno químico; los rasgos geométricos del contorno de un objeto cuya imagen es proyectada sobre la retina, disparando operaciones cognitivas específicas; el conjunto de armónicos de un tono musical que produce específicas zonas de resonancia sobre la membrana basilar, dando lugar a una específica sensación de altura). En última instancia, es este modo de procesar información lo que permite a un sistema viviente mantener orden en forma consistente y pasar información pertinente a sus descendientes.

A medida que las especies evolucionaron, información sobre el medio ambiente se fue incorporando gradualmente en las estructuras de memoria genética de los organismos. Cambios muy lentos en el entorno, que insumen tiempos mucho mayores que el tiempo de vida de una generación, pudieron ser incorporados en las estructuras de memoria de los genes, a través de la acción combinada de mutación, diversificación y selección natural. Cuando las especies se tornaron más complejas, y cuando características cada vez más imprevisibles del ambiente terrestre aparecieron como factores determinantes para la supervivencia, la capacidad de adaptación ontogénica, durante el lapso de vida de un mismo organismo, llegó a ser un requerimiento fundamental para la supervivencia.

Cuando surgió la locomoción en los organismos multicelulares, el número de variables importantes en el entorno aumentó drásticamente, con cambios en escala de tiempo de fracciones de segundo; se hizo necesario absorber una increíble cantidad de información a través de los sentidos. Aunque la mayor parte de esta afluencia de información sensorial es irrelevante, lleva incluidas de una manera 'a priori' impredecible, aquellas señales de patrones que son decisivos para la supervivencia del organismo. El sistema nervioso evolucionó con el fin de dotar a los organismos superiores de la capacidad de detectar, clasificar, identificar y almacenar la información relevante contenida en un complejo total de estímulos que llegan al organismo, y anticipar y reaccionar apropiadamente a los rápidos cambios del entorno. En el curso de este desarrollo, lo que empezó como un simple aparato de conversión y transmisión de señales ambientales, se convirtió en el sistema nervioso central (SNC) de los vertebrados superiores, con sus sofisticadas capacidades de análisis y elaboración de respuestas basadas en el estado momentáneo del entorno, del organismo, de las instrucciones innatas (instintos), y de la información aprendida (Fig. 4.29). El cerebro surgió como el procesador central que lleva a cabo las operaciones necesarias para controlar las condiciones vitales del organismo y optimizar su rendimiento en un entorno complejo y cambiante.

Un desarrollo crucial fué, finalmente, el surgimiento del cerebro humano. Ya Aristóteles señaló que «los animales tienen memoria y son capaces de aprender, pero ningún otro animal salvo el hombre puede

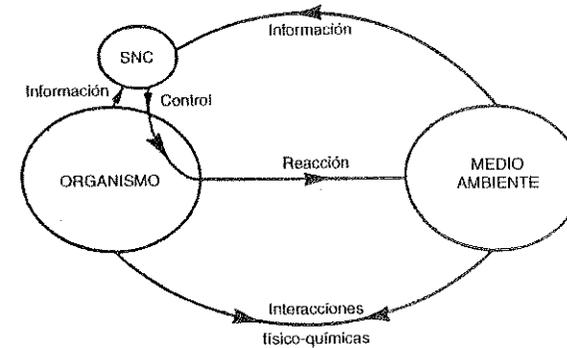


Figura 4.29 Información y control en el Sistema Nervioso Central (Roederer, 1987).

recordar el pasado a voluntad». En términos más modernos y específicos, la operación más característica que puede ejecutar el cerebro humano, y sólo éste, es la de recobrar información, imágenes o representaciones almacenadas, analizarlas, manipularlas y almacenar en memoria versiones modificadas o corregidas de las mismas, *sin* la concurrencia simultánea de estimulación externa. Llamamos a estas acciones de alteración de imágenes y su realmacenamiento sin concurrente información exterior, «el proceso del pensar humano». Este implica más que el mero procesamiento de información existente: representa la generación de información nueva (Roederer, 1978). Ligado a este proceso surgieron el habla y otros aspectos distintivos del lenguaje humano.

La capacidad de reacomodar y alterar imágenes condujo a la posibilidad de realizar predicciones y planificar a *largo plazo*, y a la conciencia del futuro. Con esta capacidad se hizo posible el posponer metas de conducta, o reacomodar órdenes de prioridad: el organismo empezó a servir al cerebro y no lo contrario. Finalmente, el almacenamiento planificado y deliberado de información en el medio ambiente para uso posterior es otro rasgo característico de las capacidades del cerebro humano (Roederer, 1978); esta «externalización de la memoria» llevó a la documentación de hechos y sentimientos a través del lenguaje escrito y de las artes y, de este modo, al desarrollo de la cultura humana como tal.²²

²² Los animales, desde luego, también dejan información «estampada» en el entorno. Sin embargo, hay una diferencia fundamental. Si repentinamente destruyéramos todos los vestigios de información dejada por los animales sobre la Tierra (nidos, reservas de alimento, huellas, etc.) la vida animal podría, esencialmente, continuar como si fuera hoy: la información clave para la supervivencia y la preservación está almacenada genéticamente en el organismo; los vestigios de «cultura» animal no son usados por generaciones posteriores como fuentes de información. Si en cambio destruyéramos de un golpe todos los vestigios de información dejada por los seres humanos, nuestra civilización desaparecería tal como hoy la conocemos, y las generaciones futuras quedarían en un estado similar a aquel en que estuvo la humanidad decenas de miles de años atrás.



Durante las dos últimas décadas hubo grandes progresos en el conocimiento de las funciones cerebrales. Mucho se fue logrando de a poco, y lo que hay en este momento es más bien un conjunto de hipótesis científicas que un modelo cuantitativo, experimentalmente verificado, de la función cerebral.

Antes que nada, examinemos qué clase de información debe ser reconocida en los diversos sistemas sensoriales. En el sentido de la vista, son las entidades discretas que llamamos «objetos en el espacio», con límites, superficies y texturas habitualmente bien definidas, y sus mutuas relaciones en el espacio y cambios en el tiempo. El sentido de la audición reconoce «objetos en el espacio acústico» (que en realidad son «objetos en el tiempo»), es decir, trenes discretos de ondas sonoras de rasgos bien definidos, y las relaciones que guardan entre sí y con las fuentes en el medio ambiente que los originan. Ya hemos mencionado repetidas veces que la percepción sensorial procede en varias etapas fisiológicas y operacionales. Las señales externas son primero convertidas en señales nerviosas que luego son transmitidas a las áreas corticales de recepción primaria, desde donde son transferidas a áreas secundarias y de asociación, los lóbulos frontales y al resto del cerebro. La figura 4.30 bosqueja un cuadro muy simplificado de los principales pasos del procesamiento de la información sensorial en el sistema visual; etapas equivalentes existen en el sistema acústico (Roederer, 1987).

Los circuitos neurales en la periferia y en las vías aferentes hasta las áreas corticales de recepción primaria, están en su mayor parte «cableados» (en forma innata) para el preprocesamiento de la detección de rasgos (por ej., en visión, detección de límites y movimiento, fusión binocular; en la audición, altura espectral y detección de transientes, audición binaural). Se han realizado mediciones que muestran que las neuronas en las áreas corticales, a las cuales está ligado el sistema de transmisión aferente de un órgano sensorial, son en realidad «detectores de rasgos», que responden a rasgos físicos bien definidos, pero complejos, del estímulo sensorial original. En la corteza visual, por ejemplo, se han descubierto neuronas que sólo responden a bordes o rayas oscuras o claras en una parte determinada del campo visual, inclinadas con un ángulo determinado, o a una línea moviéndose de determinada manera (Hubel, 1971). En la corteza auditiva, también, hay neuronas que parecen responder sólo a ciertos tipos de sonidos complejos (Simmons, 1970). La integración de rasgos («feature integration» o «binding») es la operación necesaria para extraer de la complejísima totalidad de información sensorial aquellos rasgos que pertenecen a un mismo objeto (por ej., unir aquellos bordes que definen los límites de un objeto dado; separar aquellas regiones de resonancia sobre la membrana basilar que pertenecen a un tono compuesto determinado). Esta operación involucra también transformaciones geométricas y topológicas (lo más probable, operaciones aprendidas) que

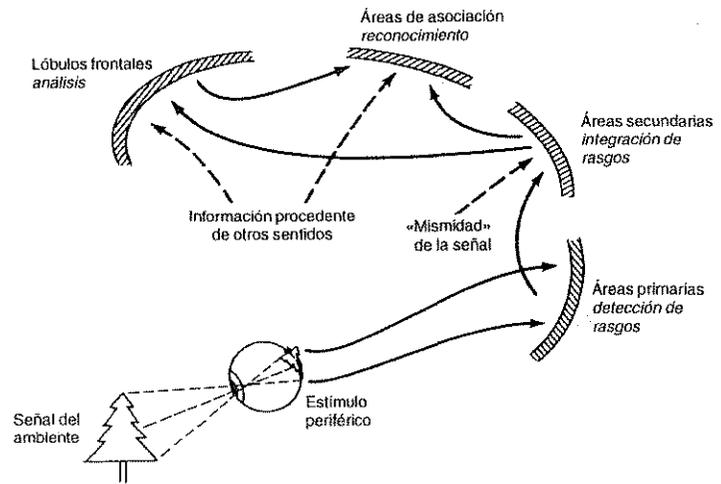


Figura 4.30 Etapas en la percepción de un objeto (según Roederer, 1987).

informan al cerebro sobre la «unicidad» o invariancia de un objeto visual dado, cuando éste es visto desde diversos ángulos o trasladado en el espacio, o de un objeto acústico, tal como un tono compuesto, cuando se cambia la altura fundamental o la intensidad de éste (el proceso de ajuste de moldes mencionado en la pág. 72.). En esta etapa el cerebro «sabe» que se trata de un objeto; el reconocimiento, es decir, la determinación de cuál es el objeto en cuestión, requiere un proceso ulterior de comparación con información ya existente, previamente adquirida (identificación del objeto; identificación del instrumento al cual pertenece el tono). El reconocimiento puede ser «automático» (memoria asociativa) o requerir un análisis de la totalidad de información sensorial que llega a los lóbulos frontales. La etapa final corresponde a la integración de los objetos reconocidos en la escena total percibida (el paisaje que se está viendo, la música que se está escuchando).

A medida que avanzamos a través de los distintos estadios de la figura 4.30, el procesamiento de la información se hace menos automático y cada vez más centralmente controlado; más necesarias se tornan las acciones y decisiones controladas por motivación (atención, eliminación de información irrelevante), y la información previamente almacenada (aprendida) influirá más el proceso. En las etapas superiores los aspectos evocativos afectan más y más cada fase del proceso cognitivo. El automatismo de los estadios inferiores del procesamiento de información cede a un control basado en la motivación de ejecutar o no una dada tarea. En los seres humanos, este control por motivación puede a su vez ser

influenciado por operaciones cognitivas de orden superior, pero no puede ser *eliminado* totalmente del proceso (véase Sec. 5.6). Esta es, tal vez, la diferencia más importante entre el procesamiento de información sensorial por robots de «inteligencia artificial» y las «operaciones cognitivas reales» del cerebro: el último tiene un propósito primordial muy diferente de aquel de los robots y computadoras: ¡optimizar las condiciones para la supervivencia!

¿Qué es lo que realmente ocurre en el sistema nervioso, desde el punto de vista físico, cuando la información sensorial es procesada y los objetos son representados, reconocidos e integrados en la escena ambiental? ¿Qué ocurre cuando objetos y eventos son imaginados? En la sección 2.9 señalamos que la información sensorial (y *cualquier* otro tipo de información) es codificada en el sistema nervioso en forma de una distribución espacial y temporal de impulsos eléctricos.²³ Definiremos como la «representación» mental o cerebral de un objeto (visual, acústico, olfativo, táctil) a la distribución específica de señales eléctricas en la red neural de la corteza cerebral, que aparece en correspondencia causal, unívoca con el conjunto de rasgos recibidos por los sentidos durante la presentación de ese objeto en particular (Roederer, 1979). Según esta definición, la «cognición» no es más que la aparición de una distribución específica de actividad neural en ciertas áreas de la corteza que está en una correspondencia unívoca con el objeto, evento o concepto que está siendo reconocido o imaginado. Por ejemplo, las distribuciones de actividad nerviosa que surgen en ciertas áreas corticales por la percepción de los siguientes objetos: una gran manzana roja, un montón de manzanas verdes, un pastel de manzana, un vergel de manzanos, aunque bien diferentes, contendrían todas un subconjunto de actividad neural cuya distribución espacio-temporal sería la misma, a saber, aquella que aparece en correspondencia con, y que define, la cognición de «manzana». Cada vez que reconocemos los tonos de un clarinete (pág. 171), está ocurriendo una distribución de actividad neural única, y que es la misma para todos los tonos de clarinete.

El acto de recordar un suceso consiste en la reactivación o «replay» de esa particular distribución o patrón de señales neurales que fueron específicas del evento sensorial original. Cuando este patrón es disparado por acción externa, por ejemplo, escuchando un tono, reconocemos o recordamos que este tono proviene, digamos, de un clarinete. Cuando esta actividad es disparada internamente (por una asociación o por un

²³ Es importante señalar que también existe un sistema de información *neuroquímico* (péptidos cerebrales, endorfinas, etc., liberados por un mecanismo nervioso de secreción especializado) que regula el estado general de la función cerebral (y somática), actuando como una suerte de «control de volumen» sobre ciertos centros específicos del cerebro. Este sistema de neuromoduladores es «lento» y, en realidad, sólo *transmite* información (a través de la corriente sanguínea); no procesa información.

acto de voluntad), estamos recordando el sonido de un clarinete en ausencia de un sonido externo verdadero. Esto representa la forma más simple de activación del *mecanismo de imaginación acústica*. Experimentos realizados en el campo visual han mostrado que, por ejemplo, la mera imaginación de una forma geométrica provoca una actividad en la corteza visual muy similar a aquella producida cuando el sujeto efectivamente ve esa forma (Herrington y Schneidau, 1968). Hoy existe evidencia experimental de que todo el proceso bosquejado en la Fig. 4.30 puede operar en sentido inverso: cuando se nos dice que imaginemos una manzana roja, se envía un «estímulo» apropiado desde los lóbulos frontales hacia las áreas de asociación y recepción visual, donde se despliega casi la misma actividad neural que sería evocada si estuviéramos realmente viendo esa manzana. Un compositor que «imagina» una secuencia de sonidos evoca una actividad nerviosa específica en las áreas de procesamiento acústico idéntica a la que aparecería si realmente escuchara esa secuencia. La actividad neural específica que se presenta en un proceso imaginativo también puede ser disparada por un proceso alucinatorio, durante un sueño, o por medio de una estimulación electrofisiológica del cerebro durante una neurocirugía.

Mientras que en las etapas periféricas e inferiores del procesamiento de información nerviosa, grupos relativamente pequeños de neuronas espacialmente contiguas ejecutan tareas definidas (por ej., detectores del «miau» en la vía auditiva del gato), en las etapas más avanzadas, incluso la más simple de las representaciones demanda la intervención de *cientos de millones* de neuronas corticales de manera simultánea y cooperativa, y no hay contigüidad espacial alguna en la distribución de la actividad correspondiente.²⁴ Estos hechos hacen que la verificación experimental de los procesos de representación neural de cognición sea una empresa extremadamente difícil. Hay experimentos, sin embargo, basados en el empleo de microelectrodos y la registración de potenciales evocados de patrones macroscópicos coherentes de actividad nerviosa en el cerebro, y más recientemente, en técnicas de resonancia magnética nuclear y tomografía por emisión de positrones, que dan soporte sustancial a esta descripción.

²⁴ Esto significa que cuando se imagina, digamos, un árbol, la «distribución de actividad neural específica» arriba mencionada *no* es un área con forma de árbol en las áreas corticales visuales de los lóbulos occipitales, sino un conglomerado monstruosamente complejo, discontinuo y distribuido de neuronas disparando impulsos. Es la *especificidad* de esta distribución monstruosamente compleja, que es única y que efectivamente *representa* la imaginación (o percepción real) de ese árbol. ¿No es maravilloso el hecho de que no tengamos que preocuparnos de ese horrible «caos altamente organizado» que ocurre en nuestro cerebro? Ese hecho se llama «simplicidad natural de la función mental» ¿No es lógico, por otra parte, que para comprender con claridad el proceso neural de la representación mental sea necesario tener familiaridad, propia de un físico, con conceptos tales como funciones de distribución y sus transformaciones, procesos no lineales, fenómenos cooperativos y teoría del caos?

¿Cómo funciona todo esto? Desde que los organismos superiores tuvieron que contar más y más con la información adquirida a lo largo del tiempo de sus propias vidas y almacenarla en el cerebro, fue surgiendo la necesidad de una memoria adecuadamente protegida y de rápido acceso. Esto condujo a una memoria distribuida y al modo de *representación holológica* (Pribram, 1971; Kohonen, 1988), en lugar de una codificación «fotográfica» de las escenas ambientales. Una representación fotográfica sería una en la cual hay una correspondencia «punto a punto» entre los rasgos del estímulo (objeto) y los rasgos de la distribución de la actividad neural (imagen). En una representación holográfica u «holológica» en nuestro caso, la información sobre cada rasgo (punto) del estímulo-objeto es mapeado sobre el dominio entero de la imagen (como en un holograma óptico, donde la información sobre un punto del objeto es distribuida sobre todo el dominio de la imagen en la película fotográfica). Un daño parcial en una imagen fotográfica puede conducir a la pérdida total de la información (por ej., recortar la cara de la fotografía de una persona), mientras que el daño parcial a un holograma conduce normalmente sólo a un borronero, pero no a una eliminación completa de información en la imagen.

Ya hemos señalado en la página 77 que en los estadios periféricos de la vía auditiva hay una organización tonotópica de las señales nerviosas, en correspondencia espacial (y temporal) con las regiones de excitación sobre la membrana basilar (por ej., las distintas regiones de resonancia producidas por los primeros siete u ocho armónicos de un tono musical). Pero a medida que nos desplazamos hacia el área cortical, esta representación fotográfica, punto a punto, se va perdiendo gradualmente, y la representación neural se torna holológica, proyectando la actividad de un solo punto de la membrana basilar sobre un conjunto grande y difuso de neuronas.

Una consecuencia natural del modo holológico de almacenamiento es el proceso de *recuerdo asociativo* (Kohonen, 1989). Efectivamente, el 'replay' de una actividad nerviosa específica puede ser desencadenado por causas o claves que no corresponden a la *totalidad* de las características sensoriales del evento original; una reactivación *parcial* de la actividad neural que se dio durante el acto original de almacenamiento, puede ser suficiente para disparar la totalidad de la actividad neural específica (por ej., el fenómeno de la fundamental ausente; todas las ilusiones ópticas). Sistemas comunes de almacenamiento de información, tales como libros, cintas magnéticas, registros fonográficos, y películas, operan en el modo fotográfico de almacenamiento de memoria; para extraer del sistema un trozo específico de información debemos conocer su dirección (ubicación o posición en el registro), o bien revisar todo el registro hasta encontrar lo que buscamos (esta es la razón por la cual este proceso se llama registro de memoria *direcciona*l). Pero en un sistema holológico de memoria de música grabada (¡que no existe todavía en la práctica!) ¡se podría cantar

o tocar el famoso *ta-ta-ta-taah* al sistema y obtener la quinta sinfonía de Beethoven completa (*memoria distribuida*, «*content-addressable*»)! ¡Y éste es el modo, precisamente, en que opera el cerebro!

¿Cómo funcionan en el cerebro el almacenamiento y la recuperación («recall») holológicos de información? Durante una fase determinada de aprendizaje, las áreas de asociación del cerebro y los sistemas de evaluación subjetiva en los lóbulos frontales muestran patrones de actividad neural disparados por una serie de constelaciones de estímulos que llegan simultáneamente a través de varios canales de entrada diferentes (por ejemplo, distintas modalidades sensoriales). Una vez completado el aprendizaje, se observa que en los circuitos nerviosos han tenido lugar apropiados cambios sinápticos (esto se llama *potenciación de largo plazo*), de tal manera que la constelación-estímulo de un solo canal de entrada (estímulo condicionado o «palabra clave») será suficiente para disparar el patrón de actividad nerviosa completo que fuera específico del evento original. En otras palabras, el aprendizaje no está representado por el almacenamiento 'per se' de información o de imágenes, sino por apropiadas modificaciones de la red neural de procesamiento de información²⁵; la memoria aparece en la forma de almacenamiento de *instrucciones* para el procesamiento de información. Esta es la razón por la cual se dice que el cerebro es un sistema «auto-organizado». Los devotos de la computación reconocerán que en la operación cerebral no hay ningún 'software': memoria, instrucciones y operaciones están todas basadas en cambios adecuados en el 'hardware' (la arquitectura y eficiencia de las conexiones sinápticas entre neuronas). La auto-organización es también el principio por el cual operan las «computadoras neuronales» de hoy (aunque los «cambios de hardware» son todavía simulados con programas apropiados, es decir, con 'software' apropiado - véase Hinton, 1992). Más información sobre el particular, así como un simple modelo de procesador central de altura que opera en forma holológica, podrá encontrarse en el Apéndice II.

²⁵ Una demostración cruda, pero muy convincente de esto, está dada por el hecho de que la densidad sináptica observada en las áreas corticales de un animal depende mucho de la complejidad del ambiente en el que el animal ha vivido, y de la «resolución de problemas» requeridos durante su vida.

5 Superposición y sucesiones de tonos compuestos y la percepción musical

«... se ha convertido en tabú para los teóricos de la música preguntar por qué nos gusta aquello que nos gusta... si diferentes personas tienen preferencias diferentes, no debemos simplemente ignorar el problema, sino intentar explicar cómo y por qué eso ocurre.»

M. Minsky, en *Music, Mind and Brain*,
M. Clynes, ed., Plenum Press, 1982.

En el curso de los capítulos 2, 3 y 4 hemos ido ascendiendo gradualmente la escalera del procesamiento neural de señales acústicas, partiendo de los mecanismos que llevan a la percepción de altura primaria, sonoridad, altura subjetiva y timbre, hasta llegar a la identificación de un instrumento musical. Del lado de la física, hemos analizado cómo las características del sonido que llevan a estas sensaciones son generadas en los instrumentos musicales. Estos atributos psicológicos son necesarios, pero de ninguna manera ingredientes suficientes de la música. La música está formada por *sucesiones* y *superposiciones* de tonos que transmiten ciertos *significados* que pueden ser analizados, almacenados y comparados en el cerebro. Sin embargo, es muy difícil realizar un estudio neuropsicológico cuantitativo de estos procesos cognitivos y afectivos de más alto nivel neural.

5.1 Superposición de tonos compuestos

La música polifónica consiste de superposiciones de tonos compuestos. Incluso en el caso de una sola melodía, como en la música monódica, la reverberación de los sonidos que llegan a nuestros oídos conducirá generalmente a la superposición de tonos compuestos. El estudio psicofísico de los efectos derivados de superposiciones de tonos compuestos dista mucho todavía de ser completo. Esto es verdad sobre todo en lo pertinente a la comprensión de cómo el cerebro es capaz de desenmarañar el «embrollo» de frecuencias de diferentes tonos

compuestos que suenan simultáneamente, de manera tal que las sensaciones de estos tonos permanezcan separadas.

Cuando se superponen dos tonos compuestos de diferente altura, pueden darse dos situaciones: o bien la frecuencia fundamental del tono más alto es igual a uno de los armónicos superiores del tono más bajo, o bien no lo es. En el primer caso, el tono más alto reforzará algunos armónicos superiores del tono más bajo. ¿Por qué, entonces, no percibimos simplemente un cambio de timbre en el tono más bajo, en lugar de percibir separadamente el tono más alto del más bajo, y por qué incluso podemos discriminar los timbres correspondientes a cada uno? Un problema similar surge en el segundo caso, donde cada tono produce su propia multiplicidad de zonas de resonancia sobre la membrana basilar. ¿Cómo distingue nuestro cerebro en la mezcla resultante cuál secuencia pertenece a qué tono? Por ejemplo, consideremos una superposición de tonos complejos, La_2 (110Hz) y $Do\#_3$ (140Hz). Los armónicos de ambos tonos están indicados en la Fig. 5.1 sobre una escala lineal de frecuencias. Cada una de estas frecuencias tiene una región de resonancia correspondiente sobre la membrana basilar (Fig. 2.8). A menos que haya pequeños cambios en la intensidad o en la altura (lo cual ocurre en la música real, véase más abajo), las células receptoras no tienen la menor clave para decidir a qué tono pertenece cada una de las zonas de resonancia. Por lo tanto, esta discriminación debe ser ejecutada en un centro más alto en el sistema nervioso auditivo.

La altura de los dos tonos es discriminada por el procesador central de altura (Sec. 2.9 y Apéndice II). Una capacidad de lo más notable del sistema nervioso auditivo es poder distinguir el *timbre* de dos tonos compuestos que suenan simultáneamente. Ninguna música sería posible sin esa capacidad. Por ejemplo, supongamos que escuchamos monoauralmente, con auriculares, el instrumento 1 tocando *exactamente* el La_4 , y al instrumento 2 tocando *exactamente* el La_5 , a prácticamente el mismo nivel de intensidad. La Fig. 5.2 muestra la superposición hipotética. La longitud total de las barras verticales representa la intensidad total de cada armónico que realmente llega a nuestro oído. ¿Cómo hace nuestro cerebro para mantener estos tonos separados en términos de timbre? El funcionamiento de este mecanismo de discriminación todavía no se entiende bien; un *elemento temporal* parece tener un rol clave. En primer

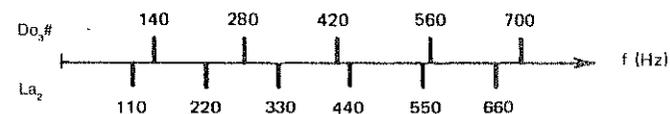


Figura 5.1 Frecuencias armónicas de dos tonos compuestos formando un intervalo de tercera mayor.

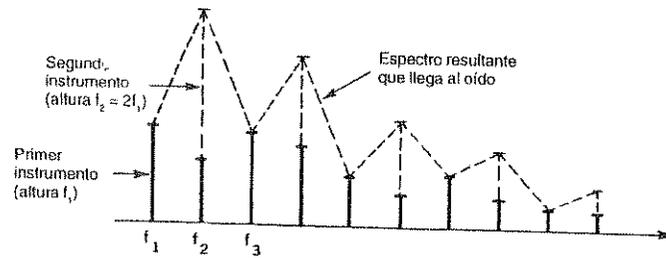


Figura 5.2 Espectro resultante de dos tonos compuestos de timbres diferentes (espectros diferentes) separados por una octava.

lugar, el comienzo (ataque) de dos tonos «simultáneos» nunca está perfectamente sincronizado, ni el modo en que se configura cada tono se desarrolla de la misma manera, especialmente cuando los tonos proceden de direcciones diferentes (efecto estéreo). Como se dijo en la pág. 169, el comienzo de un tono es el atributo más importante para la percepción del timbre y la identificación del tono (Iverson y Krumhansl, 1993). Durante este período transiente, el mecanismo de procesamiento de nuestro cerebro parece ser capaz de centrarse sobre ciertos rasgos característicos del patrón vibratorio de cada instrumento, y seguir el paso de estos rasgos, aun cuando ellos estén distorsionados y borroneados por la señal de otro instrumento. Durante la ejecución, el mecanismo de procesamiento auditivo también usa pequeños desvíos en la altura y la intensidad (el así llamado efecto *chorus*) que son coherentes para la serie completa de los armónicos de cada uno de los instrumentos (Secs. 4.9 y 4.10). Otras variaciones, más pronunciadas, que parecen dar importantes claves al mecanismo de discriminación del timbre, son las variaciones periódicas de altura (vibrato) y de intensidad (trémolo, vibrato) que se producen (en forma voluntaria o involuntaria) en tonos de muchos instrumentos musicales. Superposiciones de tonos compuestos desprovistas totalmente de estas leves perturbaciones coherentes dependientes del tiempo - como ocurre cuando juegos múltiples son combinados en la música de órgano- son en realidad mucho más difíciles de discriminar tímbricamente.

Lo que probablemente más ayude al mecanismo de discriminación, tanto de la altura como del timbre de tonos complejos, es la información recibida de una *progresión* de superposiciones de tonos. En ese caso no sólo pueden ser usadas las claves primarias arriba mencionadas, dadas por las fluctuaciones coherentes en 'timing' y frecuencia de cada tono, sino que será también aprovechable la información «de orden más alto» o información secundaria extraída de las líneas melódicas ejecutadas por cada instrumento («mensajes» musicales) (véase asimismo pág. 166).

El mecanismo de discriminación de tonos complejos tiene su equivalente (y probablemente su raíz primordial) en el mecanismo que

gobierna nuestra percepción auditiva cuando seguimos el discurso de una persona determinada entre muchas conversaciones diferentes simultáneas a niveles sonoros similares. Esta habilidad ha sido llamada el «efecto cóctel party», y muy probablemente use las mismas claves, primarias y secundarias, que el mecanismo de discriminación de tonos compuestos. Por último, es probablemente este mismo mecanismo el que nos permite desenredar los sonidos individuales de una confusa superposición de tonos en una sala muy reverberante (Sec. 4.7). En este último caso, es otra vez un efecto temporal el que parece tener el rol principal: la «primera» llegada del sonido directo (figs. 4.26 y 4.27) suministra claves que nuestro sistema perceptivo usa para determinar la real sensación del tono y su discriminación (efecto de precedencia).

Tal como ocurre con dos tonos simples, para discriminar dos tonos compuestos debe existir una diferencia mínima entre sus frecuencias fundamentales. Cuando dos tonos complejos difieren en su altura menos que el límite de discriminación, pueden surgir batidos de primer orden (sec. 2.4) entre todos los armónicos. Si por ejemplo ambos tonos complejos forman un unísono desafinado, con frecuencias fundamentales f_1 y $f_1 + \epsilon$, respectivamente, *todas* las zonas de resonancia se superpondrán sobre la membrana basilar y producirán batidos de diferentes frecuencias. Las fundamentales batirán con frecuencia ϵ , los segundos armónicos con frecuencia 2ϵ , y así sucesivamente. Solo los primeros pocos armónicos son importantes; habitualmente los batidos de la fundamental (ϵ) son los más pronunciados. Si superponemos dos tonos complejos formando otros intervalos musicales desafinados, también surgirán batidos de primer orden entre los armónicos que se correspondan. Estos batidos son diferentes (aunque iguales en frecuencia) de los batidos de segundo orden que aparecen en las superposiciones de tonos simples cuando forman intervalos desafinados (Sec. 2.6).

5.2 Sensación musical de consonancia y disonancia

Consonancia y disonancia son sensaciones subjetivas asociadas con dos o más tonos que suenan simultáneamente, de una naturaleza mucho menos definida que las variables psicofísicas de altura y sonoridad, incluso la del timbre. Sin embargo, la música basada en sonidos de altura definida (tonos) de *todas* las culturas parece indicar que el sistema auditivo de los seres humanos tiene preferencia para determinados intervalos de frecuencia: la octava, la quinta, la cuarta, etc. Lo más significativo es que estos intervalos son «valuados» en el mismo orden en que aparecen en la serie de los armónicos (véase fig. 2.19).

Cuando dos tonos compuestos suenan al unísono o a la octava perfectamente afinados, *todos* los armónicos del segundo tono estarán

exactamente apareados con armónicos del primer tono; ninguna frecuencia «intermedia» será introducida por el segundo tono. Esta propiedad pone a la octava en una situación muy especial como intervalo musical (además de los argumentos concernientes a la simplicidad del patrón vibratorio, expuestos en el análisis de superposiciones de tonos simples en la Sec. 2.6). La situación cambia cuando dos tonos compuestos suenan a la quinta justa (tabla 5.1). Todos los armónicos impares de la quinta tienen frecuencias que caen *entre* los armónicos de la tónica; sólo coinciden sus armónicos pares. En particular, el tercer armónico de la quinta, de frecuencia $\frac{9}{2} f_1$, cae «peligrosamente cerca» de las frecuencias que corresponden al cuarto y quinto armónico de la tónica: sus regiones de resonancia pueden superponerse y originar de este modo batidos o rugosidad (Sec. 2.4), incluso si el intervalo dado por las frecuencias fundamentales está perfectamente afinado. Construyendo tablas similares a la 5.1, el lector podrá verificar que para otros intervalos musicales tales como la cuarta, las terceras y las sextas la proporción de armónicos que «chocan» se incrementa rápidamente y se desplaza hacia los bajos órdenes de los armónicos. Históricamente, este efecto fue considerado como la principal causa de la sensación de creciente disonancia en la serie de intervalos musicales.

Efectivamente, desde los tiempos de von Helmholtz, la disonancia se solía asociar con el número, la intensidad y la frecuencia de los armónicos que baten, y la consonancia con la ausencia de armónicos batientes. En otras palabras, se asumía que por alguna razón nuestro sistema auditivo «no gusta de los batidos». Como resultado, se prefiere, por encima de todo, el unísono y la octava justa, porque en estos intervalos todos los armónicos del tono más alto coinciden exactamente con los armónicos de la tónica. En la quinta, según la tabla 5.1, el tercer armónico del tono superior puede batir con el cuarto y el quinto armónico de la tónica. La proporción cada vez mayor de pares de armónicos que baten, a medida

Tabla 5.1 Comparación de los primeros armónicos de dos tonos compuestos separados con una quinta. Líneas continuas: posibles pares «batientes».

Tónica	Quinta justa
f_1	
	$f'_1 = \frac{3}{2} f_1$
$2f_1$	
$3f_1$	$f'_2 = 3f_1$
$4f_1$	
$5f_1$	$f'_3 = \frac{9}{2} f_1$
$6f_1$	$f'_4 = 6f_1$

que uno va procediendo hacia la cuarta, las sextas y las terceras, la séptima, la segunda, etc., explicaría de este modo la disminución de consonancia - o el incremento de disonancia - en estos intervalos. Esta suposición era particularmente atractiva porque - como es fácil de demostrar matemáticamente - para maximizar el número de armónicos coincidentes de dos tonos complejos (y de aquí minimizar aquel de los no coincidentes) es necesario que sus frecuencias fundamentales f_1 y f_1' estén en razón de números enteros, y que estos números sean tan pequeños como sea posible. Efectivamente, si

$$\frac{f_1'}{f_1} = \frac{n}{m} \quad n, m: \text{ enteros} \quad (5.1)$$

el armónico de orden m de f_1' tendrá la misma frecuencia que el armónico de orden n de f_1 : $mf_1' = nf_1$ (y del mismo modo, el de orden $2 \times m$ tendrá la misma frecuencia que el de orden $2 \times n$, etc.). Todos los otros armónicos no coincidirán, y eventualmente originarán batidos si sus frecuencias están lo suficientemente cerca. La tabla 5.2 muestra los intervalos dentro de la octava que pueden formarse con números m y n pequeños y que son aceptados en nuestra cultura musical occidental como consonancias (en un orden de «perfección» decreciente).

En base a sofisticados experimentos monoaurales y dicóticos (Plomp y Levelt, 1965) en los que los sujetos tenían que juzgar la consonancia de pares de tonos simples y tonos compuestos inarmónicos, se descubrió que los batidos entre armónicos son el factor determinante en la percepción de la consonancia. Se presentaron superposiciones de dos tonos *simples* separados por intervalos menores que una octava a sujetos no entrenados musicalmente, pidiéndoles que calificaran la superposición como «consonante» o «placentera». Se obtuvo un resultado *continuo* que no revelaba preferencia por ningún intervalo en particular. Un ejemplo es la figura 5.3. Siempre que los tonos simples estuvieran separados por menos que una tercera menor, el juicio fue «disonante» (excepto para el unísono); los intervalos iguales o mayores que la tercera menor fueron juzgados

Tabla 5.2 Razones de frecuencias de los intervalos musicales.

Razón de frecuencias (n/m)	Intervalo
Consonancias «perfectas»	<ul style="list-style-type: none"> $1/1$ Unísono $2/1$ Octava $3/2$ Quinta $4/3$ Cuarta
Consonancias «imperfectas»	<ul style="list-style-type: none"> $5/3$ Sexta mayor $5/4$ Tercera mayor $6/5$ Tercera menor $8/5$ Sexta menor

94

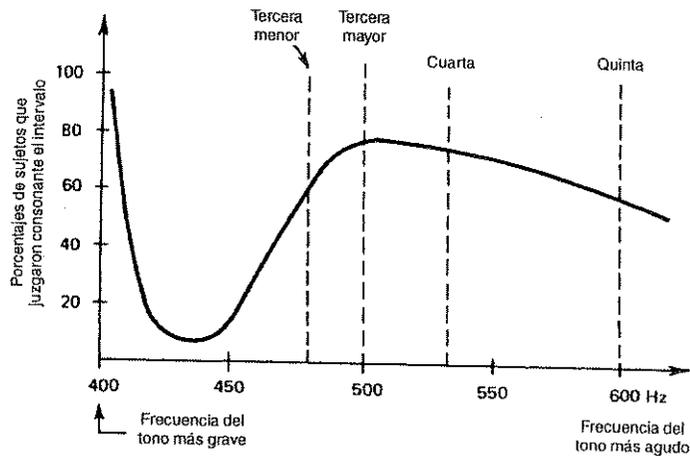


Figura 5.3 «Índice» de consonancia de superposiciones de dos tonos puros (según Plomp y Levelt, 1965).

más o menos como consonancias, independientemente de la actual relación de sus frecuencias¹. En realidad, la forma de la curva depende de la frecuencia absoluta del tono fijo. Todo esto se debe a la sensación de rugosidad de los unísonos desafinados y a la banda crítica discutida en la sección 2.4. Podemos sintetizar los resultados de estos experimentos de la manera siguiente: 1) Cuando la frecuencia de los dos tonos simples cae fuera de la banda crítica, el intervalo correspondiente de tonos simples será considerado consonante. 2) Cuando las frecuencias coinciden, el juicio será «perfectamente» consonante. 3) Cuando la frecuencia de los dos tonos simples difieren en una cantidad comprendida en el rango que va desde alrededor de un cinco a un cincuenta por ciento del ancho de la banda crítica correspondiente, el juicio será «no consonante». Llamaremos «disonancia básica» a un intervalo de dos tonos simples que cae dentro de esta última condición.

Volvamos ahora al caso musicalmente más significativo de dos tonos compuestos que suenan simultáneamente, y apliquemos individualmente los resultados de arriba a cada par de armónicos vecinos superiores. Si sopesamos el número de pares que son más o menos consonantes (véase más arriba caso 1) y perfectamente consonantes (caso 2) con el número de disonancias básicas (caso 3), puede obtenerse un «índice de consonancia» para cada intervalo de tonos compuestos (Plomp y Levelt, 1965; Kameoka y Kuriyagawa, 1969). Puede demostrarse que este índice

¹ Los músicos entrenados fueron excluidos de este experimento, porque ellos hubieran tenido una fuerte tendencia a identificar las consonancias sobre la base de su experiencia musical.

alcanza valores máximos para tonos cuyas frecuencias fundamentales satisfagan la condición (5.1): la altura de los picos (grado de consonancia) sigue, aproximadamente, el orden decreciente dado en la tabla 5.2. Lo que es más, en vista de la dependencia entre el ancho de la banda crítica y la frecuencia (Fig. 2.13), un intervalo musical determinado tiene un grado de consonancia que varía a lo largo del rango de frecuencias. Específicamente, al desplazarse hacia las frecuencias más bajas, un intervalo musical determinado va siendo cada vez menos consonante; un hecho bien conocido en la música polifónica, donde en el registro grave se usan principalmente octavas y, eventualmente, quintas, pero no terceras.

El grado de consonancia también depende del timbre o espectro de los tonos componentes, es decir, de la intensidad relativa de los pares de armónicos superiores perturbadores. Esto, también, es bien conocido en música: hay combinaciones de instrumentos que se «mezclan» mejor que otras. Incluso es importante el orden en que dos instrumentos definen un intervalo. Si por ejemplo un clarinete y un violín suenan a un intervalo de tercera mayor, con el clarinete tocando la nota inferior, el primer par disonante de armónicos será el séptimo armónico del clarinete con el sexto armónico del violín (porque los modos pares más bajos del clarinete están muy atenuados, secciones 4.4 y 4.5). Este intervalo suena «suave». Si, en cambio, tenemos al clarinete tocando el tono superior, el tercer armónico de este último chocará con el cuarto armónico del tono de violín, y el intervalo sonará «áspero».

En su teoría sobre la percepción de consonancia, Terhardt (1974) postula que la música basada en sonidos de altura determinada (tonos) está basada en los mecanismos de reconocimiento de patrones que operan en el sistema auditivo (Seccs. 2.9 y 4.8). Uno de estos - el procesador central de altura -, responsable de la extracción de una sensación de altura única a partir de la complicada distribución de la actividad neural producida por un tono musical, toma conocimiento de las relaciones específicas que existen entre las zonas de máxima resonancia y los focos de actividad neural resultantes producidos por los primeros seis a ocho armónicos de dicho tono (Sec. 4.8). Los intervalos correspondientes de altura primaria (octava, quinta, cuarta, tercera mayor, tercera menor) van haciéndose «familiares» al procesador central del sistema auditivo y transmiten significados tonales a todos los estímulos externos cuyas frecuencias (fundamentales) guardan tales relaciones (Apéndice II)². Según esta teoría, tanto la rugosidad

² Si esta hipótesis de que el procesador central de altura es una unidad neural que debe aprender a extraer información significativa de los tonos compuestos (Terhardt, 1972; Sec. 2.9) fuese correcta, podría tener consecuencias interesantes. Por ejemplo, uno podría reaprender un nuevo conjunto de características «invariantes» pertenecientes, digamos, a una clase determinada de tonos inarmónicos, con la posibilidad de construir escalas y esquemas tonales enteramente nuevos (Terhardt, 1974). Desde un punto de vista más práctico, esta capacidad intrínseca de aprendizaje acrecentaría las expectativas para el

mínima como el significado tonal tienen un rol determinante en la sensación de consonancia subjetiva. Sin embargo, en vista del fenómeno de cambio de altura primaria de los armónicos componentes percibidos individualmente (pág. 167), estos dos principios pueden impartir instrucciones contrarias y, en situaciones musicales reales, forzar al procesador central a un compromiso (Terhardt, 1974). La afinación «alargada» que se prefiere para los pianos (comparada con la escala igualmente temperada, Sec. 5.3), y el hecho observado de que la nota más alta de un intervalo melódico de tonos sucesivos se entona preferencialmente aguda (Sec. 5.4), pueden ser ambos un resultado de este compromiso.

Existen factores más complicados que influyen en la sensación de consonancia, notablemente la experiencia y el entrenamiento, así como el prejuicio resultante (es decir, la tradición musical). Es interesante observar que, históricamente, los intervalos musicales, como ingredientes armónicos explícitos, han sido «aceptados» gradualmente en nuestra civilización occidental, y esta aceptación siguió un orden que se parece mucho al dado en la tabla 5.2. Esto parece sugerir una gradual tolerancia de nuestra habilidad de procesamiento auditivo. Desde luego, esto no fue el resultado de una evolución biológica, sino más bien el de una sofisticación de las experiencias de *aprendizaje* a las cuales los seres humanos fueron expuestos a lo largo del tiempo. Este desarrollo, así como el de la civilización en un todo, procedió escalonadamente, por saltos; fue siempre la mente de un genio revolucionario que tuvo la osadía de introducir innovaciones, que requerían operaciones de procesamiento cada vez más complejas por parte del cerebro, y fue el carisma del genio lo que se necesitó para persuadir a la gente a aprender, y de este modo a aceptar y *preservar*, estas atrevidas innovaciones.

Hasta ahora hemos considerado intervalos iguales o más pequeños que la octava. Para intervalos grandes (por ej. $Do_3 - Sol_5$), es habitual transportar el tono más alto a las octavas inferiores ($Sol_5 - Sol_4 - Sol_3$) hasta obtener un intervalo más pequeño que la octava ($Do_3 - Sol_3$). El grado de consonancia del último intervalo es considerado entonces «equivalente» a aquel del intervalo original. Esta propiedad cíclica de los intervalos que se repite en las sucesivas octavas ha sido llamada el *croma* de los tonos musicales. Es una propiedad básica que asigna un rango similar a todos los tonos cuya altura difiere en una o más octavas, y que hace que llamemos a todos esos sonidos con el mismo nombre. ¿Qué cosa es responsable de este curioso carácter cíclico de los tonos musicales, que se repite cada octava (cada vez que la frecuencia se duplica)? No

desarrollo de prótesis electrónicas para sordos, basadas en el implante de *microelectrodos* en el *nervio acústico*. Mientras que el patrón espacial de actividad de estos implantes es extremadamente difícil de predecir, la interpretación de los patrones de excitación producidos podría ser aprendida por el procesador central del paciente.

existe ningún equivalente de esta propiedad en otras modalidades sensoriales. Obviamente, está relacionado con esa propiedad clave de la todopoderosa octava: de tener todos sus armónicos coincidentes con armónicos superiores (de orden par) del tono más bajo. No hay ningún otro intervalo musical que posea esta propiedad (exceptuando al unísono, por supuesto). En términos generales, la existencia del croma, es decir, del hecho que las alturas que difieren en una octava tengan un grado de similitud idéntica a aquella del unísono, indica que el proceso de reconocimiento de patrones en nuestro sistema auditivo responde de alguna manera «especial», tal vez simplificada, al presentársele las octavas. Obsérvese de nuevo que la octava es el primer intervalo en la serie de los armónicos, y que tiene una frecuencia fundamental de repetición *idéntica* a la del tono más bajo. Cualquier otro intervalo musical consonante (quinta, cuarta, etc.) tiene una frecuencia fundamental de repetición asociada (relaciones [2.7]) que *no* está presente en ninguno de los dos tonos del estímulo original. Si recordamos cómo podría funcionar el procesador de altura (Sec. 2.9 y Apéndice II), nos damos cuenta que, siempre que se presenten *dos* tonos complejos cuyas frecuencias fundamentales f_1 y f_2 estén formando un intervalo musical, la salida [output] del procesador de altura debería contener dos señales prominentes, (correspondientes a f_1 y f_2), más otras señales *menos* prominentes que representan la frecuencia de repetición (2.7) correspondiente al par de primeros armónicos de f_1 y f_2 y sus múltiplos (Apéndice II, Fig. AII.2). Bajo condiciones normales, estas señales adicionales son *descartadas* como sensaciones de altura, proceso que requiere una operación adicional de *filtrado*. Nótese, sin embargo, que esta operación adicional no es necesaria cuando el intervalo presentado es una *octava*, ¡porque no se presenta ninguna tercera señal! (Fig. AII.2). De hecho, el «significado tonal» arriba mencionado puede estar estrechamente relacionado con el número, la intensidad y posición de señales «parasitarias» a la salida del procesador de altura. Cuanto más compleja sea la multiplicidad de estas señales (es decir, cuanto más complejo sea el patrón de vibración sonora), tanto «más bajo» será el significado tonal de la superposición original de tonos.

Finalmente, cuando *tres o más* tonos suenan juntos, es habitual analizar el acorde resultante en pares de tonos, considerando individualmente sus valores de consonancia. Es obvio que, cuanto más tonos compuestos sean combinados, tanto más complicada será la configuración de zonas de resonancia originada sobre la membrana basilar. Habrá cada vez mayor superposición, al punto de llegar a una situación en que todas las células receptoras sobre una amplia región de la membrana basilar responden al mismo tiempo. A la luz de las varias teorías sobre la altura (Secs. 2.9, 4.8, y Apéndice II), podemos esperar que también en este caso, el grado de consonancia (o de disonancia) puede estar relacionado tanto con la proporción de armónicos que batan como con el número, intensidad y

posición de señales parasitarias a la salida del procesador de alturas. Nótese que la tríada mayor es una combinación de tres tonos cuyos componentes, tomados de dos a la vez, producen frecuencias de repetición que sólo difieren de la tónica por octavas (es decir que tienen el mismo cromá).

Existe sin embargo un límite para la inteligibilidad de superposiciones de tonos. Cuando los patrones vibratorios son aleatorios (o sea, cuando se destruye su periodicidad), o cuando su complejidad excede cierto umbral, el mecanismo de procesamiento nervioso simplemente se rinde: no puede establecer sensaciones definidas de altura o timbre. La sensación resultante se llama *ruido*. Cualquier oscilación no periódica de la presión conduce a la sensación de ruido. Sin embargo, el ruido puede ser altamente organizado. Del mismo modo que una oscilación periódica puede ser analizada en una superposición discreta de oscilaciones armónicas puras de frecuencias que son múltiplos enteros de la frecuencia fundamental (Sec. 4.3), las vibraciones aperiódicas pueden ser analizadas como una superposición *continua* de vibraciones puras a *todas* las frecuencias posibles. Según sea la manera en que la intensidad es distribuida en las frecuencias posibles obtendremos diferentes *espectros de ruido*. El ruido tiene un rol clave en la formación de consonantes en el habla. El ruido también desempeña un rol importante en música; la importancia de los componentes de ruido de los instrumentos de percusión es obvia. Está también demostrado que el ruido detectado en las primeras décimas de segundo en un sonido de piano o de clavecín es un elemento básico en el proceso de reconocimiento. El efecto de los ruidos de los espectros controlados electrónicamente sobre nuestro sistema auditivo está siendo estudiado ampliamente por medio de sintetizadores. Un vasto y nuevo territorio de sensaciones auditivas (¿música?) está siendo revelado (véase también Sec. 5.7).

5.3 Construyendo escalas musicales

Con fin puramente práctico, definamos una escala como *un conjunto discreto de alturas dispuestas de manera tal que se obtenga un máximo número posible de combinaciones consonantes* (o un mínimo número de disonancias) *cuando dos o más notas del conjunto son tocadas juntas*. Con esta definición, y teniendo en mente la tabla 5.2, es posible generar a la vez dos escalas, de una manera casi inequívoca, dependiendo de si se tienen en cuenta todos los intervalos consonantes, o si solo se considerarán las consonancias perfectas. En el primer caso obtenemos la *escala justa*; en el segundo, la *escala pitagórica*³.

³ Aquí concebimos una escala (también llamada temperamento) como un conjunto de tonos con *relaciones de frecuencia definidas matemáticamente*. Esto debe ser distinguido de los diversos *modos* de una escala, definidos por el orden particular en que los tonos y los semitonos se suceden entre sí.

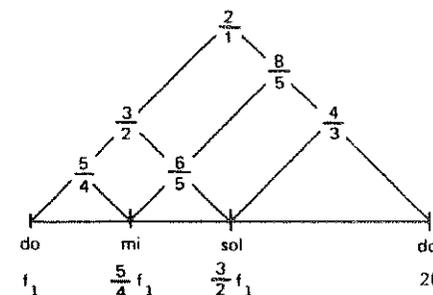


Fig. 5.4 Primer conjunto de intervalos consonantes obtenidos en el proceso de generación de una escala justa (ver texto).

1. La escala justa

Empezamos con un tono de frecuencia f_1 que llamamos *do*⁴. El primer paso obvio es introducir la octava superior, que indicaremos *do'*. Esto produce el intervalo más consonante de todos. El siguiente paso obvio es agregar la quinta, de frecuencias $\frac{3}{2}f_1$, que llamamos *sol*. Esto da dos nuevos intervalos consonantes, además de la octava, de razones de frecuencias $\frac{3}{2}$ (*do - sol*) y $\frac{4}{3}$ (la cuarta *sol - do'*), respectivamente. Para el próximo paso hay dos posibilidades, si queremos obtener un máximo número de intervalos consonantes. Estas son las notas $\frac{5}{4}f_1$ o $\frac{6}{5}f_1$, que llamamos *mi* y *mib*, respectivamente. Nosotros elegiremos la primera posibilidad, el *mi*, porque esto garantizará un número de consonancias de grado más alto. La figura 5.4 muestra los intervalos que resultan, todos ellos consonantes. Las notas *do - mi - sol* constituyen la *tríada mayor*, piedra fundamental de la armonía occidental (la segunda posibilidad $\frac{6}{5}f_1$ o *mib*, hubiera originado la *tríada menor*).

Podemos seguir «introduciendo» notas, intentando en cada paso mantener el número de disonancias a un mínimo y el número de consonancias (tabla 5.2) a un máximo. Esto nos lleva a la *escala diatónica justa* de siete notas en la octava (fig. 5.5). Estas siete notas pueden proyectarse a las octavas superiores e inferiores para formar la escala diatónica completa, a lo largo de todo el rango de alturas audibles.

⁴ La notación del solfeo *do - re - mi - fa - sol - la - si - do* es usada aquí para determinar la posición *relativa* en una escala (es decir, el cromá), no la altura real.

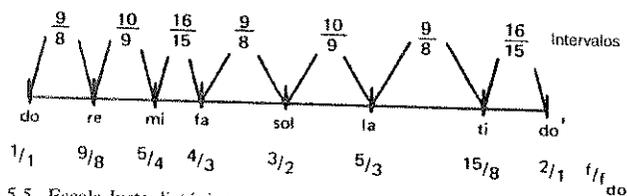


Figura 5.5 Escala Justa diatónica.

Obsérvese en la fig. 5.5 los dos intervalos con razones cercanas de frecuencias $\frac{9}{8}$ y $\frac{10}{9}$, representando *tonos enteros*. El intervalo $\frac{16}{15}$ define el *semitono*. Con las notas de esta escala, tomadas en pares, podemos formar 16 intervalos consonantes, 10 intervalos disonantes (séptimas menores y mayores, quinta disminuída, tonos, semitonos) y - desafortunadamente - dos consonancias «desafinadas»: la 1,5% demasiado grande tercera menor *re - fa* ($\frac{32}{27}$) y la 1,9% demasiado chica quinta *re - la* ($\frac{40}{27}$). Finalmente, y tal vez lo más importante, cabe subrayar que con la escala diatónica justa podemos formar tres tríadas mayores justas: do- mi- sol, do- fa- la, y re- sol- si; dos tríadas menores justas: mi- sol- si y do- mi- la, y una tríada menor desafinada: re- fa- la.

Considerando la existencia de distancias desiguales entre notas vecinas, es posible continuar implementando esta escala, partiendo las distancias más grandes (tonos enteros) en dos semitonos cada una. Desafortunadamente, los nuevos intervalos que resultan van siendo cada vez más complicados (por ej., varias clases de semitonos, más consonancias desafinadas); las elecciones no son únicas, y se obtienen diferentes valores de frecuencia para las así llamadas equivalencias enarmónicas do# - reb, re# - mib, etc. Tratando de mantener al máximo la proporción de consonancias posibles, se introducen las siguientes notas: mib ($\frac{6}{5}f_1$), sib ($\frac{9}{5}f_1$), sol# ($\frac{25}{16}f_1$) (o lab, $\frac{8}{5}f_1$), do# ($\frac{25}{24}f_1$), y fa# ($\frac{45}{32}f_1$). El resultado es una *escala cromática justa*, de 12 notas en la octava.

2. La escala pitagórica

Ahora nos vamos a restringir a las así llamadas consonancias perfectas, la quinta justa y la cuarta justa (y la octava, por supuesto), construyendo una escala sobre la base de estos intervalos exclusivamente. Podemos proceder de la siguiente manera. Después de introducir el *sol* nos desplazamos una quinta justa hacia abajo desde el *do'* para introducir el *fa* ($\frac{2}{3} \times 2f_1 = \frac{4}{3}f_1$). Luego nos desplazamos una cuarta justa descendente

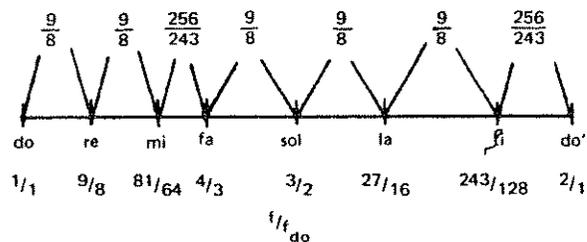


Figura 5.6 Escala diatónica pitagórica.

desde el *sol* para obtener el *re* ($\frac{3}{4} \times \frac{3}{2}f_1 = \frac{9}{8}f_1$), y una quinta hacia arriba desde esta última nota, para llegar al *la* ($\frac{3}{2} \times \frac{9}{8}f_1 = \frac{27}{16}f_1$). Finalmente llenamos los espacios que quedan saltando una cuarta hacia abajo desde el *la* para obtener *mi* ($\frac{3}{4} \times \frac{27}{16}f_1 = \frac{81}{64}f_1$), y una quinta hacia arriba desde este *mi* para llegar al *si* ($\frac{3}{2} \times \frac{81}{64}f_1 = \frac{243}{128}f_1$). El resultado es la llamada escala pitagórica (Fig. 5.6). Nótese que hay solo un tipo de intervalo entero: el tono pitagórico, de razón de frecuencias $\frac{9}{8}$ (igual al tono «chico» de la escala justa). El intervalo $\frac{265}{243}$ es el *semitono diatónico pitagórico*.

Podemos convertir esta escala en una cromática, continuando por saltos ascendentes o descendentes en cuartas y quintas justas. De este modo obtenemos *fa#* (una cuarta debajo del si); *do#* (una cuarta abajo del fa#); *sol#* (una quinta arriba del do#); *sib* (una cuarta arriba del fa), y *mib* (una cuarta arriba del sib). De esta manera aparece un nuevo tipo de semitono (por ej., fa - fa#) definido por la curiosa razón de frecuencias 2187/2048, llamada *semitono cromático pitagórico*. Este procedimiento, al continuarse, vuelve a conducir a equivalencias enarmónicas de diferentes frecuencias. Particularmente, si continuamos procediendo por saltos ascendentes y descendentes de cuartas y quintas volveremos a nuestro *do* inicial - ¡pero no exactamente! En otras palabras, vamos a llegar al equivalente enarmónico «si#», cuya frecuencia *no* es igual a la del *do'* (= $2f_1$).

De esta forma, basándonos en algunos principios «lógicos», hemos generado dos escalas, cada una con sus propios problemas. El más serio, de lejos, es común a ambas y consiste en el hecho de que con estas escalas sólo puede tocarse un grupo muy reducido de tonalidades sin caer en el problema de consonancias desafinadas. En otras palabras, *ambas escalas imponen serias restricciones a la transposición y a la modulación*. Esto ya era bien conocido en el siglo diecisiete. Con todo, no hay duda de que estas dos escalas imprimen un carácter muy específico cuando se toca música en instrumentos afinados sobre ellas; pero el tipo de música que se puede ejecutar es extremadamente limitado.

3. La escala de temperamento igual

Se hizo evidente la necesidad de una nueva escala, la cual, con un compromiso razonable, cediendo algo en lo que se refiere a la «justeza» de los intervalos, llevaría a *intervalos igualmente espaciados*, sin importar la tonalidad particular. En otras palabras, un semitono tendría la misma razón de frecuencias, se trate de do - do#, de mi - fa, o de la - sib, y una quinta sería la misma, ya se trate de fa - do' o de do# - sol#. Esto llevó a la *escala temperada*, que recibió el ferviente apoyo de nada menos que J.S. Bach, quien compuso una colección de preludios y fugas («Das Wohltemperierte Clavier») con el propósito específico de sacar el máximo provecho de las nuevas fronteras abiertas por la posibilidad de cambiar, sin restricciones, de una tonalidad a otra.

En la escala temperada la razón de frecuencias es la misma para los 12 semitonos entre do y do'. Llamemos s a esta razón. Esto significa que:

$$f_{do\#} = sf_{do}, \quad f_{re} = sf_{do\#} = s^2 f_{do}, \dots, f_{do'} = s^{12} f_{do}$$

Desde que sabemos que $f_{do'} = 2f_{do}$ (¡solo la octava es mantenida como intervalo justo!), la potencia duodécima de s debe ser igual a 2. O sea

$$s = \sqrt[12]{2} = 1,0595 \tag{5.2}$$

Esta es la razón de frecuencias del *semitono temperado*. Las frecuencias correspondientes a las notas de la escala cromática temperada son las potencias enteras de s multiplicadas por f_{do} . La tabla 5.3 muestra las razones de frecuencias de los intervalos consonantes en las tres escalas.

Es conveniente introducir una subdivisión standard del intervalo básico de la escala temperada con el fin de poder expresar numéricamente las pequeñas diferencias entre los intervalos pertenecientes a las diferentes escalas. Esta subdivisión se usa para describir pequeños cambios de frecuencia (vibrato), cambios en la entonación (altura), y desafinación de notas o intervalos. El procedimiento más aceptado hoy día es dividir el semitono temperado en *100 intervalos iguales*, o, lo que es lo mismo, dividir la octava en 1200 partes iguales. Dado que lo que define un intervalo musical es la *razón* de frecuencias fundamentales de sus tonos componentes (no su diferencia), lo que debemos hacer aquí es dividir la razón de frecuencias del semitono, es decir, dividir s (5.2), en *100 factores* iguales c :

$$\underbrace{c \times c \times c \times \dots \times c}_{100 \text{ veces}} = c^{100} = s$$

En vista de la relación (5.2), el valor de c resulta

$$c = \sqrt[100]{1,0595} = 1,000578 \tag{5.3}$$

La unidad de esta subdivisión se llama *cent*. Para determinar cuántos cents

«contiene» un intervalo dado de razón de frecuencias r , debemos determinar cuántas veces debemos multiplicar c por sí mismo para obtener r :

$$c^n = r \tag{5.4}$$

n es entonces el valor de r expresado en cents. Por definición, un semitono temperado son cien cents, un tono temperado (s^2) son 200 cents, una quinta temperada (s^7) son 700 cents, etc. Para conocer el valor en cents de cualquier otro intervalo debemos usar logaritmos. Teniendo en cuenta las propiedades descritas en la sección 3.4, tomando logaritmos a la relación (5.4) obtenemos $n \log c = \log r$. Por lo tanto

$$n = \frac{\log r}{\log c} = 3986 \log r \tag{5.5}$$

Usando esta relación, encontramos los valores en cents de los diversos intervalos consonantes dados en la tabla 5.3.

5.4 La escala standard y el standard de altura

La escala temperada ha sido usada por más de doscientos años y se convirtió de hecho en la escala patrón a la cual se afinan todos los instrumentos con notas de altura fija. Sin embargo, desde su aparición hasta el presente, esta escala ha sido el blanco de críticas. Lo que se cuestiona principalmente es la «no justeza» de los intervalos consonantes, especialmente las terceras y sextas (tabla 5.3), las cuales efectivamente suenan algo desafinadas al ser escuchadas con atención y persistentemente, en particular en el registro bajo.

Comparemos entre sí las escalas comentadas en la sección previa. Indudablemente, para una tonalidad determinada, la escala justa es «teóricamente» la escala perfecta, dando el máximo número posible de combinaciones de intervalos justos o puros (o sea, intervalos que no

Tabla 5.3 Razones de frecuencias y valores en cents de los intervalos musicales, para las tres escalas consideradas en el texto.

Intervalo	Escala justa		Escala pitagórica		Escala temperada	
	Razón	Cents	Razón	Cents	Razón	Cents
Octava	2,000	1200	2,000	1200	2,000	1200
Quinta	1,500	702	1,500	702	1,498	700
Cuarta	1,333	498	1,333	498	1,335	500
Tercera mayor	1,250	386	1,265	408	1,260	400
Tercera menor	1,200	316	1,184	294	1,189	300
Sexta mayor	1,667	884	1,687	906	1,682	900
Sexta menor	1,600	814	1,580	792	1,587	800

generan batidos). Por esta razón, la escala justa debería ser tomada como una especie de escala de referencia; es precisamente por eso que la hemos introducido en primer lugar. Pero cabe preguntarse lo siguiente ¿Le importa realmente a nuestro sistema auditivo que los intervalos no batan? Además, ¿sacrificaríamos las posibilidades de transposición y de modulación a cambio de obtener estos intervalos puros? Doscientos años de historia musical han contestado, sin titubeo, con un fuerte y claro ¡no! Así que podemos dejar de lado a la escala justa.

La escala pitagórica puede ser un paso hacia adelante (mientras las quintas y cuartas se mantienen como intervalos justos, las terceras y sextas están desafinadas levemente, tabla 5.3), pero todavía sigue presentando limitaciones a la posibilidad de transporte y modulación. Ha habido otras escalas, introducidas como modificaciones leves de la escala pitagórica, que nosotros ni siquiera vamos a mencionar. Ninguna de ellas ha logrado desplazar a la escala temperada, que sigue siendo aceptada universalmente como el mejor compromiso.

Ha habido intentos de dilucidar experimentalmente la cuestión de cuál es realmente la escala preferida por el sistema auditivo (dejando a un lado la cuestión de la modulación). Hay dos modos posibles de encarar el problema: 1) Usar instrumentos de frecuencia fija (piano, órgano), y comparar cuidadosamente las impresiones subjetivas al tocar una determinada pieza de música sucesivamente en instrumentos de la misma clase, afinados, respectivamente, a distintas escalas. Desde luego, la pieza debería ser muy simple, sin modulaciones a tonalidades lejanas. El instrumento, por otra parte, debería ser del tipo como el órgano, cuyos sonidos no presentan caída, con el fin de que los batidos o la rugosidad se presenten con mayor claridad. 2) Medir experimentalmente el promedio de frecuencias de las alturas elegidas por los cantantes o por ejecutantes de instrumentos de frecuencia variable (cuerdas), y determinar si prefieren uno u otro temperamento.

El segundo enfoque es más apropiado para aportar resultados cuantitativos, medibles. Instrumentos electrónicos pueden determinar con gran precisión la frecuencia instantánea de los tonos ejecutados. Los intervalos musicales a examinar de cerca son la tercera mayor y la sexta mayor, pues son los que presentan las diferencias más pronunciadas de una escala a otra (tabla 5.3). Obsérvese especialmente, que la nota superior en ambos intervalos resulta ser baja en la escala justa y aguda en la escala pitagórica (con respecto a la escala temperada). Los resultados experimentales muestran convincentemente que, en general, cantantes e instrumentistas de cuerda, ejecutan estos intervalos *subiendo* la entonación de las notas superiores de dichos intervalos (Ward, 1970). Esto parece señalar una preferencia por la escala pitagórica. No hay, sin embargo, que apresurarse en sacar conclusiones: ¡los mismos experimentos indican que también las quintas y las cuartas, e incluso la omnipotente octava, fueron

cantadas o tocadas, en promedio, aumentando la altura de la nota superior!⁵ Más que revelar la preferencia por una escala determinada (la pitagórica), estos experimentos señalan la existencia de una *tendencia universal*, hasta entonces inesperada, a *cantar o tocar todos los intervalos melódicos subiendo la entonación de las notas superiores*. Esta entonación 'agrandada' podría ser causada por el ligero desplazamiento de altura primaria de los componentes armónicos de un tono musical (Seccs. 4.3 y 5.2), lo cual lleva a un registro «levemente equivocado» en el procesador central de altura (para un análisis reciente y detallado, véase Hartmann, 1993). Otro resultado de estos experimentos, tal vez más significativo, es el hecho de que las fluctuaciones de altura de una nota dada son muy grandes a lo largo de una ejecución. Esto incluye tanto el vibrato como las variaciones de la altura media de una nota dada, a medida que ésta reaparece en el curso de la misma pieza de música. Estas fluctuaciones de altura de una misma nota escrita, cubren un rango de frecuencias mucho mayor que aquel determinado por las diferencias de frecuencia entre escalas diferentes (tabla 5.3); ¡en realidad, hace que estas últimas no tengan ninguna importancia práctica! De manera general, todos estos resultados apuntan al hecho que *los intervalos musicales son percibidos como categorías*; las fluctuaciones o diferencias de entonación reales son fácilmente ignoradas.

Hasta ahora hemos tratado de intervalos, es decir, de razones de frecuencias. ¿Qué decir con respecto a la altura absoluta per se? Una vez que una escala ha sido adoptada, es suficiente establecer la frecuencia de una sola nota, no importa cuál. No obstante, si exigimos que los instrumentos de frecuencias fijas sean fácilmente intercambiables en todo el mundo, la adopción de una altura «patrón» debe hacerse sobre la base de un acuerdo. Se ha convenido en establecer una frecuencia fundamental de 440 Hz para el «La central» del piano (La_4). Desde la aparición del diapason de horquilla, en el siglo diecisiete, distintos standards de frecuencias han sido usados. A lo largo de los últimos dos siglos, el patrón de frecuencia fue subiendo gradualmente desde alrededor de 415 Hz hasta tanto como 461 Hz⁶. Solo podemos esperar que el patrón actual se mantenga constante.

⁵ Existe un efecto recíproco: los intervalos justos son juzgados como sonando demasiado 'bajos' (Terhardt y Zick, 1975).

⁶ Esto tiene consecuencias serias para instrumentos históricos famosos que todavía se están usando en la actualidad. Un violín Stradivarius, por ejemplo, construido originalmente sobre la base de un standard de frecuencia de 415 Hz, hoy tiene que ser afinado más alto, lo cual implica una *mayor tensión* en sus cuerdas (relación (4.3)). Esto altera la cualidad (espectro) del tono. Un órgano barroco, construido también para un $La_4 = 415$ Hz, al subírsele la afinación a $La_4 = 440$ Hz, tiene que tener sus tubos labiales cortados parcialmente, con el fin de acortar el largo efectivo de los mismos (relación [4.6]).

En la escala temperada, todos los intervalos del mismo tipo (por ej., quintas, terceras mayores, etc.) son exactamente la «misma cosa», excepto en lo que se refiere a la altura real de sus componentes. Una melodía tocada en *Do* mayor no es en ningún aspecto diferente a la misma melodía tocada en *Re* mayor (salvo en lo que se refiere al rango de alturas cubierto). La idea de que ciertas tonalidades tienen cierto «carácter» o «color» no reviste ningún fundamento psicoacústico, como ha sido demostrado por experimentos realizados hace ya mucho tiempo (Corso, 1957). Puede haber, sin embargo, pequeñas diferencias entre los sonidos de diversas tonalidades, debido a razones físicas: la mayor aparición de teclas negras en el piano (las cuales son golpeadas de manera levemente diferente) en ciertas tonalidades,⁷ la mayor o menor aparición de cuerdas al aire en ciertas tonalidades en los instrumentos de cuerda, o el efecto de resonancias fijas o rango de formantes (Sec. 4.3) en los distintos elementos resonantes.

Una palabra final con relación a la *percepción absoluta de la altura*. Las pocas personas dotadas de la habilidad de reconocer o de reproducir con su voz una nota determinada en forma absoluta (esto se llama «oído absoluto») son, habitualmente, muy admiradas. Pero ocurre que hemos sido entrenados para prestar atención y almacenar en la memoria de largo término solo intervalos de altura relativos, porque ésta es la información más pertinente en un «mensaje musical». En otras palabras, nuestro cerebro ha sido entrenado para interpretar y almacenar una melodía como una secuencia de «transiciones» de altura, más que los valores de altura mismos; la información sobre la altura absoluta, si bien alcanza nuestro oído, es descartada como no esencial en el proceso cognitivo. A pesar de ello, esta información puede ser retenida por personas normales durante cortos intervalos de tiempo, que van desde los 10 segundos hasta varios minutos (Rakowski, 1972). Es muy posible que el «oído absoluto» pueda ser aprendido en una etapa temprana de la educación mental. Estas consideraciones, como los sorprendentes desplazamientos de altura detectados electrónicamente en la música «en vivo», que parecen pasar inadvertidos por el oyente, sugieren la necesidad de una redefinición operacional de la altura como *el correlato subjetivo de cada uno de los eventos acústicos en una secuencia de tonos musicalmente significativa* (Houtsma y Goldstein, 1972) ¡No debería sorprender a nadie, que esta definición haya sido - explícita o implícitamente - precisamente aquella usada por los músicos en todos los tiempos!

5.5 ¿Por qué existen las escalas musicales?

Nuestro oído es sensitivo a las ondas sonoras en un amplio rango de

⁷ ¡Por ejemplo, en la música para piano de Chopin!

frecuencias. Podemos detectar cambios de frecuencia diminutos; la DAP es típicamente del 0,5% o menos (Fig. 2.9). Sin embargo, la música occidental (y la de la mayor parte de las otras culturas) está basada en escalas, es decir, en transiciones y superposiciones de tonos que difieren entre sí en más de veinte veces el umbral de nuestra capacidad de resolución de frecuencia. ¿Por qué no hacemos música con cambios continuos de la altura, que suenen, por ejemplo, como las «canciones» de las ballenas y delfines (que tienen un sistema de comunicaciones acústico muy sofisticado, basado en «barridos» continuos de frecuencia)? ¿Por qué la altura siempre tiene que «saltar» en pasos discretos?

No hay respuestas simples a estas cuestiones. En primer lugar, recordemos que un tono musical determinado necesita un cierto período de tiempo mínimo para poder ser procesado en detalle por el cerebro (Sec. 3.4). Esto quizá haya evitado que los tonos cuya altura varíe de manera continua se hayan convertido en elementos básicos y duraderos de la música. En segundo término, observemos que diferentes culturas musicales usan, o han usado, diferentes escalas. Por lo tanto, las escalas están de algún modo relacionadas con - o fueron influenciadas por - el entrenamiento y la tradición. En tercer lugar, la mayor parte de los instrumentos primitivos eran de alturas fijas. También se justificó la aparición de las escalas sobre la base de la consonancia, lo que implicaría que las escalas aparecieron en conexión con la música polifónica. Sin embargo, las escalas ya estaban en uso cuando las melodías sólo eran cantadas (o acompañadas) monofónicamente al unísono (o a lo sumo, con octavas y quintas). Tal vez la razón neuropsicológica para la existencia de las escalas tenga que ver con el hecho que es más fácil para el cerebro procesar, identificar y almacenar en su memoria una melodía construida en base a alturas de valores discretos, que mantienen ciertas relaciones entre sí, dadas por la «familiar» serie de los armónicos. Procesar, identificar y almacenar en memoria un patrón de alturas que barre hacia arriba y abajo, pasando por todas la frecuencias posibles, de manera continua, requeriría muchos más «bits» de información que una secuencia discreta.

La explicación de la existencia de las escalas, es decir, de secuencias de tonos de alturas discretas, ha sido también intentada en función de relaciones *dinámicas en el tiempo* de un tono a otro, es decir sobre la base de intervalos *melódicos* más que armónicos. Esta línea de pensamiento se apoya en las sensaciones - tan importantes musicalmente, pero todavía poco exploradas desde el punto de vista psicofísico - de «dirección» o expectativa en secuencias de dos o más tonos, de dominio de una altura determinada, y de retorno a esa altura dominante (también llamada «tónica»). Por ejemplo, tendemos a asignar una dirección natural a una secuencia de dos tonos que es ascendente, cuando están separados por un semitono, y descendente cuando están separados por el intervalo de un tono. En ambos casos, nos inclinamos a asignar dominancia al segundo

tono; el dirigirse hacia este último equivale a la dirección natural, y el alejarse de él a la dirección no natural. De manera similar, una secuencia como *Do - Sol - Do - Sol - Do - Sol - . . .* «pide» terminar en *Do*, mientras que la secuencia *Do - Fa - Do - Fa - Do - Fa - . . .* «clama» por un *Fa* como nota final. Y si escuchamos *Mi - Sol - Mi - Sol - . . .* ninguno de los sonidos componentes se considera como final satisfactorio; ¡queremos escuchar un *Do*! Toda la orientación auditiva de la música diatónica occidental y el «significado» (Gestalt tonal) de mensajes musicales (Sec. 1.3), se basan sobre estos efectos.

A comienzos de siglo, Meyer (1900) y Lipps (1905) intentaron «explicar» la preferencia por ciertas terminaciones melódicas y el dominio de una altura tonal en términos de ciertas propiedades numéricas de la razón de frecuencias de un intervalo melódico. En los ejemplos precedentes, la nota que domina es aquella cuya frecuencia corresponde a una potencia de *dos* en la razón de números enteros. Por ejemplo, en $f_{\text{Sol}}/f_{\text{Do}} = \frac{3}{2}$ domina el tono más bajo; en $f_{\text{Fa}}/f_{\text{Do}} = \frac{4}{3}$ y en $f_{\text{Do}}/f_{\text{Si}} = \frac{16}{15}$ domina el tono más alto. Investigaciones posteriores tendieron a atribuir estos efectos principalmente al condicionamiento cultural. Pero la cuestión todavía sigue vigente: ¿Por qué éstas, y no otras preferencias, son las que surgen? Vale la pena observar, que cuando los intervalos musicales *Do - Sol*, *Do - Fa* y *Mi - Sol* dados arriba son considerados como elementos contiguos de la serie armónica (Fig. 2.19), la fundamental de esta serie es la que precisamente asume el rol dominante de punto de retorno (*Do - Fa - Do*, respectivamente). Otra vez, esta expectativa puede estar determinada por la «familiaridad» con las interrelaciones armónicas adquirida por nuestro procesador central de altura (Secs. 2.9, 4.8, 5.2, y Apéndice II) o, a un nivel cognitivo más alto, la familiaridad adquirida por estar inmersos en nuestra cultura occidental (Bharucha, 1994).

Otro fenómeno perceptivo de importancia musical, relacionado con secuencias de tonos, es el de la *fisión* melódica. Si se toca una melodía en la cual dos tonos se suceden alternando intervalos melódicos ascendentes y descendentes, la coherencia se pierde y se perciben *dos* (o más) líneas melódicas independientes cuando los intervalos son suficientemente grandes. En este caso, *nuestro cerebro tiende a agrupar los tonos según su proximidad en altura, más que por su proximidad en el tiempo*. Este efecto ha sido muy usado, especialmente en el período barroco, para hacer posible la ejecución de música a más de una voz en instrumentos monódicos. Para amplias reseñas sobre éste y otros fenómenos relacionados con secuencias temporales, véase van Noorden (1975), Deutsch (1982) y Sundberg (1992).

Podemos postular que todos estos efectos probablemente tengan su origen en el principio del «camino del menor esfuerzo» con el cual opera nuestro sistema nervioso: en el proceso de identificación de mensajes musicales (y también aquellos procedentes de otros sentidos), el sistema descarta primero todas las claves de información que juzga inútiles,

quedándose con solo un mínimo. Si logra la identificación, sigue adelante con el próximo paso; de lo contrario vuelve atrás, a la memoria de corto plazo, en busca de claves adicionales. Esto no solo sucede con estímulos de un tono individual, sino también en el procesamiento de un mensaje musical como un todo: *el sistema nervioso intenta usar toda información disponible de experiencias anteriores* (por ej., mensajes almacenados en la memoria de largo plazo) *para facilitar e incluso anticipar el proceso de identificación de nueva información entrante*. Esta estrategia de procesamiento neural es bien conocida en la percepción del habla y puede estar indicando la existencia de una relación natural entre música y lenguaje (véase la sección siguiente).

5.6 Procesos cerebrales cognitivos y afectivos

en la percepción musical:

¿por qué respondemos emocionalmente a la música?

La percepción musical, como la del habla, requiere tareas cognitivas complejas, en las cuales la información transmitida por las señales acústicas es analizada, almacenada, recobrada, comparada entre sí e interpretada. Sin embargo, todo el propósito de la música parece estar mucho más relacionado con los estados puramente afectivos que provoca, que al contenido básico de la información transmitida por los mensajes musicales. Quizá la demostración más convincente esté dada por el hecho de que disfrutamos de una pieza musical dada cada vez que la escuchamos, sin importar cuán familiarizados estemos con dicha pieza (¿quién, acaso, va a leer una novela policial dos veces?).

Con el fin de entender la percepción musical como un proceso cognitivo es necesario analizar en términos generales el rol de la motivación y de la emoción en las tareas cognitivas. Esta es todavía un área un tanto descuidada: el estudio de la conducta inteligente se centra principalmente en el procesamiento de la información por parte del cerebro, en términos de su contenido, lógica y organización, y de las reacciones resultantes en el plano de la conducta, pero con una atención mucho menos explícita a los factores somáticos que motivan la adquisición, el procesamiento y el almacenamiento de información, y a los estados afectivos designados generalmente como emoción. Justamente, es esto último lo que diferencia a la «función cerebral» de la «computación neural» (Apéndice II): las computadoras, por más sofisticadas que sean, ¡hoy día todavía no parecen estar preocupadas por su supervivencia! En todos los actos perceptivos conscientes, en última instancia es alguna *motivación* la que impulsa a la cognición; ésta conduce a un estado afectivo, el que a su vez impulsa, refuerza o modifica la motivación y atención perceptivas. En los animales,

esta relación circular es iniciada y controlada exclusivamente por la combinación de información instantánea recibida del entorno y las necesidades somáticas del momento. En los seres humanos, en cambio, este proceso puede iniciarse y controlarse exclusivamente a partir de la actividad cerebral interna, usando información sobre el entorno y el organismo recobrada de la memoria, en vez de aquella obtenida en tiempo real (Sec. 4.10).

Las funciones cognitivas del cerebro son manejadas principalmente por las redes corticales asociativas y en los lóbulos frontales, mientras que la motivación y la respuesta emocional están controladas por el *sistema límbico*. Pero todas estas estructuras cerebrales trabajan inextricablemente unidas, como queda evidenciado fisiológicamente por la estrechas interconexiones nerviosas existentes entre ellas. El sistema límbico es una parte del cerebro filogenéticamente vieja, que comprende varias estructuras localizadas entre las áreas corticales de asociación y el hipotálamo (hipocampo, amígdala, varios núcleos del tálamo). Junto con el hipotálamo (la parte del cerebro que integra las funciones del sistema nervioso autónomo y regula el sistema de información neuroquímica), el sistema límbico controla la entrada sensorial, dirige selectivamente el almacenamiento en la memoria según la importancia de la información, y activa la respuesta motriz (Fig. 5.7) con el objetivo específico de asegurar que la respuesta sea lo más provechosa para la preservación del organismo y de la especie en un medio ambiente complejo. En oposición a las redes corticales que controlan la conducta inteligente, muchos de los circuitos nerviosos del sistema límbico han sido «pre-cableados» antes del nacimiento, con funciones programadas a lo largo del lento curso de la evolución filogenética. La motivación y la emoción son manifestaciones de la función límbica, manifestaciones del principio que guía el sistema límbico, de asegurar que todos los procesos corticales sean llevados a

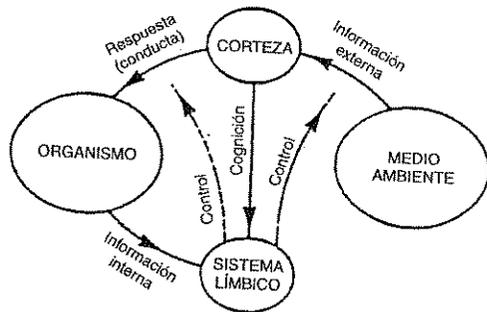


Figura 5.7 Rol del sistema límbico en la adquisición de información, cognición y conducta (Roederer, 1987).

cabo para el máximo beneficio del organismo (por ej. Panksepp, 1982, y referencias ahí dadas). El sistema límbico funciona de una manera asombrosa, dispensando sensaciones de recompensa o castigo cuando el organismo actúa de un modo que, según el dictado genético, es juzgado como beneficioso, o contrario a las posibilidades de supervivencia del organismo y/o de la especie, respectivamente.

El sistema límbico es activado tanto por señales somáticas como por aquellas procedentes del entorno (Fig. 5.7). En los seres humanos, sin embargo, este sistema puede responder a imágenes neurales evocadas internamente, disparadas en la corteza durante el proceso de pensamiento (Sec. 4.10) (Roederer, 1987). En otras palabras, motivación y emoción en los seres humanos pueden activarse sin relación explícita con el estado momentáneo del entorno y del organismo (por ej., estimulación sexual por medio de la imaginación). Recíprocamente, los seres humanos podemos desobedecer deliberadamente los dictados del sistema límbico (una dieta es un buen ejemplo).

Volvamos a la música. Los mensajes musicales no parecen transmitir ninguna información biológicamente relevante, contrariamente a lo que sucede con el habla, los sonidos emitidos por los animales y los sonidos del medio ambiente; así y todo, la mayoría de las personas reaccionan a mensajes musicales. ¿Qué pudo haber sido lo que en la evolución humana condujo a esto? ¿Ha resultado ser la música específicamente ventajosa para la raza humana? Desde luego que esta pregunta debe ser considerada parte de otra más amplia relacionada con la evolución de la motivación y reacción estética, y de la creatividad artística. En la búsqueda de respuestas (Roederer, 1984) debemos concentrarnos en esa capacidad distintiva de los seres humanos: el habla. El habla involucra al sistema auditivo y a las redes neurales centrales relacionadas; no existe un equivalente en la visión o en algún otro sentido⁸.

En la percepción del lenguaje hablado, el sistema auditivo es llevado a los límites de la percepción y la interpretación acústica. De ahí que sea concebible que con la evolución del lenguaje humano y la aparición de áreas corticales que se especializan en la percepción del lenguaje hablado, haya surgido un impulso a entrenar el sentido acústico en la tarea de reconocimiento sofisticado de patrones sonoros como parte de un *instinto humano, innato, de adquisición del lenguaje*. Durante las últimas etapas del desarrollo del feto, el sentido acústico ya empieza a registrar, pasivamente, el ambiente sonoro intrauterino. Al nacer, se produce una transición repentina a la respuesta activa, reflejada en la conducta, en la cual la comunicación acústica con la madre, o con quien cumpla ese rol,

⁸El lenguaje escrito apareció hace relativamente poco en la historia humana (algunas culturas primitivas no poseen lenguaje escrito aún hoy), y no requirió el desarrollo de centros corticales especializados en este tipo de procesamiento de información óptica.

va adquiriendo un papel fundamental. Se establece un ciclo de retroalimentación a través de la comunicación acústica, que contribuye a la relación emocional con la madre y al impulso que motiva la adquisición del lenguaje. Sonidos «musicales» sencillos y sucesiones rítmicas de sonidos (tales como los vocalizados por la madre) llaman la atención del bebé a escuchar, analizar y almacenar sonidos como preludio a la adquisición del lenguaje. Esto puede haber conducido a la aparición de una motivación por escuchar, analizar, almacenar y también vocalizar sonidos musicales, así como a una reacción emocional, o recompensa del sistema límbico, cuando eso se hace.

La motivación a descubrir simetrías y regularidades, a inferir posibilidades no explícitas, a predecir, a interpolar, a enfrentarse con la sorpresa de cambios repentinos o con la familiaridad de redundancias y repeticiones, así como también el impulso a explorar, diversificar y establecer prioridades, todos contribuyen a los elementos afectivos de la música. Hay elementos de carácter instantáneo, o de corto plazo, relacionados con las sensaciones subjetivas de timbre, consonancia, expectativa tonal, sentido de retorno a la tónica, y estructuras de largo plazo representadas por las líneas melódicas. Estos elementos afectivos pueden ser manifestaciones de recompensa del sistema límbico, en la búsqueda de contenido fonético del sonido y en la identificación de organización gramatical y contenido lógico de las señales acústicas. El hecho afortunado de que estos elementos estén presentes *en todo momento* y no puedan reprimirse, constituye el fundamento de la teoría moderna sobre la música (por ej., Lerdahl y Jackendoff, 1983; Bharucha, 1994).

En términos generales, es posible que la última «causa» de las sensaciones musicales producidas por un mensaje musical dado, resulte del efecto combinado del nivel de complejidad de la identificación del mensaje sonoro (es decir, del número total de operaciones neurales requeridas por unidad de tiempo), y de las asociaciones producidas en otros centros cerebrales. Por ejemplo, en una modulación repentina, o sea cuando se pasa de una tonalidad a otra en una pieza musical, el sistema nervioso auditivo debe construir rápidamente una nueva «lista de control» para las tareas de identificación sonora que se deberán realizar. Esta construcción rápida, una carga extra de miles de operaciones individuales, daría una sensación particular relacionada con el cambio de tonalidad. Personas sin sensibilidad musical, incapaces de experimentar estas sensaciones, son probablemente individuos cuyo mecanismo de identificación de mensajes musicales no tuvo la posibilidad de desarrollarse en todo su potencial; de ahí que, aunque el individuo escuche *todo* lo que escuchan las personas con sensibilidad musical, su sistema auditivo central no desarrolló la destreza suficiente para extraer información musicalmente relevante de superposiciones y secuencias de sonidos que no estén relacionadas directamente con el lenguaje hablado. Eventos tales como

los cambios de tonalidad son escuchados pero no interpretados, y por eso no producen ninguna respuesta afectiva.

Desde la más temprana edad, la mayor parte de las personas es expuesta solo a una clase muy limitada de estímulos musicales. El condicionamiento cultural no tarda en actuar, y la respuesta emocional empieza a estar influida por factores externos, algunos fortuitos, como ser el estado emocional que tenía una persona la primera vez que escuchó cierta pieza musical; otros más controlables, como el grado de repetición de formas musicales características pertenecientes a un estilo musical determinado; otros derivados del impulso innato a diversificar los potenciales de la creatividad humana. Esto último, el caso de la música, sería por ejemplo aprovechar desarrollos tecnológicos tales como la aparición de los instrumentos de teclado o, más recientemente, la de los sintetizadores electrónicos. Lo que permanece invariante en los instintos originales es: 1) el *hecho* de que existe una motivación a dirigir la atención hacia sonidos y formas musicales; 2) el *hecho* de que puede aparecer una reacción emocional; y 3) el *hecho* de que hay componentes de la música que son comunes a todas las culturas musicales.

Aun en el estado de desarrollo más avanzado, culturalmente condicionado, pueden encontrarse elementos de valor de supervivencia en la música. Igual que un buen discurso en público, la música puede servir para suscitar y mantener la atención de grandes masas de gente, haciendo que se ignoren impulsos límbicos normales durante largos períodos de tiempo. Todos sabemos cuál es el contenido de información del lenguaje, pero ¿cuál es el de la música? La música transmite información sobre estados afectivos. Puede contribuir a igualar los estados emocionales de un grupo de oyentes, del mismo modo que una conferencia oral puede contribuir a la igualación del estado intelectual (conocimiento) de la audiencia. Desafortunadamente, no hay todavía modelos cuantitativos para describir estados emocionales, como los que existen para los estados de información relacionados con el lenguaje. Pero el rol de la música en ritos supersticiosos y sexuales, en la religión, en el proselitismo ideológico, en la incitación militar, e incluso en la conducta antisocial, demuestra claramente el valor de la música para lograr coherencia de conducta en masas humanas. En el pasado remoto, esto pudo efectivamente haber tenido un valor importante para la supervivencia, a medida que el cada vez más complejo entorno humano demandó *acciones coherentes y colectivas* por parte de grandes grupos de personas.

Hasta ahora hemos omitido toda referencia al *ritmo* como un componente fundamental de la música; esta es una omisión particularmente crítica, por cuanto la aparición del ritmo siempre parece haber sido el primer paso en la evolución de una cultura musical dada. La propagación a través del tejido cerebral de un flujo de señales neurales cíclicamente variables, disparadas por patrones sonoros rítmicos, puede

de algún modo entrar en «resonancia» con los «relojes» naturales del cerebro que controlan las funciones periódicas del organismo y de la conducta. Estos relojes probablemente funcionan sobre la base de actividad neural que se propaga en circuitos cerrados, o que provienen de circuitos que tienen períodos propios de respuesta cíclica. Mucho más investigación científica hace falta para que podamos aventurarnos a dar una respuesta contundente a la cuestión planteada al comienzo de esta sección: ¿por qué respondemos emocionalmente a la música?

5.7 Especialización del habla y del procesamiento musical en los hemisferios cerebrales

En el capítulo introductorio aludimos brevemente a la notable división de tareas entre los hemisferios cerebrales izquierdo y derecho, en el cerebro humano (Sec. 1.6). Concluiremos este último capítulo con el tratamiento de este fenómeno, principalmente en sus aspectos pertinentes a la música (por ej., véase Scheid y Eccles, 1974; Bradshaw y Nettleton, 1981, y referencias dadas sobre el particular).

El cuerpo de los vertebrados exhibe una simetría bilateral, especialmente con respecto a los órganos involucrados en la interacción sensorial y motora con el ambiente. Esta simetría se extiende a los hemisferios del cerebro, donde la *corteza izquierda* está conectada al *lado derecho* del cuerpo y viceversa. Este cruzamiento se manifiesta principalmente en los sistemas capaces de registrar la dimensión de direccionalidad, tales como la visión y la audición (por ej., véase diagrama de flujo en la Fig. 2.26), y al control motor eferente de piernas y brazos. Este desarrollo probablemente obedezca a la necesidad de mantener juntos, en un mismo hemisferio cortical, los mecanismos de interacción con el entorno que conectan la información entrante con las instrucciones motoras salientes, para enfrentar eventos que están localizados en un mismo semiespacio del entorno. La imagen óptica es *físicamente* invertida en las lentes del ojo, proyectando el campo panorámico derecho sobre la mitad izquierda de la retina y viceversa, en cada ojo. Las mitades *izquierdas* de ambas retinas están conectadas a la corteza visual *izquierda*, con el fin de reunir en un solo hemisferio cerebral la información total que proviene del mismo semiespacio.

Como se mencionó en la Sec. 2.9, ambos hemisferios están conectados entre sí por unos doscientos millones de fibras del cuerpo caloso (y cerca de un millón de fibras de la comisura blanca anterior), con lo cual se restituye la unidad global de la representación ambiental en el cerebro. En las vías aferentes auditivas hay conexiones de más bajo nivel entre los canales de ambos lados (por ej., Fig. 2.26), a través de los cuales las

señales de los lados derecho e izquierdo pueden interactuar para dar información sobre la dirección del sonido.

En la evolución del cerebro humano, los vastos requerimientos de procesamiento de información que se presentaron con el desarrollo de la comunicación verbal, condujeron a la *especialización hemisférica*. En esta división de tareas, *el hemisferio «dominante» se ocupa de las funciones analíticas y secuenciales del lenguaje* (en el lado izquierdo en alrededor del noventa y siete por ciento de las personas - Penfield y Roberts, 1959). *El hemisferio menor se especializa en la percepción de relaciones holísticas, globales, sintéticas*⁹. Los centros del habla están localizados en un único hemisferio, un hecho bien conocido desde hace más de cien años, principalmente basado en autopsias realizadas en pacientes que habían manifestado defectos de lenguaje (afasias, alexias, anomias, agrafias) después de haber padecido hemorragias (síncope) en el hemisferio izquierdo (por ej., Geschwind, 1972). Las lesiones en el hemisferio derecho, por otra parte, causan perturbaciones en el reconocimiento de patrones visuales (Kimura, 1963), y pérdida de memoria tonal y tímbrica (Milner, 1967). En general, todas las tareas auditivas *no verbales* aparecen afectadas en estos pacientes. Más recientemente se han realizado notables experimentos con pacientes cuyo cuerpo caloso había sido seccionado transversalmente por razones terapéuticas («split-brain patients»)(Gazzaniga, 1970). Por ejemplo, estos pacientes no pueden describir verbalmente ningún objeto, palabra escrita o suceso localizado en la parte izquierda de su campo visual, porque la información sensorial pertinente, desplegada en el hemisferio derecho, no puede ser transferida a los centros del habla a causa del cercenamiento del cuerpo caloso. Un procedimiento usado en pacientes sin traumas físicos en el cerebro es la inyección de un barbitúrico en la arteria carótida, para anestesiarse brevemente uno de los hemisferios (un recurso preoperatorio que a veces se emplea para confirmar la localización izquierda/derecha de los centros del habla); una serie de pruebas con dichos pacientes (véase resumen en Borchgrevink, 1982) confirma que en música (pero no en el habla) la altura y la tonalidad son procesadas por el hemisferio derecho, mientras que la comprensión y producción normal del habla, así como el ritmo musical, son tareas del hemisferio izquierdo. Hoy día existen técnicas no invasivas que pueden ser usadas en individuos sanos y normales destinadas a estudiar la especialización hemisférica. Una de tales técnicas, llamada test de audición dicótica, se basa en el hecho, mencionado en la sección 2.9, que la información auditiva que fluye a través de los

⁹ Se podría esperar que el procesamiento del habla esté localizado en el mismo hemisferio que controla la mano predominante (por ej., hemisferio izquierdo para la mano derecha). Sin embargo, es mucho más frecuente el ser zurdo, que el mero tres por ciento de casos de procesamiento del habla en el lado derecho.

Tabla 5.4 Listado comparativo de la especialización hemisférica en las tareas auditivas (basado en Bradshaw y Nettleton, 1981).

Hemisferio izquierdo	Hemisferio derecho
Consonantes explosivas	Vocales tenidas
Atributos fonológicos, sintaxis	Atributos estereotipados, rima en poesía, contaxis o remise
Comprensión del habla	Entonación del habla, sonidos ambientales y de animales
Lenguaje	Contenido emocional del habla
Análisis de sonidos sin sentido en el habla	Altura, timbre, tonalidad, armonía
Texto hablado (contenido verbal)	Texto cantado (contenido musical y fonético)
Ritmo, secuencias melódicas de corto plazo	Memoria holística
Memoria verbal	Memoria tonal

principales canales auditivos contralaterales (Fig. 2.26) tiende a ignorar cualquier información contradictoria que se propague por la ruta ipsilateral. Por lo tanto, cuando ambos oídos reciben información mutuamente contradictoria, la corteza auditiva *izquierda* prestará más atención a la información procedente del *oído derecho* (aunque también reciba información del oído izquierdo - Fig. 2.26), y viceversa. Efectivamente, se nota una superioridad del oído derecho en tareas de reconocimiento del habla, y una superioridad del oído izquierdo para analizar melodías (por ej., Kimura, 1963). Finalmente, por medio de la electroencefalografía, de la tomografía por emisión de positrones y de la resonancia magnética nuclear, se obtienen mapas de la localización espacial de la actividad eléctrica o metabólica en el cerebro durante una tarea mental determinada que ratifican, de manera contundente, los resultados arriba mencionados. La tabla 5.4 (basada en una reseña de Bradshaw y Nettleton, 1981, y en sus referencias) resume algunas características básicas de la especialización hemisférica en las tareas auditivas.

¿Por qué surgió esta curiosa dicotomía en las funciones hemisféricas en el curso de la evolución humana?¹⁰ La explicación más plausible de este desarrollo es la necesidad de mantener las áreas responsables del procesamiento acústico del habla y las responsables de dirigir el 'output' gestual, vocal y mímico lo más cerca posible entre sí, con el fin de *minimizar las demoras de transmisión* entre las redes neurales participantes. ¡Las complejas operaciones secuenciales en el procesamiento del habla simplemente no podían tolerar el retardo que involucra la transmisión de las señales neurales de un hemisferio a otro! (aproximadamente

¹⁰Nuestros antepasados primates no exhiben una especialización hemisférica diferenciada (esto, sin embargo, está sujeto a controversias: algunos animales presentan una asimetría operacional en sus hemisferios cuando se trata de tareas de procesamiento secuencial en oposición a tareas de naturaleza holística).

cincuenta milésimas de segundo)¹¹. En consecuencia, una porción apreciable del «espacio de procesamiento» en el hemisferio izquierdo, quedó vedada a otras tareas, más lentas, de naturaleza integradora, holística, las cuales «por defecto» fueron asumidas por el hemisferio derecho. Es importante notar que la especialización de los hemisferios cerebrales es de una naturaleza muy fundamental: *involucra dos modos de operar muy diferentes*. Un modo representa el *análisis secuencial* de partes de información, tal como se requiere en el procesamiento del lenguaje. El otro modo involucra *integración espacial* o síntesis de patrones momentáneos de actividad neural, para determinar las cualidades holísticas de los estímulos (por ej., Papçun y otros, 1974). Sin embargo, ambos modos deben coexistir y cooperar para procesar la información, y programar la respuesta del organismo en un complejo entorno humano¹². En particular, las tareas secuenciales (como el escudriñamiento visual) pueden ser necesarias para el reconocimiento de patrones y la construcción de imágenes y, recíprocamente, la imaginación holística puede ser una operación colateral necesaria para la programación secuencial.

Dado que la música se procesa principalmente en el hemisferio menor, ¿significa ésto que la música requiere principalmente operaciones sintéticas de reconocimiento de cualidades holísticas? Considerando el reconocimiento de tonos compuestos, esto efectivamente parece estar de acuerdo con las teorías de «ajuste de moldes» [«template-fitting»] sobre la percepción de la altura (Sec. 4.8 y Apéndice II). *La magnitud holística en un estímulo musical es la distribución espacial momentánea de actividad neural* (correspondiente a las máximas resonancias sobre la membrana basilar), que conduce a la altura de un tono compuesto (Sec. 4.8), a la discriminación de sonidos múltiple (Sec. 5.1), a la consonancia (Sec. 5.2), y a la expectativa y sensación de retorno a la tónica (Sec. 5.5). Otra magnitud holística es la distribución relativa de la *intensidad* de actividad, dada por el espectro de potencia, que lleva al timbre y al reconocimiento de la fuente (Sec. 4.9). Podemos trazar aquí una buena analogía con la visión: el patrón sonoro (patrón temporal) es «proyectado» como un patrón espacial sobre la membrana basilar. El resultado es una imagen *espacial*, muy parecida a la proyectada sobre la retina. De aquí en adelante, ambos sistemas operan sobre sus respectivas informaciones de manera formalmente análoga, lo que conduce eventualmente a las sensaciones musicales y pictóricas, respectivamente.

Surge una aparente paradoja cuando consideramos *melodías* y la dependencia temporal de los mensajes musicales. ¿No requieren ellas

¹¹ ¡Obsérvese que, por la misma razón, lo que limita la velocidad de procesamiento de las computadoras actuales es la extensión espacial!

¹² Por «entorno humano» entendemos a un ambiente que incluye otras personas con las cuales nos debemos comunicar.

tareas de secuenciación, es decir, operaciones del hemisferio *dominante*? No necesariamente. *Nuestro cerebro reconoce que los mensajes musicales son de naturaleza holística*, patrones temporales de largo plazo, más que secuencias de corto plazo. El fenómeno de fisión melódica (pag. 200) es un ejemplo convincente de esto. Dicho con otras palabras, *la música parece ser reconocida por nuestro cerebro como la representación de imágenes auditivas integrales, holísticas* (la estructura armónica), *cuya sucesión (de largo plazo) en el tiempo tiene en sí misma un valor holístico como Gestalt* (contornos melódicos). Aunque inicialmente la atención humana a los sonidos y mensajes musicales puede ser el resultado de algún temprano instinto a prepararse para la adquisición del lenguaje (Sec. 5.6), la estructura y gramática musical son bien diferentes de aquellas del lenguaje (Jackendoff y Lerdahl, 1982).

Como se dijo antes, el lóbulo temporal izquierdo se especializa en el procesamiento de entradas [inputs] «verbales», y el derecho en el análisis de entradas «musicales». Un resultado interesante es el siguiente: un discurso totalmente carente de significado, obtenido, por ejemplo, reproduciendo en sentido inverso una cinta de grabación, va a ser tratado preferentemente por el hemisferio dominante (Scheid y Eccles, 1974). Esto confirma el carácter de corto plazo de las operaciones secuenciales pertinentes al habla mencionadas arriba: estas operaciones representan una etapa de procesamiento de la información acústica *previa* al reconocimiento del contenido conceptual. En una línea similar, tests de audición dicótica (Bever y Chiarello, 1974) demostraron que cuando sujetos *con experiencia musical* se centran sobre los detalles estructurales de sucesiones de sonidos de corto plazo en una melodía, las estrategias de procesamiento analítico correspondientes son manejadas principalmente por el hemisferio *dominante*. Una situación inversa se da con un texto *cantado*. Se encontró que pacientes con severos problemas en el habla (afasias) son capaces de cantar, con palabras claramente comprensibles, una canción aprendida antes de que el trauma hubiera surgido - pero no pueden hablar ese mismo texto. Esto sugiere que el habla en la expresión musical es procesada preferentemente en el hemisferio *menor*. En estos casos, el lenguaje presenta una característica que tiene mucho de sintético y holístico (Scheid y Eccles, 1974) como en una poesía.

Todo esto está muy relacionado con la comprensión de la evolución de la música occidental. En un sentido amplio, podemos describir esta evolución como una transición gradual entre dos configuraciones extremas. En un extremo encontramos estructuras sonoras espaciales (armónicas) y temporales (melódicas), claramente definidas, repetidas enfáticamente, cada una de las cuales representa la cualidad de un todo indivisible (por ej. un acorde determinado, y la progresión de una voz o de acordes, respectivamente). En el otro extremo (al cual nos estamos dirigiendo en la música contemporánea) identificamos formas sonoras cuyo valor

fundamental se reconoce en el estado momentáneo de rasgos temporales de corto plazo. A la luz de lo dicho con relación a la especialización hemisférica, podemos especular que estas dos configuraciones extremas están íntimamente ligadas a las dos estrategias fundamentales de procesamiento del cerebro humano. Solo el futuro dirá si las actuales tendencias en música representan un esfuerzo más o menos al azar para romper con las formas tradicionales (que en parte habían surgido de modo natural como resultado de las propiedades físicas y neuropsicológicas del sistema auditivo humano), o si estas tendencias pueden ser canalizadas hacia una exploración y explotación premeditadas de un vasto territorio, todavía virgen, de las capacidades de procesamiento de nuestro sistema nervioso central.

Para terminar, volvamos a la cuestión de la simetría bilateral que presentan los cuerpos y los cerebros en los vertebrados. Gur et al. (1980) encontró que hay más materia gris (consistente de neuronas estrechamente conectadas) que materia blanca (líneas de transmisión axónica mielinadas) en la corteza humana del hemisferio izquierdo que en el derecho. Esto puede indicar que el hemisferio izquierdo tiende a procesar información *dentro* de las regiones corticales, mientras que el hemisferio derecho opera *entre* regiones corticales. Otra notable asimetría encontrada¹³, muestra que el *plano temporal* - área cortical que tiene un rol clave en la comprensión del lenguaje - es significativamente mayor en el lóbulo temporal izquierdo que en el derecho, en el 65% de los cerebros examinados (Geschwind y Levitzky, 1968). En el 24% de los casos no se encontró una asimetría significativa, y en el 11% restante la asimetría estaba invertida. Tal asimetría no se encuentra en los cerebros de primates no humanos. Es muy sugestivo relacionar esta asimetría con la asimetría de los modos de operación de ambos hemisferios. Si bien un 65% es mucho menor que el porcentaje (alrededor del 97%) de casos con los centros del habla localizados en el hemisferio izquierdo, es muy posible que un plano temporal derecho relativamente grande en el restante 32% sea un indicio de un talento más grande para el procesamiento sonoro no verbal. Es más, se ha sugerido (Scheid y Eccles, 1974) que *un plano temporal derecho agrandado podría ser indicio de una habilidad musical innata*. Esta hipótesis, de ser confirmada estadísticamente a través de análisis sistemáticos post-mortem¹⁴, daría un fundamento anatómico a la transmisión hereditaria de la musicalidad.

¹³ Detectada incluso en bebés y fetos (Scheid y Eccles, 1974).

¹⁴ Músicos: ¡donad vuestros restos a una causa científica!

Apéndice I

Algunos aspectos cuantitativos del mecanismo de frotamiento

Consideremos una situación ideal: una cuerda muy, muy larga, frotada en el punto A por un arco infinitesimalmente angosto (Fig. AI.1). El arco se mueve con una velocidad b hacia arriba. Supongamos además que al comienzo (tiempo t_0) la cuerda está adherida al arco. Esto significa que el punto de contacto A (en realidad deberíamos decir el segmento de contacto) también se mueve hacia arriba con velocidad b . El resultado será una deformación de la cuerda en forma de una onda transversal que se propaga del punto A tal como se muestra para distintos instantes t_1, t_2, t_3 , en la figura AI.1. Dado que las ondas transversales se propagan con una velocidad V , dada por la relación (3.3), mucho más alta que la velocidad de frotamiento b , la pendiente b/V de las porciones inclinadas de la cuerda AP, AQ será en realidad muy pequeña. Bajo estas condiciones, la fuerza transversal F aplicada por el arco (no debe ser confundida con la presión de arco, la cual sería perpendicular al papel) está en equilibrio con las proyecciones sobre OA de ambas fuerzas de tensión T . Esto significa que $F = 2T b/V$. En un régimen de fricción estática, con la cuerda adherida al arco, la fuerza F debe ser menor que un cierto umbral F_s , llamado límite de fricción estática. Resultados experimentales muestran que este límite es proporcional a la «presión» de arco: $F_s = \mu_s P$. μ_s (letra griega mu) es el coeficiente de fricción estática; depende de la «rugosidad» de las superficies en contacto (en este caso, de la cantidad de resina en las cerdas del arco). La condición para la «adherencia» es entonces $F = 2T b/V < \mu_s P$.¹ Del mismo modo, la condición para el deslizamiento del arco sobre la cuerda será $F = 2T b/V > \mu_s P$. Dado que las cantidades V, T y μ_s son parámetros constantes para una cuerda dada, podemos sintetizar ambas expresiones de la siguiente manera:

¹ < significa «menor que»; > significa «mayor que».

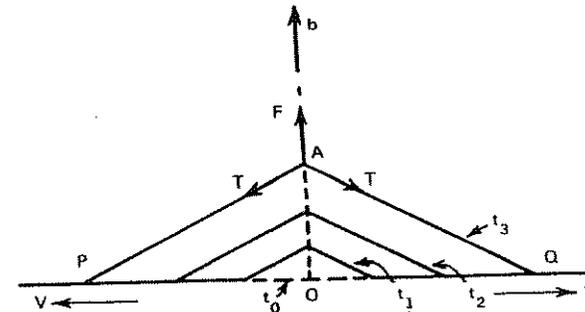


Figura AI.1 Progresión idealizada de la deformación de una cuerda larga frotada en el punto A con velocidad constante b .

Cantidad controlable por el ejecutante	Cantidad fija para cada cuerda	Tipo de movimiento de la cuerda con respecto al arco	
$\frac{b}{P}$	$< \frac{\mu_s V}{2T}$	adherencia	(AI.1)
	$> \frac{\mu_s V}{2T}$	deslizamiento	

Adviértase que en las relaciones (AI.1) lo que importa es la razón entre la velocidad del arco y la presión del arco, no b o P aisladamente. El cociente b/P define de este modo la naturaleza del movimiento de la cuerda frotada. ¿Qué pasa si ya desde el principio la cuerda se está deslizando (relación inferior en la ecuación (AI.1))? En ese caso la velocidad v del punto de frotamiento A será menor que la velocidad b del arco (incluso puede ser opuesta a ésta). Tendremos un régimen de fricción dinámica, en el cual la fuerza $F = 2T b/V$ (Fig. AI.1) es proporcional a P , pero también depende de la velocidad relativa $b - v$ entre arco y cuerda (velocidad del deslizamiento). Representamos esto en la forma $F = \mu_D P$, donde μ_D es el coeficiente de fricción dinámica, que depende de la velocidad relativa $b - v$ (es una función de $b - v$). Por lo tanto, durante el régimen de deslizamiento:

$$\frac{b}{P} = \mu_D \frac{V}{2T} \tag{AI.2}$$

Si conociéramos la dependencia de μ_D con la velocidad de resbalamiento $b - v$, podríamos usar la expresión (AI.2) para determinar la velocidad v del punto de contacto A de la cuerda. Nuevamente, esta relación está regida por la razón b/P . Obsérvese cuidadosamente: lo que está regido por este cociente es la diferencia $b - v$, es decir, la velocidad de

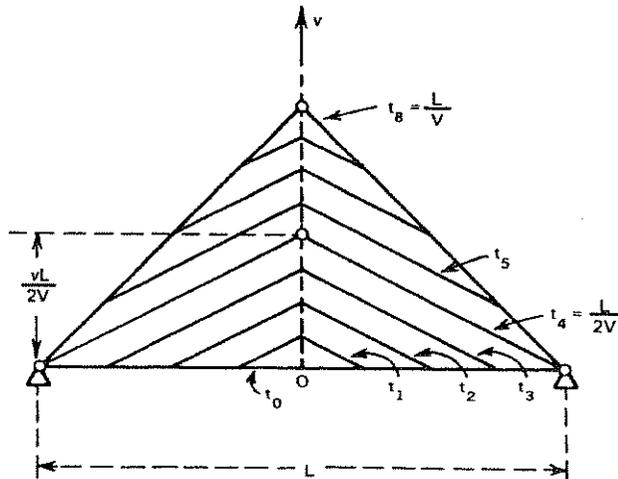


Figura A1.2 Lo mismo que en la figura A1.1, para una cuerda con extremos fijos.

la cuerda *con respecto al arco*. Cuanto mayor sea b , tanto más grande será v , para un valor determinado de b/P .

Mientras este cociente determina el tipo de movimiento de la cuerda (adherencia vs. deslizamiento, véase figura A1.1), la velocidad del arco determina la velocidad real de la cuerda (para un valor dado de b/P). Por lo tanto, si se incrementa la velocidad del arco, pero *al mismo tiempo* se incrementa la presión de arco de manera tal que el cociente b/P se mantenga constante, la naturaleza del movimiento de la cuerda no cambiará en absoluto - solo su velocidad se incrementará linealmente con b . Esto lleva a un aumento de la amplitud, es decir, de la intensidad del sonido. En otras palabras: *la amplitud de la vibración de una cuerda frotada (sonoridad del tono) está controlada exclusivamente por la velocidad del arco, pero con el fin de mantener constante el tipo de movimiento de la cuerda (o sea, el timbre del tono), la presión de arco debe mantenerse proporcional a la velocidad de frotamiento.*

Consideremos un caso que se acerca un poco más a la realidad: una cuerda de longitud finita L , frotada con un arco angosto en el punto medio O (Fig. A1.2). Esta figura muestra esquemáticamente la forma de la cuerda cuando comenzamos su frotamiento (otra vez, las pendientes en la figura aparecen *altamente* exageradas). v es la velocidad del punto medio (deslizamiento si $v < b$, o adherencia si $v = b$). Obsérvese que en el instante $t_4 = L/2V$, la primera «onda» (de pendiente v/V) ha alcanzado los extremos de la cuerda. Ahí, la onda se refleja y se superpone con la onda incidente, produciendo los perfiles doblados mostrados para los

instantes t_5 a t_7 . Luego, en $t_8 = L/V$, ocurre algo nuevo (Fig. A1.2): *la pendiente cambia repentinamente en el punto de frotamiento*. Esto cambia la expresión de la fuerza F , pudiendo surgir un nuevo régimen (por ej., deslizamiento, si previamente había adherencia). No podemos seguir este tratamiento sin considerables complicaciones matemáticas (Keller, 1953). Simplemente, nótese que estos cambios fundamentales en la forma de la onda (cuando la onda se refleja en los extremos fijos) ocurren siempre en instantes que son múltiplos enteros de L/V , una cantidad que es totalmente independiente del mecanismo de frotamiento. Nótese que la inversa de L/V aparece en la expresión (4.3) de la frecuencia fundamental de la cuerda vibrante; el lector puede visualizar, por lo tanto, cómo esa frecuencia (y todos los armónicos superiores) puede ser producida (y mantenida) por medio del mecanismo de frotamiento, e inferir de la figura A1.2 (con un poco de imaginación extra) que, *en su movimiento vibratorio, una cuerda frotada siempre tendrá una forma instantánea compuesta por segmentos rectos*; este resultado ha sido verificado experimentalmente hace ya largo tiempo.

Apéndice II

Algunos aspectos cuantitativos de los modelos del procesador central de altura tonal

En este Apéndice continuamos nuestra discusión de la percepción de la altura de sonidos compuestos, cuestión tratada en las secciones 2.9, 4.8 y 4.10. Mostraremos cómo usando solo un poco de álgebra, el modelo del 'ajuste de moldes' [template fitting] (Goldstein, 1973) puede explicar algunas características cuantitativas de la percepción de la altura de sonidos compuestos. En la segunda parte especularemos acerca de modelos neurales que pueden ejecutar las funciones de un procesador central de altura, así como otras funciones cognitivas.

El modelo de Goldstein (Sec. 2.9) se basa en la suposición de que la información neural acerca de las posiciones de los máximos de resonancia sobre la membrana basilar no está definida de manera precisa, fluctuando estadísticamente alrededor de valores promedio. El ajuste del molde, que conducirá a la sensación de altura subjetiva será aquel que minimice la diferencia entre los valores del molde y los de la señal real. Para ilustrar cómo esto podría funcionar, apliquemos este proceso a los experimentos de Smoorenburg (1970) (Sec. 2.7, pág., 58). Tomemos un estímulo que consiste en dos tonos de frecuencias $f_a = 1000$ Hz y $f_b = 1200$ Hz. Estas frecuencias corresponden *exactamente* a los armónicos quinto y sexto de una fundamental $f_1 = 200$ Hz. Nuestro molde está representado por un conjunto de frecuencias en relación de armónicos $f_1, 2f_1, \dots, nf_1, \dots$, cuya fundamental f_1 puede variarse a voluntad. El proceso de ajuste consiste en encontrar una frecuencia f_1 para la cual dos armónicos *sucesivos* nf_1 y $(n+1)f_1$ coincidan con (o estén lo más cerca posible de) las frecuencias del estímulo f_a y f_b . En esta etapa no tiene importancia el número de orden de los dos armónicos (o sea el valor de n), siempre que la correspondencia sea la mejor entre todas las posibles (que las diferencias $[nf_1 - f_a]$ y $[(n+1)f_1 - f_b]$ sean lo más pequeñas posible). En nuestro ejemplo, un ajuste es el mejor de todos: para $n = 5$ y $f_1 = 200$ Hz ambas diferencias de frecuencia son exactamente cero: el ajuste es perfecto. El lector puede verificar fácilmente que *no existe* otra frecuencia fundamental

f_1 , ni otro valor n , que den una correspondencia perfecta. Obsérvese con atención que un ajuste determinado requiere la estimación de *dos* valores: el orden armónico n y la frecuencia fundamental f_1 . La tesis de la teoría es que esta última frecuencia es la que corresponde a la sensación de altura subjetiva, y única, producida por el estímulo bitonal. Todo esto debería valer para un estímulo de más de dos tonos pertenecientes a una misma serie de armónicos, como ocurre en un sonido compuesto real, llevando a la frecuencia f_1 correspondiente a la altura percibida en ese sonido compuesto (sin importar si la frecuencia f_1 está realmente presente en el estímulo original).

Como en los experimentos de Smoorenburg, cambiemos ahora las frecuencias del estímulo al siguiente par *inarmónico*: $f_a = 1050$ Hz y $f_b = 1250$ Hz. No hay *ninguna* serie de armónicos que contenga estas frecuencias como componentes vecinos. ¿Cómo funciona en este caso el 'template matching'? Según la teoría descrita más arriba (en una versión muy simplificada), debemos hallar dos valores, n y f_1 , de tal manera que los armónicos vecinos nf_1 y $(n+1)f_1$ minimicen el *error cuadrático medio relativo* ϵ :

$$\epsilon^2(n) = \frac{1}{2} \left[\left(\frac{f_a - nf_1}{f_a} \right)^2 + \left(\frac{f_b - (n+1)f_1}{f_b} \right)^2 \right] \quad (\text{AII.1})$$

Para cada valor de n , habrá una frecuencia f_1 que minimice el valor del error $\epsilon(n)$, el cual puede ser hallado usando cálculo diferencial elemental¹:

$$f_1 = \frac{f_a}{n} \frac{[1 + (n+1)/n(f_a/f_b)]}{[1 + (n+1)/n(f_a/f_b)]^2} \quad (\text{AII.2})$$

Insertando esta expresión de f_1 en la expresión (aII.1), obtenemos el valor que corresponde al error de ajuste $\epsilon(n)$:

$$\epsilon^2(n) = \frac{1}{2} \frac{[1 - (n+1)/n(f_a/f_b)]^2}{[1 + (n+1)/n(f_a/f_b)]} \quad (\text{AII.3})$$

En lugar de trabajar con el error $\epsilon(n)$ podemos hacerlo con su recíproco

$$Q(n) = 1/\epsilon(n) \quad (\text{AII.4})$$

que llamaremos «calidad de ajuste». Un ajuste perfecto ($\epsilon = 0$) da $Q = \infty$ («ajuste infinitamente bueno»).

Para diferentes valores de n , obtenemos diferentes frecuencias fundamentales f_1 y diferentes valores de calidad de ajuste $Q(n)$. De este modo llegamos a una serie de valores $Q(n), Q(n+1), \dots$. Si entre éstos hay

¹ El lector puede verificar que esto da la respuesta correcta para el ejemplo armónico precedente, con $n = 5$ y $f_1 = 200$ Hz, valores para los cuales resulta $\epsilon = 0$ (ajuste perfecto).

110

uno que se destaque como el más alto, los valores correspondientes de f_1 y n representan la mejor correspondencia posible y f_1 efectivamente corresponde a la altura que se escucha! En nuestro ejemplo de una señal de dos frecuencias $f_a = 1050$ Hz y $f_b = 1250$ Hz, obtenemos, para $n = 4$ $f_1 = 256,0$ Hz y $Q(4) = 41$; para $n = 5$ $f_1 = 209,2$ Hz y $Q(5) = 251$; para $n = 6$ $f_1 = 176,8$ Hz y $Q(6) = 99$. Es evidente que $n = 5$ conduce al valor más alto de Q . La frecuencia correspondiente (209,2 Hz) efectivamente es la altura subjetiva que se percibe con mayor claridad cuando se presenta este complejo inarmónico de dos tonos (véase Smoorenburg, 1970). Obsérvese que $n = 6$ y, en menor medida, $n = 4$, también dan calidades de ajuste no despreciables; esto explica el hecho experimental de que las frecuencias correspondientes $f_1 = 176,8$ Hz y $f_1 = 256,0$ Hz pueden ser identificadas como sensaciones de altura «secundarias», aunque con mucho mayor dificultad. Este modelo por lo tanto ofrece una explicación cuantitativa de las sensaciones de alturas múltiples o ambiguas que aparecen en la audición de sonidos inarmónicos. Si desplazamos las frecuencias de los tonos que componen el estímulo, alejándolas aún más de la armonicidad (pero manteniendo siempre constante la diferencia de 200 Hz entre ambas frecuencias), obtenemos los resultados expuestos en la Tabla AII.1.

Nótese que para el par 1100/1300 hay *dos* ajustes que se destacan: los correspondientes a $n = 5$ y $n = 6$. Obsérvese también en la Tabla AII.1 cómo el orden armónico n que da el valor más alto de Q salta de 5 a 6 cuando aumenta la frecuencia central del par f_a y f_b , mientras que la altura que corresponde a mejor ajuste salta hacia abajo, desde un valor mayor que 200 Hz a otro debajo de dicha frecuencia. En general, a medida que f_a y f_b ascienden (manteniendo fija la diferencia de 200 Hz entre ellas), la sensación de altura subjetiva principal «oscila» alrededor de la frecuencia de repetición dada por $f_a - f_b$ (200 Hz), coincidiendo con ésta en las posiciones armónicas en las que f_a y f_b son armónicos contiguos. Las alturas múltiples o ambiguas aparecen con mayor claridad cuando el

Tabla AII.1 Valores de f_1 (AII.2) y $Q(n)$ (AII.4) (resaltados en negrita), para un tono compuesto de dos tonos puros, de frecuencias f_a y f_b .

f_a/f_b (Hz)	$n = 4$	$n = 5$	$n = 6$	$n = 7$
1000/1200	$f_1 = 240,0$ Hz $Q(n) = 49$	200,0 ∞	169,0 71	146,0 41
1050/1250	256,0 41	209,2 251	176,8 99	153,0 49
1100/1300	267,1 36	218,3 131	184,5 155	159,7 60
1150/1350	278,2 32	227,4 91	192,3 323	166,5 75
1200/1400	289,3 29	236,6 71	200,0 ∞	173,2 97

estímulo está aproximadamente en la mitad entre dos situaciones armónicas. Se puede hacer la suposición de que *la calidad de ajuste Q representa la «claridad» o inteligibilidad de la correspondiente sensación de altura*. Es importante señalar que los valores teóricos obtenidos mediante este procedimiento de «ajuste de molde» no coinciden con los resultados experimentales cuando el valor de n es mayor que 7. Los *sonidos resultantes* del tipo (2.5) y (2.6) (Sec. 2.5) influyen en el proceso como *señales adicionales* de altura bajas (por ej., Plomp, 1976).

Este método del ajuste puede extenderse a estímulos multitonales. Es ilustrativo usar una versión extendida para predecir la altura de combinaciones tales como una superposición de sonidos compuestos, en los acordes musicales² (véase más adelante). Es interesante notar que este modelo de extracción de altura por medio del 'ajuste de molde' [template matching] funciona de manera formalmente análoga al mecanismo de acomodación de altura en un instrumento de viento (pág. 154).

Una consecuencia musicalmente importante es la siguiente. Las relaciones (AII.2) o (AII.3) muestran que cuando se desafina uno de los armónicos superiores en una vibración compleja ($f_n = nf_1 + \delta$) el efecto sobre la altura subjetiva es *muy* pequeño (del orden de $\delta/2n$ o menos). Por lo tanto, las pequeñas inarmonicidades que presentan los sobretonos más altos en cuerdas vibrantes (pág. 125) ejercen un efecto despreciable sobre la altura subjetiva resultante.

Esbozemos ahora un modelo neural para el procesador central de altura, siguiendo la línea de la «matriz que aprende» invocada por Terhardt (1974). Haremos esto meramente como un *ejercicio académico*; no pretendemos formular otra teoría sobre un procesador central de altura. La figura AII.1 muestra un 'esquema de cableado neural' capaz de llevar a cabo las operaciones necesarias para la extracción de altura y el rastreo de fundamental, basadas en la «matriz que aprende» de Terhardt (1974). En esta figura se supone que las fibras horizontales conducen señales nerviosas, combinadas de ambas cócleas, al procesador de altura *primaria* (altura espectral). El output correspondiente (hacia la derecha, en la figura) lleva información sobre cada componente armónico presente en la vibración compleja, pero se supone que esta información es normalmente descartada en las etapas superiores del procesamiento del sonido. En nuestro modelo, estos axones horizontales son interceptados por dendritas (Sec. 2.8) pertenecientes a un haz de neuronas vertical, tal como se muestra en la Fig. AII.1. Supongamos que inicialmente (al nacer), las conexiones sinápticas activas están distribuidas como se muestra para las neuronas *K*, *L* y *M*. También supondremos que para alcanzar el umbral de disparo,

² Hay que tener sumo cuidado al extender las relaciones (AII.1) y (AII.2) al caso multitónico; sumar términos en el numerador y el denominador *no da* el resultado correcto, salvo que los componentes del estímulo sean *casi* armónicos.

cada neurona vertical debe ser activada en muchos contactos sinápticos prácticamente al mismo tiempo. Se desprende claramente de esta figura que, en un cerebro acústicamente virgen, la distribución de la actividad neural de salida correspondiente a las neuronas verticales (a lo largo de la dimensión x) será idéntica a la distribución de entrada en las fibras horizontales (a lo largo de la dimensión y).

Nuestra suposición siguiente es, de acuerdo a la teoría de Terhardt, que a medida que los oídos son expuestos repetidamente a tonos armónicos, se activarán nuevos contactos sinápticos entre una neurona vertical y todos aquellos axones horizontales que tienen la mayor probabilidad de estar disparando al mismo tiempo (la «esencia» del proceso de aprendizaje en el sistema nervioso, pág. 179). Cuando se presenta un tono compuesto de espectro armónico, cuya fundamental es, por ejemplo, f_1 , la neurona K (Fig. AII.1) responderá, inicialmente, solamente a esa fundamental. Pero a medida que se repita ese estímulo (el tono compuesto), la dendrita K desarrollará contactos sinápticos activos con todas aquellas fibras horizontales cuyas frecuencias características correspondan a los armónicos superiores de ese tono (pág. 65). Como resultado, la neurona vertical queda «sintonizada» a la serie completa de los armónicos de f_1 , como se muestra para la neurona K' en la Fig. AII.1. Desde el punto de vista físico, las neuronas verticales representarían de este modo una primera aproximación a los «moldes» [«templates»] postulados en la teoría de Goldstein (véase más arriba). La respuesta será la más alta para aquellas neuronas «molde» cuya arquitectura de conexión sináptica se ajuste más al pattern de excitación de la señal incidente. Finalmente, se supone que la ubicación de la máxima actividad de salida de las neuronas verticales (a lo largo de la dimensión x , Fig. AII.1) conduce a la sensación de altura subjetiva o de periodicidad. Después del proceso de aprendizaje, esta salida es muy diferente de aquella de las fibras horizontales (en la dimensión y). Por ejemplo, si se activan suficientes sinapsis por los armónicos superiores de f_1 , la neurona K' responderá aun en el caso de que la fundamental f_1 esté ausente en el tono estímulo. Esto representa el mecanismo de seguimiento de la fundamental. Cuanto más alto sea el orden de los armónicos, tanto menos definida será la información «horizontal» de entrada, a causa de la proximidad de las zonas de máxima excitación (Fig. AII.1). Las neuronas verticales pueden ser llevadas entonces a responder a una señal de entrada «equivocada» (una que no corresponde a la frecuencia fundamental a la cual el árbol dendrítico apical de la neurona estuvo ligado originalmente). De aquí que sea posible la percepción de alturas múltiples, ambiguas, como lo hemos demostrado cuantitativamente en la primera parte de este Apéndice.

Sin embargo, nuestro modelo necesita algunos ajustes. Como se ve en la Fig. AII.1, la neurona sintonizada K' también respondería a todos los tonos compuestos que tengan por frecuencia fundamental un múltiplo

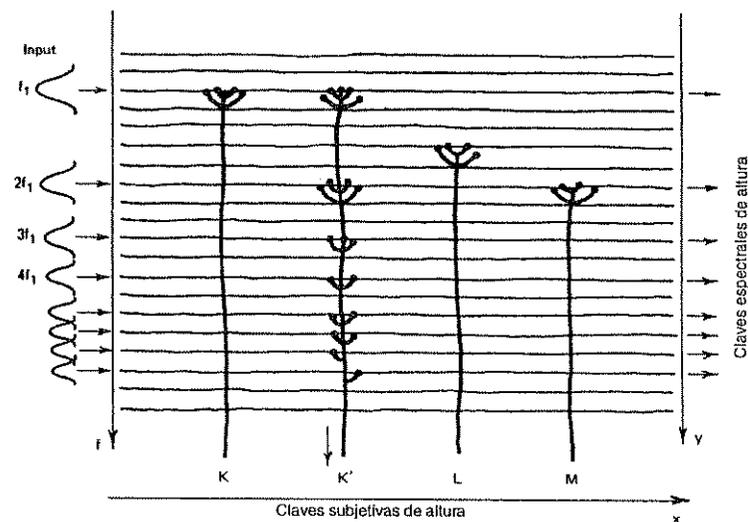


Figura AII.1 Modelo esquemático de un cableado neural para la extracción de altura y seguimiento de la fundamental. K , L , M : neuronas en una etapa inicial, «no entrenadas». K' : configuración sináptica de una neurona después de múltiples exposiciones a un tono compuesto de armónicos. Estos se muestran a lo largo del eje izquierdo (véase el texto).

entero de f_1 . Incluso podrá disparar en el caso de que esté presente solo uno de los armónicos superiores. Para evitar este efecto indeseable, podemos introducir un conjunto intermedio de neuronas verticales, capaz de detectar coincidencias entre armónicos vecinos. No es difícil escribir un simple programa de computación³ que simule estas operaciones. Algunos resultados se muestran en la Fig. AII.2 (para suposiciones simples pero realistas sobre la distribución de excitación primaria alrededor de cada armónico y sobre la degradación de respuesta para armónicos más altos). En la parte superior de la figura se dan los espectros de potencia de entrada primarios (en escalas lineales), correspondientes a superposiciones de tonos compuestos formando intervalos de octava, quinta y tercera menor respectivamente. Los gráficos inferiores representan la distribución computada de la actividad neuronal a lo largo de la dimensión x (compárese con Fig. AII.1). Nótese los pronunciados picos correspondientes a cada una de las frecuencias fundamentales de cada tono compuesto. Se supone

³ Este programa no hace suposiciones sobre transformadas de Fourier o funciones de autocorrelación. Simplemente desplaza un 'molde' [template] de frecuencias armónicas a lo largo de un dominio de frecuencia fundamental determinado, y cuenta, para cada posición, el número de excitaciones simultáneas en posiciones armónicas vecinas. El número total de pares de excitaciones simultáneas representa la salida para una frecuencia fundamental dada (la intensidad o probabilidad de activación de la neurona vertical correspondiente).

que estos picos son reconocidos en un nivel más alto del procesamiento neural, conduciendo a las dos sensaciones de altura asociadas a los dos tonos compuestos incidentes. La posición, y en gran medida la forma de estos picos primarios son independientes de los espectros de potencia de los tonos componentes, dependiendo *solo* de sus frecuencias fundamentales f_a y f_b . Obsérvese también que, a medida que disminuye el grado de consonancia, empiezan a aparecer picos «parásito» en posiciones que corresponden a la frecuencia de repetición r y sus múltiplos. Estos picos parásito (que deben ser inhibidos deliberadamente en alguna etapa superior) están *ausentes* para la octava. Además hay una actividad de fondo o «nivel de ruido» debajo de la frecuencia más baja que aumenta a medida que decrece el grado de consonancia⁴.

El modelo neural arriba tratado es muy primitivo, por cuanto su concepción dista mucho de reflejar la realidad fisiológica, salvo en aquello que hemos llamado «la esencia» del proceso de aprendizaje: el establecimiento de nuevas sinapsis o el cambio en la eficiencia en las sinapsis existentes entre neuronas (Sec. 4.10). En los últimos años hubo importantes avances en el desarrollo de modelos más realistas de redes neurales (por ej. véase Hinton, 1992, o para más detalles, Arbib, 1987; Kohonen, 1988), en parte promovidos por científicos en computación e ingenieros en robótica, interesados en el diseño de computadoras neurales. La computación neural está empezando a ser usada en el estudio de la percepción de altura y en otras tareas musicales cognitivas. Una discusión detallada excede el alcance de este libro y solo podemos sintetizar algunos conceptos fundamentales.

Los modelos de redes neurales están compuestos de unidades interconectadas que representan neuronas típicas. Cada conexión tiene un «peso» que puede variarse y que determina la eficiencia de la transmisión de la «frecuencia de disparo» de una neurona a otra (un peso negativo corresponde a una sinapsis inhibitoria). Cada unidad recibe la actividad de muchas otras, multiplica cada una con el factor de peso correspondiente a esa sinapsis, las suma, y transforma la totalidad de esta entrada en una actividad única de salida (véanse pág. 62 y Fig. 2.21). Los modelos neurales usuales tienen sus unidades distribuidas en varios estratos diferentes. El primer nivel recibe información de entrada del «mundo externo», representando, por ejemplo en la percepción de altura, la distribución de la actividad sobre la membrana basilar (véase eje vertical en la Fig. AII.1). El estrato de salida presenta distribuciones de actividad neural que son específicas a los patrones para cuyo reconocimiento la red ha sido diseñada (por ej. una «señal» de altura separada para cada uno de los tonos compuestos que forman el estímulo en la Fig. AII.2, o, en un

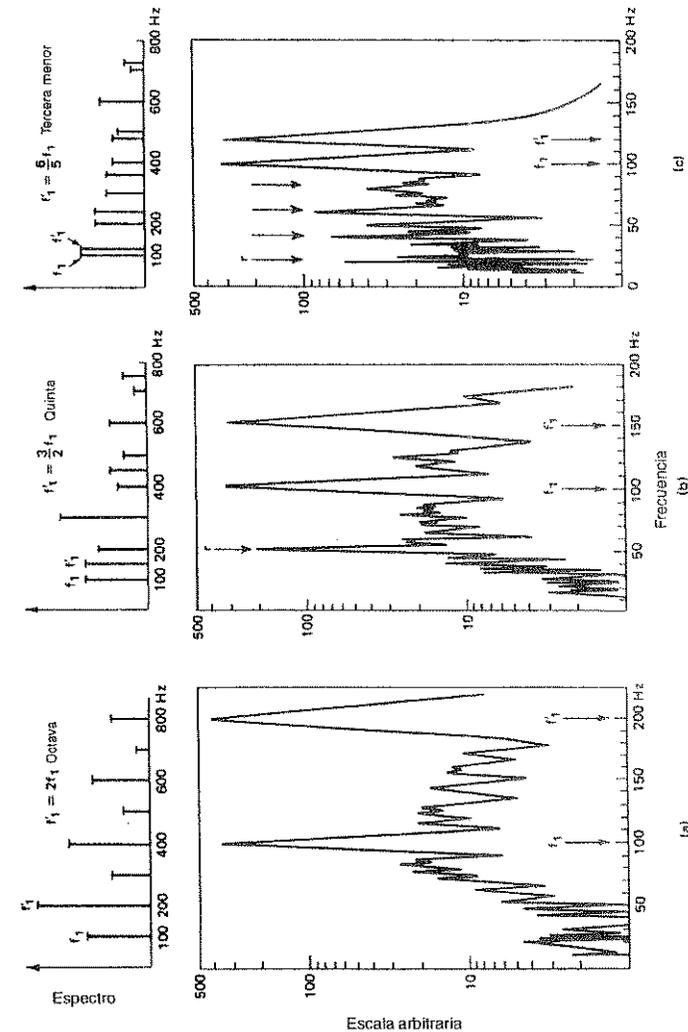


Figura AII.2. Gráficos superiores: espectros hipotéticos de la superposición de dos tonos compuestos formando una octava exacta, una quinta y una tercera menor. Gráficos inferiores: señales 'output' computadas usando un modelo simple de procesador de alturas (véase el texto). Nótese cuán pronunciados son los picos en las frecuencias fundamentales de los tonos componentes.

⁴ Un procedimiento mucho más sofisticado, basado en líneas generales en el esquema esbozado arriba, fue desarrollado por Terhardt et. al. (1982).

caso mucho más complejo, una «señal cognitiva» separada para cada instrumento musical en un sonido orquestal que llega al oído).

Inicialmente, el peso dado a cada sinapsis está preasignado (por ej., el mismo peso para todas, o un peso asignado aleatoriamente). Luego, estos pesos se modifican paulatinamente en sucesivos pasos de entrenamiento; en cada paso el output real se compara con la salida deseada, y los pesos se reajustan, según complejos protocolos numéricos, para producir una aproximación mejor. Este procedimiento se repite incontables veces hasta lograr un nivel de precisión predeterminado. Después de haber completado el «aprendizaje», se ve con frecuencia que el sistema es capaz de reconocer correctamente patrones incompletos (¡la fundamental ausente!) o ciertos patrones que no han sido presentados en absoluto al sistema durante el proceso de aprendizaje (Kohonen, 1988). Esta capacidad de interpretar situaciones nuevas (no aprendidas) es una propiedad fundamental de los modelos avanzados de redes neurales. Es esencial para la descripción de los modos de procesamiento paralelo, distributivo, holológico de la función cerebral (Sec. 4.10), que le permiten a un sistema que evolucionó para llevar a cabo un conjunto limitado de tareas, ejecutar operaciones nuevas que no han jugado ningún papel en la evolución genética, tales como leer y escribir (Kosslyn y Koenig, 1992), integrar ecuaciones diferenciales... ¿escuchar música? (véase Sec. 5.6).

¿Cómo se ajustan los modelos neurales actuales a la realidad neuronal? En primer lugar está la llamada retropropagación [«back-propagation»], responsable de efectuar los cambios sinápticos necesarios en los circuitos para adecuarlos a producir la salida deseada. Esto requeriría que la información de salida sea usada para realimentar la red neural en una dirección opuesta a la propagación de actividad que «baja» del estrato de entrada; ¡la estructura neural de la corteza y de las estructuras subcorticales presenta características que parecen ejecutar precisamente este proceso!⁵. En segundo lugar, estos modelos requieren un «maestro» que suministre la salida deseada, verifique cuán bueno es el ajuste, y ordene los cambios apropiados en los pesos sinápticos durante la retropropagación. Por otra parte, el modelo primitivo de la figura AII.1 es un ejemplo de una red que aprende «sin supervisión», en la cual la aparición repetida de rasgos

⁵ Estudios recientes muestran que las neuronas activadas liberan una sustancia química - óxido nítrico - (por ej. Shuman y Madison, 1994) que se propaga a neuronas vecinas y de algún modo sirve para reforzar aquellas sinapsis que estén recibiendo impulsos simultáneamente (es decir, aquellas que participan en un proceso cerebral específico relacionado con aquel en el que está involucrada la neurona que libera la sustancia). Esto representaría un mecanismo específico de potenciación a largo plazo (véase pág. 179). Por ejemplo, consideremos la Fig. AII.1: suponiendo que en un comienzo las neuronas «verticales» de salida estén conectadas débilmente a todas las neuronas de entrada que ellas interceptan, la liberación de una sustancia química activante por parte de la neurona *K* reforzaría todos los contactos sinápticos con aquellas fibras de entrada que disparen en el momento de ser estimuladas por el tono compuesto mostrado a la izquierda. Después de sucesivas exposiciones a ese tono, *K* surgiría como la neurona sintonizada *K*'.

específicos conduce automáticamente a los cambios sinápticos después de varias exposiciones, produciendo una salida que es respuesta específica a esos rasgos después del período de entrenamiento. Finalmente, en el sistema nervioso central hay redes 'precableadas' que procesan información sensorial, cuya arquitectura sináptica surgió en la evolución.

Mucho queda por hacer en el desarrollo de la computación neural. Por ejemplo, por ahora, los modelos de redes neurales, en su mayoría, están diseñados para un análisis de patrones espaciales (incluso los sistemas de reconocimiento de voz trabajan principalmente sobre la base del análisis de espectros sonoros instantáneos). Sin embargo, el cerebro debe manipular configuraciones secuenciales y trabajar tanto con distribuciones espaciales como temporales de actividad neuronal (véanse Secs. 2.8 y 4.10). Por lo tanto es necesario diseñar modelos en los cuales la actividad sea mantenida durante un cierto tiempo en circuitos cerrados (loops). Una cuestión mucho más difícil es la de comprender la capacidad de «aprendizaje instantáneo» de nuestro cerebro por la cual la información es almacenada en memoria *sin* la necesidad de exposiciones repetidas⁶, y tratar de emular este proceso con modelos de redes neurales. Finalmente, deben desarrollarse las herramientas matemáticas que permitan la comprensión cabal y la descripción cuantitativa de los algoritmos involucrados (por ej., Arbib, 1987). Cuando se cumpla todo esto, habremos completado una parte importante de... ¡meramente la etapa inicial en el desarrollo de la computación neural!

A causa de su potencial impacto en las neurociencias, ciencias del comportamiento, robótica, matemáticas y sus muchas aplicaciones en física, ecología, etc., el desarrollo de modelos de redes y computadoras neurales representa hoy en día uno de los grandes desafíos interdisciplinarios de la ciencia y la tecnología. ¿Se beneficiará la música con todo esto? Una mirada hacia el pasado nos enseña que desde los comienzos mismos de su desarrollo la música *siempre* abrazó los últimos avances de la ciencia y la tecnología. ¡Lo atestiguan las flautas, arpas y trompetas del antiguo Egipto, las maravillas mecánicas y acústicas de los órganos del siglo XV, el uso de las computadoras en la teoría y composición musicales, así como también los refinados instrumentos electroacústicos de hoy!

⁶ Esto involucra al *hipocampo* y estructuras subcorticales relacionadas que tienen un importante papel en las operaciones de almacenamiento y recuperación de información en la memoria.

Apéndice III

Algunas observaciones acerca de la enseñanza de la física y la psicofísica de la música *

No sería sensato hacer recomendaciones demasiado elaboradas acerca de cómo organizar un curso verdaderamente interdisciplinario sobre esta materia. La principal razón radica en la diversa preparación, impredecible, del estudiante típico que podría interesarse en tal curso, así como también en el amplio espectro de intereses que el docente deberá enfrentar. Suponiendo que estuviera abierto a estudiantes de todas las disciplinas universitarias, el curso podría incluir estudiantes de cinco áreas principales: 1) música; 2) psicología; 3) ciencias biológicas; 4) ciencias de la comunicación e ingeniería; y (5) física y matemáticas. La principal dificultad radica en lograr que el curso sea igualmente interesante, útil y fácilmente comprensible para todos. Esto requiere tres condiciones: 1) Minimizar el uso de las matemáticas, sin llegar a que la exposición resulte ridícula para los especialistas en ciencias e ingeniería (*Sugerencia: usar el curso para mostrar explícitamente a los expertos cómo enseñar ciencia sin usar matemática*). 2) Explicar cada tema «desde cero», sea de física, psicofísica o música, sin llegar a ser condescendiente con los expertos en cada campo (*Sugerencia: usar el curso para mostrar explícitamente a los expertos de cada campo como hacer reseñas concisas y completas sobre sus respectivas especialidades*). 3) Llevar a cabo demostraciones en clase, organizar trabajos prácticos de laboratorio así como también plantear problemas y situaciones de tal manera que el estudiante (independientemente de su área de especialización) pueda responder sin hesitar a la siguiente pregunta: ¿Qué aprendí viendo esa demostración o resolviendo tal o cual problema? (*Sugerencia: en los experimentos, evitar que los estudiantes tengan que realizar o presenciar tediosas mediciones; mostrarles cómo las magnitudes físicas se relacionan entre sí, cómo cambia una con respecto a la otra, y como están conectadas entre sí por relaciones*

* En este Apéndice se formulan recomendaciones para un curso universitario del tipo electivo, ofrecido a nivel del ciclo básico de universidades norteamericanas.

causa y efecto. En los problemas, evitar que los alumnos tengan que «resolver ecuaciones»; mostrar cómo una determinada relación matemática une dos o más cantidades «dinámicamente» a lo largo de todo un rango de variabilidad; dar una *visión intuitiva de las relaciones cuantitativas entre magnitudes*; mostrar cómo se pueden usar las matemáticas para predecir la conducta de un sistema. En las preguntas, hacer pensar al alumno intuitivamente sin dejar de exigir, sin embargo, precisión científica en las respuestas).

Una seria dificultad es que muchas de las personas especializadas en música (y en otras artes) tienen «miedo» acerca de todo lo que sea rigor científico, en el sentido de suponer de antemano que ellas «no podrán comprender». Esto no es más que un bloqueo mental que puede ser eliminado con paciencia, persuasión y dedicación por parte del docente.

La inclusión de temas de psicoacústica en un curso introductorio de acústica musical presenta una serie de desafíos adicionales para quien esté a cargo del mismo. El más obvio es desarrollar adecuadamente el contenido dentro del tiempo disponible. No importa cuanto tiempo se disponga: siempre habrá que tomar decisiones difíciles con respecto a qué temas dejar de lado y cuáles incluir. En segundo lugar, para un estudiante de acústica musical, la psicoacústica y la neuropsicología, son aún menos familiares que la física. Esto hace necesario restringir los temas en estas disciplinas a los más relevantes.

Será útil señalar ya desde el comienzo del curso algunos aspectos pertinentes de la psicoacústica. Por ejemplo, hacer notar que recientes conquistas en la comprensión de la percepción musical pueden conducir a abrir nuevas fronteras en el área de la composición musical. Señalar que muchos de los errores conceptuales que existen en torno a la ejecución musical tienen su raíz en los modos particulares en que la información acústica es procesada en el oído y en el cerebro. Indicar que muchos requisitos técnicos en los equipos electroacústicos de alta calidad están directamente relacionados con aspectos particulares de procesamiento de señales en el sistema nervioso. Hacer notar que el estudio de la percepción musical no es de interés solo para los músicos, sino también para neuropsicólogos y psicólogos, quienes pueden obtener de este campo información cuantitativa sobre funciones del cerebro y aplicaciones en musicoterapia.

Una dificultad general consiste en que las demostraciones experimentales y trabajos de laboratorio en general requieren equipos muy costosos. Sin embargo, es posible organizar un curso razonable con instrumentos prestados de otros cursos o departamentos. Describimos a continuación este mínimo equipamiento, con el cual pueden demostrarse la mayor parte de los experimentos mencionados en este libro.

1. Experimentación psicoacústica

1) Dos generadores de onda sinusoidal, un amplificador 'stereo' de buena calidad, buenos auriculares, dos parlantes de buena calidad. 2) Un osciloscopio para cada grupo de cuatro a seis estudiantes; de ser posible, de doble haz con memoria. 3) Un sintetizador electrónico (es suficiente una versión portátil¹). Con este equipo es perfectamente factible demostrar casi todo lo que se mencionó en las páginas 27, 29, 37, 41, 44, 46, 49, 53, 107, 168 - 169, 185 - 187 y 196. Si se cuenta con un órgano de tubos grande en un buen entorno acústico, se podrán hacer demostraciones útiles adicionales (véanse por ej., págs. 55, 106 - 107, 110, 162, 166 y 181). En todos estos experimentos o demostraciones, como regla, *todo lo que esté sonando debería a la vez presentarse en la pantalla de un osciloscopio.*

2. Experimentación acústica

1) Un «sonómetro»² por grupo de cuatro a seis estudiantes con estroboscopio y circuitería adecuada para llevar a cabo los experimentos descritos en las páginas 121 - 125. Con este equipo también se pueden estudiar experimentalmente las relaciones (4.2) y (4.3), así como los mecanismos de frotamiento y punteado de una cuerda. 2) Un piano es útil para realizar las simples demostraciones descritas en las páginas 124 y 126. 3) Tubos de órgano aislados (habitualmente disponibles en los departamentos de física) para explorar curvas de resonancia del tipo mostrado en la Fig. 4.24, y las relaciones (4.5) y (4.6), usando un pequeño parlante de buena calidad «implantado» en el tubo. 4) Sinfines y «ripple tanks» también habitualmente disponibles en laboratorios de física, son muy útiles para la demostración de ondas viajeras, estacionarias, y de acústica óptica en general. 5) Toda experimentación acústica debe estar acompañada de adecuadas clases de problemas. Un conjunto excelente de problemas se encuentra en Savage (1977).

Además, es aconsejable proponer individualmente a cada alumno un trabajo de seminario sobre un instrumento musical de su elección, lo cual, por supuesto, requerirá acceso a la apropiada literatura.

Resumiendo, éste es un curso que plantea al docente muchos desafíos y, a la vez, una enseñanza entretenida, tal vez el curso más interdisciplinario que pueda ofrecer una universidad a un nivel básico. Tanto al alumno como al docente le brinda la oportunidad de hacer volar la imaginación ¡dentro de los estrictos límites de la ciencia!

¹ Véase la excelente descripción de experimentos seleccionados y de demostraciones en clase dada por Hartmann (1975).

² Un nombre pomposo para una cuerda montada sobre una caja de resonancia, con tensión

Referencias

- Allen, J.B., and S.T. Neely. 1992. Micromechanical models of the cochlea. *Phys. Today*, July 1992, 40.
- Ando, Y. 1985. *Concert Hall Acoustics*. Springer-Verlag, Berlin.
- Arbib, M.A. 1987. *Brains, Machines, and Mathematics*. Springer-Verlag, New York.
- Askenfelt, A., and E. Jansson. 1990. In *The Acoustics of the Piano*. Publ. of the Royal Swedish Academy of Music 64:36.
- Backus, J. 1974. Input impedance curves for the reed woodwind instruments. *J. Acoust. Soc. Am.* 56:1266.
- Backus, J., and T.C. Hundley. 1971. Harmonic generation in the trumpet. *J. Acoust. Soc. Am.* 49:509.
- Benade, A.H. 1971. Physics of wind instrument tone and response. In *Symposium on Sound and Music, December 1971*. American Association for the Advancement of Science, Washington, D.C.
- Benade, A.H. 1973. The physics of brasses. *Sci. Am.* 229(1):24.
- Benade, A.H. 1976. *Fundamentals of Musical Acoustics*. Oxford University Press, London.
- Bever, T.G., and R.J. Chiarello. 1974. Cerebral dominance in musicians and nonmusicians. *Science* 185:537.
- Bharucha, J.J. 1994. Tonality and expectation. In *Musical Perceptions*. R. Aiello, ed. Oxford University Press, Oxford. 213.
- Bilsen, F.A., and J.L. Goldstein. 1974. Pitch of dichotically delayed noise and its possible spectral basis. *J. Acoust. Soc. Am.* 55:292.
- Borchgrevink. 1982. Prosody and musical rhythm are controlled by the speech hemisphere. In *Music, Mind, and Brain*. M. Clynes, ed. Plenum Press, New York. 151.
- Bradshaw, J.L., and N.C. Nettleton. 1981. The nature of hemispheric specialization in man. *Behavioral and Brain Sci.* 4:51.
- Bredberg, G., H.H. Lindemann, H.W. Ades, R. West, and H. Engstrom. 1970. Scanning electron microscopy of the organ of Corti. *Science* 170:861.
- Brodal, A. 1969. *Neurological Anatomy*. Oxford University Press, London.
- Corso, J.F. 1957. Absolute judgments of musical tonality. *J. Acoust. Soc. Am.* 29:138.

- Dallos, P. 1992. The active cochlea. *J. Neurosci.* **12**:4575.
- Damaske, P. 1971. Head-related two-channel stereophony with loud-speaker reproduction. *J. Acoust. Soc. Am.* **50**:1109.
- Davis, H. 1962. Advances in the neurophysiology and neuroanatomy of the cochlea. *J. Acoust. Soc. Am.* **34**:1377.
- deBoer, 1983. No sharpening? A challenge to cochlear mechanics. *J. Acoust. Soc. Am.* **73**:567.
- Denenberg, V.H. 1981. Hemispheric laterality in animals and the effects of early experience. *Brain and Behavioral Sci.* **4**:1.
- Deutsch, D. 1982. Organizational processes in music. In *Music, Mind, and Brain*. M. Clynes, ed. Plenum Press, New York. 119.
- Deutsch, D., ed. 1983. *The Psychology of Music*. Academic Press, New York.
- Egan, J.P., and H.W. Hake. 1950. On the masking pattern of a simple auditory stimulus. *J. Acoust. Soc. Am.* **22**:622.
- Flanagan, J.L. 1972. *Speech Analysis, Synthesis and Perception*, 2nd ed. Springer-Verlag, New York.
- Fletcher, H., and W.A. Munson. 1933. Loudness, its definition, measurement, and calculation. *J. Acoust. Soc. Am.* **5**:82.
- Friedlander, F.G. 1953. On the oscillations of a bowed string. *Cambridge Philos. Soc. Proc.* **49**:516.
- Gazzaniga, M.S. 1970. *The Bisected Brain*. Meredith Corporation.
- Gelfand, S.A. 1990. *Hearing*. Marcel Dekker, New York.
- Geschwind, N. 1972. Language and the brain. *Sci. Am.* **226**(4):76.
- Geschwind, N., and W. Levitzky. 1968. Human brain: Left-right asymmetries in the temporal speech region. *Science* **161**:186.
- Goldberg, J.M., and P.B. Brown. 1969. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanisms of sound localization. *J. Neurophysiol.* **32**:613.
- Goldstein, J.L. 1970. Aural combination tones. In *Frequency Analysis and Periodicity Detection in Hearing*. R. Plomp and G.F. Smoorenburg, eds. A.W. Suithoff, Leiden. 230.
- Goldstein, J.L. 1973. An optimum processor theory for the central formation of the pitch of complex tones. *J. Acoust. Soc. Am.* **54**:1496.
- Grey, J.M., and J.W. Gordon. 1978. Perceptual effects of spectral modifications of musical timbres. *J. Acoust. Soc. Am.* **63**:1493.
- Gur, R.C., I.K. Packer, J.P. Hungerbuhler, M. Reivich, W.D. Obrist, W.S. Amarnek, and H.A. Sackeim. 1980. Differences in distribution of gray and white matter in human cerebral hemispheres. *Science* **207**:1226.
- Hall, D., and A. Askenfelt. 1988. Piano string excitation V: Spectra for real hammers and strings. *J. Acoust. Soc. Am.* **83**:1627.
- Hartmann, W.M. 1975. The electronic music synthesizer and the physics of music. *Am. J. Phys.* **43**:755.
- Hartmann, W.M. 1993. On the origin of the enlarged melodic octave. *J. Acoust. Soc. Am.* **93**:3400.
- Herrington, R.N., and P. Schneidau. 1968. The effect of imagery on the wave-shape of the visual evoked response. *Experientia* **24**:1136.
- Hinton, G.E. 1992. How neural networks learn from experience. *Sci. Am.* **267**(3):145.
- Houtsma, A.J.M. 1970. Perception of musical pitch. *J. Acoust. Soc. Am.* **48**:88(A).
- Houtsma, A.J.M., and J.L. Goldstein. 1972. Perception of musical intervals: Evidence for the central origin of the pitch of complex tones. *J. Acoust. Soc. Am.* **51**:520.
- Hubel, H. 1971. The visual cortex of the brain. In *Contemporary Psychology. Readings from Scientific American*. W-H. Freeman, San Francisco.
- Hudspeth, A.J. 1985. The cellular basis of hearing: The biophysics of hair cells. *Science* **230**:745.
- Hudspeth, A.J. 1989. How the ear's works work. *Nature* **341**:397.
- Hutchins, C.M., and F.L. Fielding. 1968. Acoustical measurements of violins. *Phys. Today* **21**(7):34.
- Iverson, P., and C.L. Krumhansl. 1993. Isolating the dynamic attributes of musical timbre. *J. Acoust. Soc. Am.* **94**:2595.
- Jackendoff, R., and F. Lerdahl. 1982. A grammatical parallel between music and language. In *Music, Mind, and Brain*. M. Clynes, ed. Plenum Press, New York. 83.
- Jansson, E., N.-E. Molin, and H. Sundin. 1970. Resonances of a violin body studied by hologram interferometry and acoustical methods. *Phys. Scr.* **2**:243.
- Johnstone, B.M., R. Patuzzi, and P. Sellick. 1983. Comparison of basilar membrane, hair cell and neural responses. In *Hearing-Psychological Bases and Psychophysics*, R. Klinke and R. Hartmann, eds. Springer-Verlag, Berlin. 46.
- Johnstone, B.M., R. Patuzzi, and G.K. Yates. 1986. Basilar membrane measurements and the traveling wave. *Hearing Res.* **22**:147.
- Kachar, B., W.E. Brownell, W.E. Altschuler, and J. Fex. 1986. Electrokinetic shape changes of cochlear outer hair cells. *Nature* **322**:365.
- Kameoka, A., and M. Kuriyagawa. 1969. Consonance theory Part II: Consonance of complex tones and its calculation method. *J. Acoust. Soc. Am.* **45**:1460.
- Keller, J.B. 1953. Bowing of violin strings. *Comm. Pure Appl. Math.* **6**:483.
- Kemp, D.T. 1978. Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am.* **64**:1386.
- Kiang, N.Y.-S., T. Watanabe, E.C. Thomas, and L.F. Clark. 1965. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. MIT Press, Cambridge, Massachusetts.
- Kimura, D. 1963. Right temporal lobe damage. *Arch. Neurol.* **8**:264.
- Klein, W., R. Plomp, and L.C.W. Pols. 1970. Vowel spectra, vowel spaces, and vowel identification. *J. Acoust. Soc. Am.* **48**:999.
- Kohonen, T. 1988. *Self-Organization and Associative Memory*. Springer-Verlag, Berlin.
- Kosslyn, S.M., and O. Koenig. 1992. *Wet Mind: The New Cognitive Neuroscience*. The Free Press.
- Lerdahl, F., and R. Jackendoff. 1983. *A Generative Theory of Tonal Music*. MIT Press, Cambridge.
- Lieberman, M.C. 1978. Auditory-nerve response from cats raised in a low noise chamber. *J. Acoust. Soc. Am.* **63**:442.
- Licklider, J.C.R. 1959. Three auditory theories. In *Psychology: A Study of a Science*. Vol. I. S. Koch, ed. McGraw-Hill, New York.
- Lipps, T. 1905. *Psychologische Studien*. Durr'sche Buchhandlung. Leipzig.
- Matthews, M.V., and J. Kohut. 1973. Electronic stimulation of violin resonances. *J. Acoust. Soc. Am.* **53**:1620.

- Meyer, M. 1900. Elements of a psychological theory of melody. *Psych. Rev.* 7:241.
- Milner, B. 1967. Brain mechanisms suggested by studies of temporal lobes. In *Brain Mechanisms Underlying Speech and Language*. C.H. Millikan and F.L. Darley, eds. Grune and Stratton, New York.
- Milner, B., L. Taylor, and R.W. Sperry. 1968. Lateralized suppression of dichotically presented digits after commissural section in man. *Science* 161:184.
- Molino, J.A. 1973. Pure-tone equal-loudness contours for standard tones of different frequencies. *Percept. Psychophys.* 14:1.
- Molino, J.A. 1974. Psychophysical verification of predicted interaural differences in localizing distant sound sources. *J. Acoust. Soc. Am.* 55:139.
- Moore, B.C.J. 1973. Frequency difference limens for short-duration tones. *J. Acoust. Soc. Am.* 54:610.
- Panksepp, J. 1982. Toward a general psychobiological theory of emotions. *Behavioral and Brain Sci.* 5:407.
- Papçun, G., S. Krashen, D. Terbeek, R. Remington, and R. Harshman. 1974. Is the left hemisphere specialized for speech, language, and/or something else? *J. Acoust. Soc. Am.* 55:319.
- Patterson, B. 1974. Musical dynamics. *Sci. Am.* 231(5):78.
- Penfield, W., and L. Roberts. 1959. *Speech and Brain Mechanisms*. Princeton University Press, Princeton, New Jersey.
- Pierce, J.R. 1983. *The Science of Musical Sound*. Scientific American Books, W.H. Freeman, San Francisco.
- Plomp, R. 1964. The ear as a frequency analyzer. *J. Acoust. Soc. Am.* 36:1628.
- Plomp, R. 1965. Detectability threshold for combination tones. *J. Acoust. Soc. Am.* 37:1110.
- Plomp, R. 1967. Beats of mistuned consonances. *J. Acoust. Soc. Am.* 42:462.
- Plomp, R. 1967a. Pitch of complex tones. *J. Acoust. Soc. Am.* 41:1526.
- Plomp, R. 1970. Timbre as a multidimensional attribute of complex tones. In *Frequency Analysis and Periodicity Detection in Hearing*. R. Plomp and F.G. Smoorenburg, eds. A.W. Suithoff, Leiden. 397.
- Plomp, R. 1976. *Aspects of Tone Sensations*. Academic Press, New York.
- Plomp, R., and H.J.M. Steeneken. 1973. Place dependence of timbre in reverberant sound fields. *Acustica* 28:50.
- Plomp, R., and M.A. Bouman. 1959. Relation between hearing threshold and duration for tone pulses. *J. Acoust. Soc. Am.* 31:749.
- Plomp, R., and W.J.M. Levelt. 1965. Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38:548.
- Pribram, K.H. 1971. *Languages of the Brain*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Raiford, C.A., and E.D. Schubert. 1971. Recognition of phase changes in octave complexes. *J. Acoust. Soc. Am.* 50:559.
- Rakowski, A. 1971. Pitch discrimination at the threshold of hearing. *Proc. 7th Int. Congr. Acoust. Budapest.* 3:373.
- Rakowski, A. 1972. Direct comparison of absolute and relative pitch. *Proc. Symp. Hearing Theory*, IPO, Eindhoven.
- Ratliff, F. 1972. Contour and contrast. *Sci. Am.* 226(6):91.
- Reinicke, W., and L. Cremer. 1970. Application of holographic interferometry to

- vibrations of the bodies of string instruments. *J. Acoust. Soc. Am.* 48:988.
- Rhode, W.S., and L. Robles. 1974. Evidence from Mössbauer experiments for nonlinear vibration in the cochlea. *J. Acoust. Soc. Am.* 55:588.
- Richards, A.M. 1977. Loudness perception for short-duration tones in masking noise. *J. Speech Hearing Res.* 20:684.
- Ritsma, R.J. 1967. Frequencies dominant in the perception of the pitch of complex sounds. *J. Acoust. Soc. Am.* 42:191.
- Roederer, J.G. 1978. On the relationship between human brain functions and the foundations of physics, science, and technology. *Found. of Phys.* 8:423.
- Roederer, J.G. 1979. Human brain functions and the foundations of science. *Endeavour* 3:92.
- Roederer, J.G. 1984. The search for a survival value of music. *Music Perception* 1:350.
- Roederer, J.G. 1987. Neuropsychological processes relevant to the perception of music: An introduction. In *Music in Medicine*. R. Spintge and R. Droh, eds. Springer-Verlag, Berlin. 81.
- Rose, J.E., J.F. Brugge, D.J. Anderson, and J.E. Hind. 1969. Some possible neural correlates of combination tones. *J. Neurophys.* 32:402.
- Ruggero, M.A., and N. Rich. 1991. Application of a commercially manufactured Doppler shift laser velocimeter to the measurement of basilar membrane motion. *Hearing Res.* 51:215.
- Sachs, M.B., and P.J. Abbas. 1974. Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli. *J. Acoust. Soc. Am.* 56:1835.
- Saunders, F.A. 1946. The mechanical action of instruments of the violin family. *J. Acoust. Soc. Am.* 17:169.
- Savage, W.R. 1977. *Problems for Musical Acoustics*. Oxford University Press, New York.
- Scharf, B. 1983. Loudness adaptation. In *Hearing Research and Theory*, Vol. 2. J.V. Tobias and E.D. Schubert, eds. Academic Press, New York.
- Scheid, P., and J.C. Eccles. 1975. Music and speech: Artistic functions of the human brain. In *Psychology of Music* 3:21.
- Schelleng, J.C. 1973. The bowed string and the player. *J. Acoust. Soc. Am.* 53:26.
- Schuman, E.M., and D.V. Madison. 1994. Locally distributed synaptic potentiation in the hippocampus. *Science* 263:532.
- Siebert, W.M. 1970. Frequency discrimination in the auditory system: Place or periodicity mechanisms? *Proc. IEEE* 58:723.
- Simmons, F.B. 1970. Monaural processing. In *Foundations of Modern Auditory Theory*. J.V. Tobias, ed. Academic Press, New York. 343.
- Small, A.M. 1970. Periodicity pitch. In *Foundations of modern Auditory Theory*. J.V. Tobias, ed. Academic Press, New York. 1.
- Smoorenburg, G.F. 1970. Pitch perception of two-frequency stimuli. *J. Acoust. Soc. Am.* 48:924.
- Smoorenburg, G.F. 1972. Audibility region of combination tones. *J. Acoust. Soc. Am.* 52:603.
- Sokolich, W.G., and J.J. Zwislocki. 1974. Evidence for phase oppositions between inner and outer hair cells. *J. Acoust. Soc. Am.* 55:466.
- Sommerhoff, G. 1974. *Logic of the Living Brain*. John Wiley, New York.
- Stevens, S.S. 1955. Measurement of loudness. *J. Acoust. Soc. Am.* 27:815.

- Stevens, S.S. 1970. Neural events and the psychophysical law. *Science* 170:1043.
- Stevens, S.S., J. Volkman, and E.B. Newman. 1937. A scale for the measurement of psychological magnitude pitch. *J. Acoust. Soc. Am.* 8:185.
- Sundberg, J. 1991. *The Science of Musical Sound*. Academic Press, San Diego.
- Sundberg, J., ed. 1992. *Gluing Tones*. Royal Swedish Academy of Music (with compact disc).
- Terhardt, E. 1971. Pitch shifts of harmonics, an explanation of the octave enlargement phenomenon. *Proc. 7th Int. Congr. Acoust.* Budapest. 3:621.
- Terhardt, E. 1972. Zur Tonhöhenwahrnehmung von Klängen, I, II. *Acustica* 26:173, 187.
- Terhardt, E. 1974. Pitch, consonance and harmony. *J. Acoust. Soc. Am.* 55:1061.
- Terhardt, E., and H. Fastl. 1971. Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen. *Acustica* 25:53.
- Terhardt, E., and M. Zick. 1975. Evaluation of the tempered tone scale in normal, stretched, and contracted intonation. *Acustica* 32:268.
- Terhardt, E., G. Stoll and M. Seewann. 1982. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *J. Acoust. Soc. Am.* 71:679.
- van Noorden, L.A.P.S. 1975. *Temporal Coherence in the Perception of Tone Sequences*. Institute for Perception Research, Eindhoven (with a phonographic demonstration record).
- von Békésy, G. 1960. *Experiments in Hearing*. McGraw-Hill Book Company, New York.
- von Helmholtz, H. 1863. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. English translation, 1954. Dover Publications, New York.
- Walliser, K. 1969. Über die Abhängigkeiten der Tonhöhenempfindung von Sinustönen von Schallpegel, von überlagertem drosselnden Störschall und von der Darbietungsdauer. *Acustica* 21:211.
- Ward, W.D. 1970. Musical perception. In *Foundations of Modern Auditory Theory*. J.V. Tobias, ed. Academic Press, New York. 405.
- Whitfield, I.C. 1967. *The Auditory Pathway*. Edward Arnold Ltd., London.
- Wightman, F.L. 1973. The pattern-transformation model of pitch. *J. Acoust. Soc. Am.* 54:407.
- Yost, W.A., and C.S. Watson, eds. 1987. *Auditory Processing of Complex Sounds*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Zwicker, E., and H. Fastl. 1990. *Psychoacoustics*. Springer-Verlag, Berlin.
- Zwicker, E., and B. Scharf. 1965. A model of loudness summation. *Psych. Rev.* 72:3.
- Zwicker, E., G. Flottorp, and S.S. Stevens. 1957. Critical bandwidth in loudness summation. *J. Acoust. Soc. Am.* 29:548.
- Zwislocki, J.J. 1965. Analysis of some auditory characteristics. In *Handbook of Mathematical Psychology*. R.D. Luce, R.R. Bush, and E. Galanter, eds. Wiley, New York.
- Zwislocki, J.J. 1969. Temporal summation of loudness: An analysis. *J. Acoust. Soc. Am.* 46:431.
- Zwislocki, J.J., and W.G. Sokolich. 1973. Velocity and displacement responses in auditory-nerve fibers. *Science* 182:64.

Indice alfabético

A

- Absorción (de sonido), 161
- Acústica de salas, 159
- Adaptación, 109
- Alargamiento de intervalos, 167, 188, 196
- Altura, 11
- ambiguas (múltiple), 59, 166
 - de periodicidad. Véase subjetiva.
 - de tonos compuestos (Véase altura subjetiva), 58, 72, 164, 165
 - de tonos simples o puros, 31, 35, 67, 117
 - primaria. Véase de tonos simples o puros.
 - espectral, 34, 174, 219 (Véase también de tonos simples o puros)
- Altura de periodicidad (Véase altura), 51
- Altura espectral (Véase altura), 167, 219
- Altura subjetiva, 55, 56, 58, 73, 165, 218
- Altura virtual (Véase altura), 55
- Amplitud, 29
- Análisis de Fourier, 135 (Véase también espectro)
- Apareamiento de altura [Pitch matching], 46, 58
- Aprendizaje (como proceso neural), 73, 169, 170, 179, 188, 220
- Armónicos, 57, 120, 123
- Armónicos aurales, 49
- Armónicos superiores. Véase armónicos.

- Atenuación de la sonoridad, 109
- Audición
- Teoría auditiva de localización, 37
 - Umbral de la, 100
- Axones, 62

B

- Banda crítica, 41, 44, 46, 105, 137, 186
- Batidos, 40
- de primer orden, 50, 59, 183
 - de segundo orden, 50, 52, 53, 183
- Batidos subjetivos, 50

C

- Caída exponencial, 131
- Calidad (de un tono). Véase timbre.
- Células pilosas, 33
- motilidad, 65, 114
- Cent (unidad de intervalo musical), 195
- Ciclo, 26
- Ciclos por segundo. Véase Hertz.
- Cóclea, 10, 32, 110
- Codificación (en el sistema auditivo), 60
- Codificación auditiva. Véase codificación.
- Cognición, 176, 201
- Columna de aire
- abierta, 144
 - cerrada, 152
- Computadoras neurales, 222

Consonancia, 13, 183, 222
 Corteza, 10, 14, 21, 75, 203
 Corteza auditiva, 77
 Corteza cerebral. *Véase* corteza.
 Cromá, 188
 Cuerdas (ondas en), 121, 128
 Cuerpo calloso, 78
 Cuerpo del violín (vibración del), 142
 Curva de resonancia, 142, 152

D

DAP. *Véase* diferencia apenas perceptible.
 db. *Véase* decibel.
 Decibel, 100
 Dendritas, 61
 Detección de rasgos, 174
 Diferencia apenas perceptible, 18, 36
 de sonoridad, 99
 de altura, 36
 Difracción, 162
 Direccionalidad. *Véase* localización.
 Disonancia, 13, 46, 183
 Dispersión, 125
 Distribución (de actividad neural), 21,
 62, 67, 220
 Dominio (de un tono en una secuencia).
Véase tónica.
 Duración (efecto sobre la sonoridad).
Véase atenuación de la sonoridad.

E

Efecto «cóctel party», 183
 Efecto Chorus, 182
 Efecto de precedencia, 160, 163, 183
 Efecto Mössbauer, 35, 116
 Elemento vibrante, 10, 135
 Emisiones otoacústicas, 113
 Endolinfa, 32
 Energía
 cinética, 81
 potencial, 82
 Enmascaramiento, 102, 106
 Enseñanza, 226
 Escalas, musicales
 justa, 190, 191

pitagórica, 190, 192
 Escala justa. *Véase* escalas.
 Escala pitagórica. *Véase* escalas.
 Escala standard, 195
 Escala temperada, 194, 195, 198
 Especialización hemisférica, 207, 208
 Espectro (de un tono), 12, 135, 137,
 138, 142, 152, 165, 187
 Espectro de potencia, 136, 156, 168,
 169
 Experimentos (para laboratorios de
 enseñanza), 227, 228

F

Fase, 29, 51, 67, 128, 136, 165
 Fibras nerviosas. *Véase* neurona.
 Física
 clásica, 16
 cuántica, 16
 Fisión melódica, 200
 Fon (unidad de nivel de sonoridad), 103
 Formantes, 144
 Fraseo, 110
 Frecuencia
 análisis. *Véase* análisis de Fourier.
 discriminación, 40
 resolución, 36
 Frecuencia característica, 65
 Frecuencia central, 42
 Frecuencia de batido, 40
 Frecuencia de disparo (neural), 63, 111
 Frecuencia de repetición, 12, 54, 55, 56
 (*Véase* también frecuencia funda-
 mental)
 Frecuencia de resonancia, 89, 141
 (*Véase* también curva de resonancia)
 Frecuencia fundamental, 12, 54, 56,
 122
 Fricción, 82
 dinámica, 134
 estática, 133
 Fuerza, 80
 Fundamental ausente, 55 (*Véase* tam-
 bién altura subjetiva)
 Función cerebral, 171

H

Habilidad musical, 211
 Hemisferio dominante, 210
 Hemisferio menor, 207
 Hemisferios cerebrales, 23
 Hertz (Hz, unidad de frecuencia), 30
 Hz. *Véase* Hertz.

I

IL. *Véase* nivel de intensidad.
 Imaginación (en el cerebro), 177
 Imaginación acústica. *Véase* imaginación.
 Impedancia de entrada, 152
 Impulsos (en neuronas), 10, 20, 30, 62,
 63
 Impulsos nerviosos. *Véase* impulsos.
 Información (neural). *Véase* codificación.
 Instrumento tipo clarinete, 153
 Instrumentos de bronce, 156, 157
 Instrumentos tipo flauta, 31, 145, 152
 Instrumentos tipo oboe, 153
 Intensidad, 12, 93
 Intervalos agrandados. *Véase* alarga-
 miento de intervalos.
 Intervalos musicales, 43, 46, 184, 196,
 200

J

J. *Véase* Joule.
 Joule (J, unidad de energía), 82

L

L. *Véase* sonoridad, subjetiva.
 Lengüetas, 150
 LL. *Véase* nivel de sonoridad.
 Localización (del sonido), 69
 Longitud de onda, 87

M

Magnitud psicofísica, 18
 multidimensional, 137
 Mecanismo de excitación, 10, 120
 Mecanismo de frotamiento, 83, 133
 Mejor frecuencia. *Véase* frecuencia ca-

racterística.
 Melodía, 14
 Membrana basilar, 32
 Membrana tectorial, 33
 Memoria, 170
 Memoria asociativa, 175
 ML. *Véase* nivel de enmascaramiento.
 Modos (de vibración), 124
 Moldes [Templates], 72, 165, 216
 Motivación, 19, 175
 Movimiento
 armónico simple, 28
 periódico, 26
 sinusoidal, 28
 Música, 21
 Musicalidad. *Véase* habilidad musical.

N

N. *Véase* Newton.
 Neurona, 21, 61, 111, 219
 Neuropsicología, 19
 Newton (N, unidad de fuerza), 80
 Nivel de enmascaramiento, 102, 107
 Nivel de intensidad (IL), 100
 Nivel de presión sonora (SPL), 101
 Nivel de sonoridad (LL), 103
 Nodos, 95
 Nota pedal, 110

O

Octava, 50, 183, 188
 Oído absoluto, 198
 Oído interno. *Véase* cóclea.
 Ondas
 elásticas, 79
 intensidad de las, 93
 longitudinales, 84, 90
 estacionarias, 94, 121, 144, 147
 superposición de, 94, 124, 135
 transversales, 85, 87
 Ondas longitudinales. *Véase* ondas.
 Ondas sonoras. *Véase* ondas.
 Organo, 55, 107, 155, 163, 164, 166
 Organo de Corti, 32
 Oscilación amortiguada, 82, 130, 131

Oscilaciones armónicas. *Véase* movimiento armónico simple.
Osciloscopio, 27, 228

P

Pa. *Véase* Pascal.
Paquete olivo-coclear, 77, 117
Pascal (Pa, unidad de presión), 81
Patrón de excitación (en la membrana basilar), 37, 61, 71, 73 (*Véase* también reconocimiento de patrones)
Patrón vibratorio, 26, 53, 59, 60, 61 (*Véase* también modos de vibración)
Pensamiento (como proceso neural), 173
Percepción estereofónica (del sonido). *Véase* localización.
Perilinfia, 32
Período, 26
Piano, experimentos con el, 124, 126
Potencia, 83
Potencial de acción, 62
Potencial postsináptico, 62
Presión, 81
 nodos de, 98, 144
 vientres de, 98, 146
Presión de arco, 134
Procesador central de altura, 72, 165, 181, 187, 200, 219
Propagación (del sonido), 79, 163
Psicoacústica, 17, 19
Psicofísica, 15, 16

R

Reconocimiento (de un instrumento musical), 144, 169
Reconocimiento de patrones [Pattern Recognition], 169, 171, 187
Red eferente, 65, 77, 116
Redes neurales (modelos), 176, 219
Reflexión (del sonido), 92, 162
Registro (grave, medio, agudo), 154, 155

Representación holológica (en el cerebro), 178
Resonador, 10, 120, 130, 138
Retropropagación, 224
Reverberación, 10, 160
Ritmo, 205
Rugosidad (de dos tonos puros), 41, 184
Ruido, 13, 190

S

Seguimiento de la fundamental [Fundamental Tracking], 53 (*Véase* también altura subjetiva)
Sentido de retorno, 14, 204 (*Véase* también tónica)
Señal nerviosa. *Véase* impulsos.
Sinapsis, 62, 220
Síntesis (de tonos), 135, 163
Síntesis de Fourier, 135
Síntesis sonora. *Véase* síntesis.
Sintonización nerviosa, 115
Sistema límbico, 20, 202
Sistema nervioso central, 211
Sistema neural auditivo (sistema nervioso auditivo), 10, 61, 69
Sistemas de alta fidelidad, 50
Sobresoplo, 155
Son (unidad subjetiva de sonoridad), 105
Sonido diferencial, 47, 49
Sonido directo, 159, 160, 163, 183
Sonidos resultantes, 47, 219
Sonoridad, 11, 31, 98, 101, 103, 109, 167
 atenuación (en tonos cortos), 109
 compresión, 105
 subjetiva (L), 83, 99
 suma de, 107, 112
Sonoridad subjetiva, 103
SPL. *Véase* nivel de presión sonora.
Superposición
 de ondas. *Véase* ondas.
 de tonos compuestos, 180
 de tonos simples o puros, 37, 120

T

Temperatura (efectos sobre el sonido), 86, 146
Teoría auditiva de localización [place theory of hearing], 37 (*Véase* también altura)
Tiempo medio de caída, 131
Timbre, 10, 12, 120, 126, 129, 134, 167
Tímpano, 10, 81
Tonalidad, 14
Tónica, 14, 184, 199
Tonos simples o puros, 31
 superposición de, 37, 120
Tono de bisel, 149
Tono de lengüeta, 151
Tono residual. *Véase* fundamental ausente.
Tonos armónicos, 71
Toque (en la ejecución pianística), 132
Trabajo, 81, 82
Tríada, 190, 191
Tubo (abierto, cerrado). *Véase* columna de aire.
Tubos de órgano, 10, 146, 155

U

Umbral de enmascaramiento, 102
Umbral de la audición, 98, 100

V

Velocidad (de las ondas), 86
Vías auditivas, 76
Vibración, 26
Vibrato, 171
Vientres, 95, 97, 124, 144

W

W. *Véase* Watt.
Watt (W, unidad de potencia), 83

Z

Zona de resonancia (sobre la membrana basilar), 34, 64, 89