

USP

Reitor
Vice-reitor

Jacques Marcovitch
Adolpho José Melfi

edusp

EDITORA DA UNIVERSIDADE DE SÃO PAULO

Presidente
Diretor Editorial
Editores-assistentes

Sergio Miceli Pessoa de Barros
Plinio Martins Filho
Heitor Ferraz
Rodrigo Lacerda

Comissão Editorial

Sergio Miceli Pessoa de Barros (Presidente)
Davi Arrigucci Jr.
Hugo Aguirre Armelin
Oswaldo Paulo Forattini
Tupã Gomes Corrêa

INTRODUÇÃO À FÍSICA E PSICOFÍSICA DA MÚSICA

JUAN G. ROEDERER

TRADUÇÃO
ALBERTO LUIS DA CUNHA

edusp

Originalmente publicado em inglês, sob o título:
Introduction to the Physics and Psychophysics of Music
Copyright © 1975 by Springer-Verlag, New York
Todos os direitos reservados

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Roederer, Juan G., 1929 -

Introdução à Física e Psicofísica da Música / Juan G. Roederer; tradução Alberto Luis da Cunha. - São Paulo: Editora da Universidade de São Paulo, 1998.

Título original: Introduction to the physics and psychophysics of music.
ISBN 85-314-0457-6

1. Música - 2. Acústica e Física I. Título.

98-2026

CDD-781.22

Índice para catálogo sistemático:

1. Tempo musical :Teoria musical 781.22

Direitos reservados à

Edusp - Editora da Universidade de São Paulo
Av. Prof. Luciano Gualberto, Travessa J, 374
6º andar - Ed. da Antiga Reitoria - Cidade Universitária
05508-900 - São Paulo - SP - Brasil Fax (011) 818-4151
Tel. (011) 818-4008/818-4150 - e-mail: edusp@edu.usp.br

Printed in Brazil 1998

Foi feito o depósito legal

*Agradeço à minha esposa, Beatriz,
pela cuidadosa revisão dos manuscritos e pela
sua compreensão e tolerância do tempo "extracurricular"
que eu gastei escrevendo e reescrevendo este livro!*

SUMÁRIO

13	Introdução
17	1. Música, Física, Psicofísica e Neuropsicologia
17	1.1 Os Sistemas Físicos Intervenientes
21	1.2 Atributos Característicos dos Sons Musicais
23	1.3 O Elemento Tempo em Música
25	1.4 Física e Psicofísica
33	1.5 O que é Música
35	2. Vibrações Sonoras, Sons Puros e a Percepção da Altura
36	2.1 Movimento e Vibração
41	2.2 Movimento Harmônico Simples
43	2.3 Vibrações Acústicas e Sensações de Sons Puros
53	2.4 Superposição de Sons Puros: Batimentos de Primeira Ordem e a Banda Crítica
63	2.5 Outros Efeitos de Primeira Ordem: Sons Resultantes e Harmônicos Aurais

68	2.6	Efeitos de Segunda Ordem: Batimentos de Consonâncias Desafinadas
72	2.7	Rastreamento de Fundamentais
81	2.8	Codificação Auditiva no Sistema Nervoso Periférico
89	2.9	A Altura Subjetiva e o Papel do Sistema Nervoso Central
103	3.	Ondas Sonoras, Energia Acústica e a Percepção do Volume
103	3.1	Ondas Elásticas, Força, Energia e Potência
109	3.2	Velocidade de Propagação, Comprimento de Onda e Potência Acústica
123	3.3	Superposição de Ondas; Ondas Estacionárias
128	3.4	Intensidade, Nível Sonoro e Volume
143	3.5	O Mecanismo de Percepção do Volume e Processos Correlatos
148	3.6	Música dos Ouvidos: Emissões Otoacústicas e Mecânica Coclear
157	4.	Geração de Sons Musicais, Sons Complexos e a Percepção do Timbre
158	4.1	Ondas Estacionárias numa Corda
164	4.2	Geração de Vibrações Estacionárias Complexas em Instrumentos de Cordas
176	4.3	Espectros de Vibração Sonora e Ressonância
188	4.4	Ondas Estacionárias Longitudinais numa Coluna de Ar Idealizada
194	4.5	Geração de Vibrações Estacionárias Complexas em Instrumentos de Sopro
203	4.6	Espectros de Sons de Instrumentos de Sopro
206	4.7	Confinamento e Absorção de Ondas Sonoras num Ambiente Fechado
213	4.8	A Percepção de Altura e de Timbre de Sons Musicais
220	4.9	Identificação de Sons Musicais
223	4.10	Processos Cognitivos Cerebrais relativos à Percepção de Sons Individuais

235	5.	Superposição e Seqüências de Sons Complexos e a Percepção da Música
235	5.1	Superposição de Sons Complexos
239	5.2	A Sensação da Consonância e da Dissonância Musicais
248	5.3	Construindo Escalas Musicais
254	5.4	A Escala-padrão e o Padrão de Altura
259	5.5	Por que Existem Escalas Musicais
262	5.6	Processos Cerebrais Cognitivos e Afetivos na Percepção Musical: Por que Reagimos Emocionalmente à Música?
268	5.7	A Especialização da Fala e o Processamento da Música nos Hemisférios Cerebrais
275	Apêndice I	– Alguns Aspectos Quantitativos, do Mecanismo de Movimento do Arco
281	Apêndice II	– Alguns Aspectos Quantitativos de Modelos de Processador de Altura Central
293	Apêndice III	– Algumas Observações sobre o Ensino da Física e Psicofísica da Música
295		Experimentação Psicoacústica
296		Experimentação Acústica
299	Bibliografia	
307	Índice Remissivo	

INTRODUÇÃO

Este livro trata dos sistemas físicos e processos psicofísicos que ocorrem no que nós em geral chamamos de “música”. Vamos analisar quais propriedades físicas, objetivas, dos padrões sonoros estão associadas a quais sensações psicológicas, subjetivas, da música. Vamos descrever como esses padrões sonoros são realmente produzidos nos instrumentos, como eles se propagam pelo ambiente e como são detectados pelo ouvido e interpretados no cérebro. Faremos tudo isso utilizando a linguagem dos físicos e seus métodos de raciocínio e análise — sem usar, entretanto, matemáticas complicadas (isso, obviamente, imporá sérias limitações em nossa apresentação). Apesar de não ser exigido nenhum conhecimento prévio de física, fisiologia e neurobiologia, supõe-se, no entanto, que o leitor esteja familiarizado com música, em particular com a notação musical, escalas musicais e intervalos, que possua pelo menos algum conhecimento básico sobre instrumentos musicais e tenha experimentado “sensações” musicais típicas.

Há livros facilmente encontráveis sobre os fundamentos da física da música (p. ex., Benade, 1976; Pierce, 1983) e da psicologia da música

(p. ex., Deutsch, 1982). O propósito deste volume não é duplicar, mas sintetizar e complementar a literatura existente. Na verdade, minha meta principal ao escrever este livro foi tecer uma rede consistente entre as disciplinas da física, psicofísica e neuropsicologia, e produzir um texto interdisciplinar de um só autor sobre a chamada “ciência da música”. Também espero que ele transmita ao leitor um pouco do que chamo de “música da ciência”, ou seja, a beleza e a excitação da pesquisa científica!

Após a primeira edição de 1973, várias reimpressões e uma edição revista foram publicadas, bem como traduções para o alemão e o japonês; no momento, o livro também está sendo traduzido para o português. Estes são indicadores muito gratificantes, especialmente levando-se em conta o fato de que o assunto em questão é mais um *hobby* para mim do que uma ocupação de tempo integral! A terceira edição inclui informação sobre algumas descobertas fundamentais e novos estudos feitos durante os últimos 15 anos sobre a fisiologia da audição e da função cerebral. As seções físico-acústicas, por outro lado, permaneceram mais ou menos inalteradas desde a edição de 1979.

Uma das partes mais dolorosas ao se escrever um livro é decidir quais tópicos devem ser deixados de lado, ou desprezados, à vista das rigorosas limitações de espaço. Não importa o que o autor faça, haverá sempre alguém reclamando desta ou daquela omissão. Vamos enumerar aqui alguns dos assuntos que foram desprezados ou omitidos — sem que queiramos esboçar uma tentativa de justificção. Na discussão sobre geração de notas musicais, são analisados principalmente os mecanismos básicos, em detrimento da apresentação de situações musicais concretas. A voz humana não foi abordada, o que também ocorreu com discussões sobre tons inarmônicos (instrumentos de percussão) e geração eletrônica de sons; a música gerada por computador não é sequer mencionada. No lado psicoacústico, apenas a percepção de sons senoidais simples ou múltiplos é analisada, sem que se mencionem faixas de ruído ou experiências com estímulos de pulso.

E não há praticamente nada sobre ritmo, quase nada sobre percepção estereofônica e muito pouco sobre desenvolvimento histórico. Finalmente, nas referências bibliográficas, foi dada prioridade à citação de artigos em

fontes de disponibilidade mais ampla ao público em geral; referências detalhadas de artigos originais podem ser encontradas na maioria das fontes citadas.

A primeira edição deste livro foi o resultado de um roteiro publicado pela Universidade de Denver para os estudantes do curso de física da música, introduzido na universidade em 1970. Além dos trabalhos regulares de classe, foi proposto a esses estudantes que realizassem uma série de experimentos acústicos e psicoacústicos num laboratório modesto. A realização desses experimentos, dos quais alguns serão descritos, é essencial para uma compreensão clara dos principais conceitos físicos e psicoacústicos envolvidos. Infelizmente para o leitor não-especializado, eles com freqüência exigem um equipamento eletrônico não facilmente disponível. Peço aos leitores que acreditem na descrição dos experimentos e acreditem que eles realmente resultem no que eu disser. Sempre que possível, indicarei como pode ser realizado certo experimento com o auxílio de equipamento musical comum.

MÚSICA, FÍSICA, PSICOFÍSICA E NEUROPSICOLOGIA

UMA ABORDAGEM INTERDISCIPLINAR

1.1 Os SISTEMAS FÍSICOS INTERVENIENTES

Imagine que você esteja numa sala de concertos ouvindo a apresentação de um músico. Vamos identificar os sistemas que são relevantes para a “música” que você ouve. Primeiramente, é claro, temos o músico e o instrumento que “faz” a música. Em segundo lugar, temos o ar da sala, que transmite o som em todas as direções. Em terceiro lugar, há você, o ouvinte. Em outras palavras, temos a seguinte cadeia de sistemas: instrumento-ar-ouvinte. O que os conecta enquanto a música está sendo tocada? Um certo tipo e forma de vibração chamada som, que se propaga de um ponto a outro na forma de ondas, e à qual nosso ouvido é sensível. (Há muitos outros tipos e formas de vibração que não podemos detectar, ou que podemos fazê-lo, mas com outros sentidos, como o tato ou a visão.)

O físico usa termos mais gerais para descrever os três sistemas acima. Ele os chama de: fonte-meio-receptor. Essa cadeia de sistemas é comum

no estudo de muitos outros processos físicos: luz, radioatividade, eletricidade, gravidade, raios cósmicos etc. A fonte emite, o meio transmite e o receptor detecta, registra, ou, em geral, é afetado de alguma forma específica. O que é emitido, transmitido ou detectado é a energia — em uma de suas múltiplas formas, dependendo do caso particular considerado. No caso das ondas sonoras, trata-se da energia elástica, porque ela envolve oscilações de pressão, *i.e.*, compressões e expansões de ar que se alternam rapidamente¹.

Vamos agora proceder a uma segunda e mais detida análise dos sistemas envolvidos. Na fonte, *i.e.*, no instrumento musical, identificamos vários componentes distintos:

1. O mecanismo de excitação primária que deve ser ativado pelo músico², tal como a ação do arco ou do dedo sobre uma corda de violino, a palheta oscilante numa clarineta, os lábios do músico num instrumento de sopro de metal, ou o ar soprado na fenda de uma flauta. Esse mecanismo de excitação age como fonte de energia primária.
2. O elemento vibrante fundamental que, quando excitado pelo mecanismo primário, é capaz de sustentar certos modos bem-definidos de vibração com frequências prefixadas, como as cordas de um violino ou a coluna de ar dentro do tubo de um instrumento de sopro ou órgão. Esse elemento vibrante, na verdade, determina a altura musical da nota e, como feliz dividendo, fornece os harmônicos superiores necessários para conferir certa qualidade característica ou timbre ao som. Além disso, ele serve como um armazenador de energia vibracional. Em instrumentos de sopro, ele controla parcialmente o mecanismo de excitação primária por meio de um acoplamento de realimentação (forte nas madeiras, fraco nos metais).

1. Naturalmente, o som também se propaga em líquidos e sólidos.
 2. Para tornar a descrição completa, devemos acrescentar os componentes do músico: o córtex motor de seu cérebro, que emite os comandos aos seus músculos, as partes do seu corpo com as quais ele opera o instrumento musical ou o seu aparelho vocal, a realimentação dos ouvidos e dos músculos que o auxiliam a controlar sua *performance* etc. Mas, infelizmente, a limitação de espaço nos obriga a deixar o músico totalmente fora desse quadro.

3. Muitos instrumentos têm um ressonador adicional (a caixa de um piano ou o corpo de um instrumento de cordas), cuja função é converter mais eficientemente as oscilações do elemento vibrante primário (corda) em vibrações sonoras do ar circundante e dar ao som o seu timbre final.

Quanto ao meio, também devemos fazer uma distinção: há o meio propriamente dito, que transmite o som, e as fronteiras, *i.e.*, as paredes, o teto, o chão, as pessoas do público etc., que afetam fortemente a propagação sonora por reflexão e absorção das ondas sonoras, e cuja configuração determina a qualidade acústica da sala (reverberação).

Finalmente, podemos selecionar no ouvinte os seguintes componentes principais:

1. O tímpano, que capta as oscilações de pressão da onda sonora que atinge o ouvido e as converte em vibrações mecânicas que são transmitidas por meio da ligação de três pequenos ossos para:
2. O ouvido interno, ou cóclea, no qual as vibrações são classificadas de acordo com gamas de frequência, captadas por células receptoras, e convertidas em impulsos nervosos elétricos.
3. O sistema nervoso auditivo, que transmite os sinais neurais ao cérebro, onde a informação é processada, apresentada como uma imagem de detalhes auditivos em certa área do córtex (a superfície do cérebro e o tecido subjacente), identificada, armazenada na memória e eventualmente transferida para outros centros do cérebro. Esses últimos estágios levam à percepção consciente dos sons musicais.

Observe que podemos substituir o ouvinte por um dispositivo de gravação, como um gravador em fita magnética, um gravador em CD, ou uma gravação fotoelétrica em filme, e ainda reconhecer pelo menos três dos subsistemas: a detecção mecânica e subsequente conversão em sinais elétricos no microfone, uma quantidade limitada de processamento deliberado ou acidental no circuito eletrônico e a armazenagem em memória na fita, disco ou filme, respectivamente. O primeiro sistema, *i.e.*, o instru-

mento, também pode, naturalmente, ser substituído por um dispositivo eletrônico de execução musical.

Podemos resumir essa discussão na Tabela 1.1.

Tabela 1.1

	Sistema	Função
Fonte	Mecanismo de excitação	Suprimento de energia
	Elemento vibrante	Determinação das características do som fundamental
	Ressonador	Conversão em oscilações de pressão do ar (ondas sonoras), determinação final das características do som
Meio	Meio propriamente dito	Propagação sonora
	Fronteiras	Reflexão, absorção, reverberação
Receptor	Tímpano	Conversão em oscilações mecânicas
	Ouvido interno	Seleção de frequências primárias, conversão em impulsos nervosos
	Sistema nervoso	Processamento, configuração, identificação, armazenagem e transferência para outros centros cerebrais

O principal objetivo deste livro é analisar abrangentemente o que ocorre em cada estágio visto na Tabela 1.1 e durante cada transição de um estágio para outro, quando a música está sendo tocada.

1.2 ATRIBUTOS CARACTERÍSTICOS DOS SONS MUSICAIS

Elementos de todas as culturas confirmam que há três sensações primárias associadas a um determinado som musical: altura, intensidade e qualidade³. Não vamos tentar definir esses atributos subjetivos ou magnitudes psicológicas, ou, nesse estágio da discussão, preocuparmo-nos com o fato de eles serem ou não quantidades mensuráveis; vamos apenas observar que a altura é normalmente relacionada a graves e agudos, e a intensidade ao volume ou “força” de um som. A qualidade, ou timbre, é o que nos permite distinguir os sons de diferentes instrumentos, mesmo que eles tenham a mesma altura e a mesma intensidade. A associação não-ambígua dessas três qualidades para um determinado som é o que diferencia uma “nota” musical de um “ruído”: embora possamos seguramente atribuir intensidade a um certo ruído, é muito mais difícil identificar-lhe uma única altura ou timbre.

A atribuição de altura, intensidade e qualidade a um som musical é o resultado de operações de processamento no ouvido e no cérebro. É subjetivo e inacessível a uma medição física direta (veja a Seção 1.4). Entretanto, cada uma dessas sensações primárias pode ser associada, em princípio, a uma quantidade física bem definida do estímulo original, *i.e.*, a onda sonora, que pode ser medida e expressa numericamente por métodos físicos. Na verdade, a sensação de altura está primariamente associada à frequência fundamental (frequência de repetição de um padrão de vibração, descrito pelo número de oscilações por segundo), o volume à intensidade (fluxo de energia ou amplitude de oscilação de pressão da onda sonora que atinge o ouvido), e a qualidade ao espectro, ou seja, a proporção em que outras frequências superiores, chamadas

3. As sensações às vezes citadas de volume e densidade (ou brilho) são conceitos compostos que nunca podem ser resolvidos numa combinação de efeitos de altura e volume (uma diminuição da altura com um simultâneo aumento de volume dá uma sensação de volume aumentado; o aumento de altura com simultâneo aumento de volume dá uma sensação de densidade ou brilho aumentado).

“harmônicos superiores”, aparecem misturadas entre si, acompanhando a frequência fundamental.

Este, entretanto, é um quadro muito simplista. Em primeiro lugar, a sensação de volume de um som com intensidade constante aparentemente variará se mudarmos a frequência. Em segundo, o volume de uma superposição de várias notas de diferentes frequências (p. ex., um acorde) já não está mais relacionado de uma forma simples ao fluxo total de energia sonora; por outro lado, para uma sucessão de notas de duração muito curta (p. ex., uma frase *staccato*), a sensação dependerá de quanto realmente durar cada nota. Em terceiro, a percepção refinada de timbre necessária para a identificação de instrumentos musicais é um processo que utiliza muito mais informação do que apenas o espectro da nota: as características transientes de ataque e decaimento são igualmente importantes, como pode ser facilmente verificado quando se tenta identificar instrumentos musicais enquanto se faz a fita magnética, onde esses sons estão gravados, girar ao contrário. Além disso, as notas de um certo instrumento podem apresentar características espectrais que mudam consideravelmente ao longo da tessitura do instrumento, e a composição espectral de uma certa nota pode mudar bastante de um ponto a outro da sala de música; não obstante, elas são identificadas sem hesitação como pertencentes ao mesmo instrumento. Da mesma forma, um músico bastante treinado teria a maior dificuldade em reproduzir a altura exata de uma nota gerada eletronicamente e desprovida de harmônicos superiores, ouvindo-a com fones de ouvido, porque falta ao seu sistema nervoso central importantes informações adicionais, que normalmente existem nos sons “reais” com os quais ele está familiarizado.

Outra característica física relevante de uma nota é a direção espacial de onde chega a onda sonora correspondente. O que importa aqui é a pequena diferença de tempo entre os sinais acústicos detectados por cada ouvido, o que depende da direção de incidência. Essa diferença de tempo é medida e codificada pelo sistema nervoso para fornecer a sensação da direcionabilidade do som (estereofonia ou fusão binaural).

Quando duas ou mais notas são ouvidas simultaneamente, o nosso cérebro é capaz de separá-las individualmente, dentro de certas limita-

ções. Sensações novas, não tão bem definidas, mas, ainda assim, musicalmente importantes, surgem em conexão com duas ou mais notas superpostas, levando coletivamente ao conceito de harmonia. Entre elas estão as sensações “estáticas” de consonância e dissonância, que descrevem o caráter “agradável” ou “irritante” de certas superposições de notas, a sensação “dinâmica” da necessidade de resolver certo acorde ou intervalo dissonante, o efeito peculiar dos batimentos, e o caráter diferente dos acordes maiores e menores. Enquanto a correlação de altura, intensidade e, até certo ponto, qualidade, com certas características físicas de notas individuais, é universal — *i.e.*, independente do condicionamento cultural do indivíduo —, este não é necessariamente o caso dos atributos subjetivos das superposições de notas que mencionamos acima.

1.3 O ELEMENTO TEMPO EM MÚSICA

Um som estacionário, com frequência, intensidade e espectro constantes, é perturbador. Além disso, após certo tempo, nosso presente consciente passa a não mais registrá-lo. Apenas quando esse som desaparecer é que subitamente nos damos conta de que ele estava sendo ouvido. A música é feita de notas cujas características físicas variam de certo modo com o tempo. É apenas essa dependência temporal que torna um som realmente “musical”. Em geral, daqui para a frente, usaremos o termo mensagem musical para designar uma seqüência ou superposição de notas individuais. Essa mensagem musical poderá ser significativa (às vezes chamada de *Gestalt tonal*), se atribuirmos a ela certo valor como resultado de uma série de operações cerebrais de análise, comparação com mensagens previamente armazenadas, armazenagem na memória e associações. Uma melodia é o exemplo mais simples de mensagem musical. Alguns atributos de mensagens musicais significativas são elementos-chave na música ocidental: a tonalidade (predomínio de uma certa nota na seqüência), a sensação de retorno à tônica, modulação e ritmo. Uma característica fundamental de uma melodia é que a sucessão de

notas procede em passos de altura finita e discreta em praticamente todas as culturas musicais. Do número infinito de frequências disponíveis, nosso sistema auditivo prefere selecionar valores discretos correspondentes às notas de uma escala musical, mesmo que sejamos capazes de detectar variações de frequência muito menores que o passo natural de qualquer escala musical. O mecanismo neural que analisa uma mensagem musical presta atenção apenas nas transições de altura; o processamento absoluto de alturas (“ouvido absoluto”) é perdido em tenra idade na maioria dos indivíduos.

Vamos examinar mais detalhadamente o elemento tempo em música. Há três escalas distintas de tempo nas quais ocorrem variações temporais de relevância psicoacústica. Primeiro, temos a escala “microscópica” de tempo, na qual as vibrações reais de uma onda sonora ocorrem, cobrindo uma gama de períodos de mais ou menos 0,00007 a 0,05 s. Há então uma faixa intermediária centrada em torno de um décimo de segundo, onde ocorrem algumas variações transientes, tais como ataque e decaimento, representando as variações temporais dos detalhes microscópicos. Finalmente, temos a escala “macroscópica” de tempo, indo de cerca de 0,1 s em diante, correspondendo às durações normais de notas, sucessões e ritmo. É importante observar que cada escala temporal típica tem um “centro processador” particular com uma função específica no sistema auditivo.

1. As vibrações microscópicas são detectadas e codificadas no ouvido interno e levam principalmente às sensações primárias das notas (altura, intensidade e qualidade).
2. As vibrações intermediárias ou transientes parecem afetar principalmente mecanismos de processamento no caminho neural entre o ouvido e a área auditiva do cérebro e fornecem indicações adicionais para a percepção da qualidade, identificação e discriminação de notas.
3. As variações macroscópicas de tempo são processadas no nível neural mais elevado — o córtex cerebral, *i.e.*, a estrutura superficial envoltória e o tecido subjacente; eles determinam a mensagem musical real e seus atributos.

Quanto mais subimos nesses estágios processadores no conduto auditivo, mais difícil torna-se definir e identificar os atributos psicológicos aos quais esse processamento leva, e mais as coisas parecem estar influenciadas por aprendizagem e condicionamento cultural, bem como pelo comportamento momentâneo do indivíduo.

Por mais de 100 anos os musicólogos têm se queixado de que a física da música e a psicoacústica restringiram-se principalmente ao estudo de sons estacionários e constantes, ou agregados de sons com essas características, enquanto a essência da música é uma seqüência temporal disso. As suas queixas são justificadas, mas as razões para tal restrição também têm suas justificativas. Como foi explicado acima, o processamento de seqüências de sons ocorre no nível mais alto do sistema nervoso central, envolvendo uma cadeia de mecanismos ainda pouco explorada. Neste livro, sempre que possível, vamos tentar remediar essa situação e ir até onde pudermos para tentar disseminar um pouco dessa crítica bem posicionada.

1.4 FÍSICA E PSICOFÍSICA

Podemos descrever o principal objetivo da física da seguinte maneira: fornecendo métodos por meio dos quais seja possível prever quantitativamente a evolução de um dado sistema físico (ou rever a sua história), baseado nas condições em que esse sistema é encontrado num certo instante⁴. Por exemplo: dado um automóvel de certa massa e especificando-se as condições de freagem, a física nos permite prever quanto tempo o carro levará para parar e onde ele parará, desde que especifiquemos a posição e a velocidade no instante inicial. Dados o comprimento, a massa e a tensão de uma corda de violino, a física prevê as frequências possíveis

4. Às vezes, se afirma que o objeto da Física é “conseguir uma explicação quantitativa do universo”. Todavia, isso, na melhor das hipóteses, é um jargão filosófico que pouco tem a ver com o âmbito real da Física (bem mais preciso, realista e modesto).

em que essa corda vibrará se for puxada ou friccionada de certa maneira. Dadas a forma e as dimensões de um tubo de órgão, e a composição e temperatura do gás lá dentro (ar), a física prevê a frequência fundamental do som emitido quando ele for soprado.

Na prática, “prever” significa fornecer um aparato matemático, uma série de equações ou “receitas” que, baseadas em certas leis físicas que governam o sistema em análise, estabelecem relações matemáticas entre os valores das magnitudes físicas que caracterizam o sistema num certo instante (posição e velocidade no caso do carro; frequência e amplitude de oscilação nos outros dois exemplos). Essas relações são então usadas para se descobrir a maneira como esses valores variam com o tempo.

Para estabelecer as leis físicas que governam um dado sistema, precisamos primeiro observar o sistema e fazer medições quantitativas das magnitudes físicas relevantes para descobrirmos experimentalmente as suas inter-relações causais. Uma lei física expressa certa relação que é comum a muitos sistemas físicos diferentes e independente de circunstâncias particulares. Por exemplo, a lei que governa a gravitação é válida aqui na Terra, na Lua, no Sistema Solar e em qualquer lugar do Universo. As leis de movimento de Newton aplicam-se a todos os corpos, independentemente de sua composição química, cor, temperatura, velocidade ou posição.

A maioria dos sistemas reais estudados na Física — mesmo os exemplos “simples” e “familiares” dados acima — é tão complexa, que as previsões detalhadas e precisas são impossíveis. Assim, precisamos fazer aproximações e imaginar modelos simplificados que representem um certo sistema apenas em seus aspectos principais. A “massa puntiforme” a que um corpo normalmente é reduzido na Física — seja ele um planeta, seja um carro, seja um elétron — é o mais simplificado de todos os modelos! Muitas vezes é necessário partir o sistema em estudo em uma série de subsistemas mais elementares, interagindo fisicamente uns com os outros, cada qual sendo governado por um conjunto bem definido de leis físicas.

A física do mundo cotidiano, ou física clássica, assume que tanto as medições quanto as previsões devem ser sempre “exatas” e “únicas”, e

as únicas limitações e erros são devidos à imperfeição de nossos dispositivos de medição. No domínio atômico e subatômico, entretanto, essa visão não se aplica mais. A natureza é tal, que nunca é possível esperar que as medições e previsões de um sistema atômico sejam exatas ou únicas no sentido comum, não importando o quanto tentemos aperfeiçoar nossas técnicas: as medições sempre terão precisão limitada, e apenas probabilidades podem ser previstas para os valores das magnitudes físicas no domínio atômico. Em outras palavras é impossível prever, digamos, quando um núcleo radioativo irá decair, ou exatamente onde certo elétron se encontrará num certo instante de sua trajetória do catodo até a tela da TV — só se pode especificar probabilidades. Uma física totalmente nova, adequada a descrever sistemas atômicos e subatômicos — a chamada *física quântica* —, teve que ser construída no começo da década de 20.

O leitor a essa altura pode estar se perguntando por que diabos estamos falando de física quântica, quando esta parece totalmente irrelevante para o estudo do som e da música. Entretanto, a psicofísica opera de uma forma tal que em muitos aspectos é surpreendentemente semelhante à física quântica. Em primeiro lugar, “como a física em geral, a psicofísica tenta fazer previsões sobre a evolução de um sistema específico sujeito a certas condições iniciais. O sistema em consideração é o sistema sensorial de um indivíduo (ou de um animal) — órgão receptor e partes afins do sistema nervoso. As condições são determinadas pelos *estímulos físicos de entrada*, e a resposta é expressa pelas *sensações psicológicas* despertadas no cérebro e relatadas pelo indivíduo (ou manifestadas pelo comportamento sensorial específico do animal). A *psicoacústica*, um ramo da psicofísica, é o estudo que relaciona os estímulos acústicos com as sensações auditivas. Assim como a física, a psicofísica exige que a relação causal entre o estímulo físico de entrada e a saída psicológica (ou comportamental) seja estabelecida por experimentação e medição. Assim como a física, a psicofísica deve fazer suposições simplificadoras e construir *modelos* para estar apta a estabelecer relações quantitativas e se aventurar no negócio de fazer previsões. Nos primeiros tempos da psicofísica, as relações empíricas de entrada e saída foram condensadas em leis

psicofísicas, tratando o *hardware* interveniente como uma “caixa preta” (equivalente à “massa puntiforme” da física!). Hoje, os modelos psicofísicos levam em conta as funções fisiológicas do órgão sensorial e das partes pertinentes do sistema nervoso.

Ao contrário da física clássica, mas muito semelhante à física quântica, nunca se pode esperar que as previsões psicofísicas sejam exatas ou únicas — apenas valores de probabilidade podem ser estabelecidos. Ao contrário da física clássica, mas muito semelhante à física quântica, a maioria das medições em psicofísica perturbará substancialmente o sistema em observação (p. ex., um indivíduo relatando as sensações causadas por um dado estímulo físico), e nada poderá ser feito para eliminar completamente essa perturbação. Em conseqüência, o resultado de uma medição não reflete o estado do “sistema perse”, mas o estado mais complexo do “sistema em observação”. Também como resultado, a psicofísica exige experimentação com diversos sistemas (sujeitos) equivalentes (mas nunca idênticos) e uma interpretação estatística dos resultados⁵.

Obviamente, existem alguns limites para essas analogias. Na física, o processo ou “receita” da medição que define uma certa magnitude física, como comprimento, massa ou velocidade de um objeto, pode ser formulado de uma forma rigorosa e não-ambígua. Quando trabalhamos com uma saída fisiológica, como rapidez de impulso neural, amplitude dos arrepios provocados, ou aumento do número de batimentos cardíacos, as medições também podem ser formuladas de uma forma rigorosa e não-ambígua. Mas na psicoacústica, como podemos definir e medir as sensações subjetivas de graves e agudos, volume, ou — para tornar a questão mais delicada — a magnitude que representa a necessidade de conduzir uma certa melodia para a sua conclusão na tônica? Como podemos realizar medições na “audição interna”, *i.e.*, a ação de provocar sensações de sons musicais por volição, sem estímulos externos? Poderia isso ser feito apenas por meio de perguntas, ou precisaríamos recorrer a medições “diretas”, via implantação de microeletrodos nas células do cérebro?

5. Devemos enfatizar que estas são apenas analogias. A física quântica como tal não desempenha um papel explícito no sistema nervoso, cuja operação está baseada em processos eminentemente clássicos.

Muitas sensações podem ser classificadas em tipos mais ou menos bem definidos (as chamadas qualidades sensoriais, se forem causadas pelo mesmo órgão de sentido) — e podemos citar como exemplo o fato de que as pessoas se relacionam umas com as outras em termos de altura, volume, qualidade tonal, consonância etc., sem muitos mal-entendidos mútuos acerca do significado desses conceitos. Além disso, duas sensações pertencentes ao mesmo tipo, experimentadas uma após a outra, podem em geral ser ordenadas pelo indivíduo que as experimentou quanto ao atributo específico de uma ser “maior”, “igual” ou “menor” que a outra (ou “mais alta”, “mais forte”, “mais brilhante”, “mais pronunciada” etc.). Por exemplo, quando lhe são fornecidas duas notas em seqüência, o sujeito pode julgar se a segunda nota teve uma altura “maior”, “igual” ou “menor” que a primeira. Outro exemplo de ordenamento é o seguinte: dispondo de três sons complexos de mesma altura e intensidade, ele deve ordená-los em pares, julgando quais os dois sons que têm o timbre mais semelhante e qual o mais diferente.

A habilidade, possuída por todos os indivíduos, de classificar e ordenar sensações subjetivas dá a elas um *status* quase equivalente ao de uma magnitude física e justifica a introdução do termo magnitude psicofísica. O que podemos *a priori* é esperar que as pessoas possam julgar se uma sensação tem, digamos, em volume, “duas vezes” ou “a metade”, ou qualquer outro fator numérico, de uma sensação unitária de referência. Nosso cérebro simplesmente não está equipado *a priori* com o dispositivo operacional necessário para uma intercomparação ou medição quantitativa. É como pedir, em medições atômicas, uma precisão que exceda o limite imposto pela mãe natureza, *i.e.*, a física quântica. Entretanto, existem situações em que é possível aprender a fazer estimativas quantitativas de magnitudes psicofísicas com uma base estatística, e, em alguns casos, o cérebro pode até se tornar muito bom nisso. O sentido visual é um exemplo. Com suficiente experiência, a avaliação do tamanho dos objetos pode tornar-se bastante precisa, desde que suficiente informação esteja disponível sobre o objeto; julgamentos como “o dobro do comprimento” ou “metade da altura” são feitos sem hesitação. Fica bem claro, por esse exemplo, que uma “unidade” e um processo psicofísico corres-

pondente de comparação foram formados no cérebro apenas por meio de experiência e aprendizagem, em múltiplos contatos com as magnitudes físicas originais. O mesmo pode ser conseguido para outras sensações psicofísicas como o volume: é necessário adquirir, por aprendizagem, a habilidade de comparação e julgamento quantitativo. O fato de os músicos de todo o mundo usarem uma notação comum para volume — talvez sem nunca terem ouvido falar de sonômetros e decibéis — é um exemplo auto-evidente.

E aqui chegamos talvez às diferenças mais cruciais entre física e psicofísica:

1. Repetidas medições de mesma espécie podem condicionar a resposta do sistema psicofísico em observação: o cérebro tem a habilidade de aprender gradualmente, mudando a probabilidade de resposta a um certo estímulo de entrada, à medida que aumenta o número de exposições semelhantes.
2. A motivação do sujeito em estudo e as conseqüências, mentais ou físicas, disso podem interferir nas medições de uma forma altamente imprevisível.

Como conseqüência do primeiro ponto, um estudo psicofísico estatístico com um único indivíduo exposto a repetidas “medições” nunca será idêntico a um estudo estatístico envolvendo uma única medição tomada com vários indivíduos diferentes. Isso é devido não só às diferenças entre os indivíduos, mas também ao condicionamento que ocorre no caso de exposições repetidas. Os sistemas ultracomplexos de realimentação do sistema nervoso tornam particularmente delicado estabelecer e interpretar medições psicoacústicas. As conseqüências do ponto (2) podem ter um alcance muito maior, especialmente em pesquisas sobre o papel do seu consciente.

A psicofísica pode ser considerada parte de uma disciplina mais abrangente. A psicoacústica, por exemplo, só menciona a questão “Por que ouvimos o que ouvimos?”, quando estamos expostos a certo estímulo acústico — mas ela não lida com *o que é* que estamos ouvindo! Em outras palavras,

ela não considera os processos de mais alto nível: cognição, comportamento e resposta emocional. A *neuropsicologia* é a disciplina que estuda os *processos e funções* do sistema neural que ligam a entrada recebida do ambiente e do corpo com toda a saída mental e comportamental⁶. Em outras palavras, a neuropsicologia busca fornecer um fundamento sistemático, biológico, quantitativo da experiência e do comportamento mentais. Como a física, a neuropsicologia trabalha com modelos. Eles são principalmente modelos de inter-relações funcionais; as partes neuroanatômicas e os processos fisiológicos intervenientes, naturalmente, devem ser levados em conta realisticamente. É importante ressaltar que as definições de psicofísica, fisiologia sensorial, neuropsicologia e psicologia, e as fronteiras entre elas, são difíceis de serem formuladas em termos precisos e concisos — a palavra “neuropsicologia” nem é encontrada nos dicionários ainda!

O principal sistema em estudo na neuropsicologia é o cérebro⁷. Em resumo, as mais importantes funções globais de um cérebro animal são a representação e previsão ambiental, e o planejamento da resposta comportamental, com o objetivo de maximizar as chances de sobrevivência e perpetuação da espécie. Para conseguir isso, o cérebro deve, a longo prazo, adquirir a informação sensorial necessária para fazer “plantas-baixas” das vizinhanças espaciais e descobrir relações de causa e efeito na ocorrência de eventos temporais e, a curto prazo, avaliar o atual estado do ambiente e do corpo, identificar aspectos ou alterações relevantes, fazer previsões a curto prazo baseadas na experiência (informação aprendida) e no instinto (informação genética) e planejar e executar uma resposta comportamental que provavelmente seja benéfica para o organismo. A condução e motivação gerais para executar essas tarefas são comandadas por uma parte filogeneticamente antiga do cérebro denominada *sistema límbico*, que distribui sinais que compõem o *estado afetivo* do organismo (prazer, medo, raiva etc.).

6. Repare que ela é um tanto diferente da *psicologia tradicional*, que é uma ciência basicamente descritiva que trata do comportamento *per se* e dos conceitos mais intangíveis, como sentimentos, motivação, valores morais, consciência e alma, e demais aspectos clínicos relacionados.

7. Com poucas exceções, não trataremos da anatomia e neurofisiologia do cérebro; existem muitos livros disponíveis nas bibliotecas médicas (p. ex., Brodal, 1969; Sommerhoff, 1974).

O cérebro humano pode trabalhar com sua própria saída e planejar uma resposta comportamental que é independente do estado ambiente atual e do corpo, e cuja meta está desligada das exigências momentâneas para a sobrevivência. Além disso, ele pode recuperar informação à vontade, sem estímulo externo ou somático, analisá-la e rearmazenar na memória (ou no ambiente) versões modificadas dessa informação para uso posterior. Chamamos isso de “processo de pensamento humano”. Além disso, por causa dessas habilidades de “comando interno” o cérebro humano pode rejeitar os comandos do sistema límbico — uma dieta é um bom exemplo! — e até se engajar em operações de processamento de informações para as quais ele não está destinado, a princípio — a matemática abstrata e a arte são bons exemplos!

Todas as funções cerebrais são baseadas em impulsos elétricos gerados, transmitidos e transferidos por neurônios. Há mais de dez bilhões dessas células no cérebro; um neurônio pode estar conectado a centenas e até milhares de outros, e cada operação cerebral, por mais “simples” que seja, envolverá milhares de neurônios. É nas interconexões sinápticas e de arquitetura desse conglomerado de neurônios que estão escondidos os mistérios da consciência, memória, pensamento e emoções. Toda operação cerebral, como o reconhecimento de um objeto que está sendo visto, a imaginação de um som musical ou um prazer que está sendo sentido, é definida por uma *distribuição espacial e temporal da atividade neural* bem específica. A “representação” do ambiente acima mencionada, ou qualquer imagem mental disso, nada mais é do que o surgimento de uma distribuição de impulsos neurais em certas áreas do córtex que, embora incrivelmente complexa, contém padrões que são absolutamente específicos do que está sendo representado ou imaginado⁸.

O cérebro humano é o sistema interativo mais complexo do Universo que conhecemos atualmente. Assim, é bem compreensível que qualquer cientista, deixando de lado qualquer pessoa cientificamente não treinada, tenha a maior dificuldade em explicar por que a função do nosso cérebro

8. Observe atentamente que esses padrões, embora absolutamente específicos, não sustentam nenhuma semelhança “pictórica” com o que representam.

nos parece tão “simples” e como a totalidade disso está totalmente sob nosso controle! (Isso é chamado de “simplicidade natural da função mental” e “natureza unitária da experiência consciente”, respectivamente). Também são compreensíveis os motivos por que temos a maior dificuldade em aceitar o fato de que, para atingir os objetivos científicos da neuropsicologia, não há necessidade de invocar conceitos separados e fisicamente indefiníveis como “mente” e “alma”!

1.5 O QUE É MÚSICA?

A discussão anterior pode ter irritado alguns leitores. A música, eles dirão, é “estética pura”, uma manifestação da inata e sublime compreensão humana do belo, e não o mero efeito de certos estímulos de ondas sonoras sobre uma complexa rede de bilhões de células nervosas. Entretanto, de acordo com a seção anterior, mesmo os sentimentos estéticos estão relacionados com o processamento neural de informação. A mistura, tão característica de padrões regulares e ordenados alternados com surpresa e incerteza, comum a toda entrada sensorial classificada como “estética”, pode ser essencialmente uma manifestação inata do homem de exercitar a sua rede neural super-redundante com operações de processamento de informações não-essenciais de complexidade variável ou alternante. Sabemos nós, realmente, o que é música? Quando falamos, transmitimos mensagens concretas. O pensamento transmitido pode ser abstrato, mas o conteúdo sonoro oferece informação. Quando ouvimos sons de origem animal, trovões, água corrente, sons ambientais, enfim, nós o fazemos em resposta a um impulso inato de tomar consciência daquilo que nos cerca usando todos os nossos sentidos (veja a seção anterior). A audição confere uma vantagem para a sobrevivência: a interpretação da informação acústica oferecida pela linguagem e pelo ambiente tem uma importância biológica fundamental. Mas qual é a informação que a música transmite? Em quase todas as culturas, a música consiste em sucessões e superposições organizadas e ritmicamente estruturadas de sons selecionados a

partir de um repertório muito limitado das alturas discretas de uma determinada escala. Os sons ambientais não apresentam um equivalente direto⁹, e imitar os sons da natureza nunca foi a principal força a conduzir o desenvolvimento de uma cultura musical. Entretanto, se a música não transmite informação biologicamente relevante, por que ela nos afeta? Belas passagens musicais podem nos deixar arrepiados, música ruim pode nos deixar irritados, um bebê que chora pode ser acalmado simplesmente pelas notas musicais que sua mãe lhe entoa. Por que essas coisas acontecem? Por que existe música?

A música pode ser um subproduto bastante natural da evolução da fala e da linguagem. Nessa evolução, que indubitavelmente foi um fator essencial para o desenvolvimento da raça humana, surgiu uma rede neural capaz de executar as ultracomplexas operações de processamento, identificação, armazenagem e recuperação de som que são necessárias para o reconhecimento fonético, a identificação da voz e a compreensão da fala.

Como a percepção da música está essencialmente baseada em processamento de informações acústicas, o maior ou menor grau de complexidade de identificação de uma mensagem sonora, o grau de sucesso nas operações de previsão que são realizadas pelo cérebro para lançar esse processo de identificação e o tipo de associações invocadas por comparação com informação armazenada de experiências anteriores podem, conjuntamente, ser a “causa” essencial das sensações musicais invocadas por uma certa mensagem musical. Se isso for verdade, deve ser bastante óbvio que tanto os mecanismos neurais inatos (operações de processamento primário) quanto o condicionamento cultural (mensagens armazenadas e operações de processamento aprendidas) devem determinar nossa resposta comportamental e estética à música.

9. O canto dos pássaros é música para nós, mas para os pássaros é informação muito concreta: p. ex., “Este território tem dono” ou “Este macho está procurando uma parceira!”.

VIBRAÇÕES SONORAS, SONS PUROS E A PERCEPÇÃO DA ALTURA

Ouvimos um som quando o tímpano entra num tipo característico de movimento chamado vibração. Essa vibração é causada por pequenas oscilações de pressão do ar do canal auditivo associadas a uma onda sonora de entrada. Neste capítulo vamos, primeiro, discutir os fundamentos do movimento vibratório periódico em geral, e então concentraremos nossa atenção nos efeitos das vibrações do tímpano no nosso sentido da audição. Por enquanto, não vamos nos preocupar com a forma como o tímpano é posto em movimento. Para esse efeito, vamos imaginar que estamos com fones de ouvido e ouvimos sons aí gerados. Os tímpanos reproduzirão com grande semelhança as vibrações dos diafragmas dos fones de ouvido. Essa abordagem para introduzir o tema não é nada ortodoxa, mas nos permitirá mergulhar direto no estudo de alguns dos conceitos principais associados à vibração sonora e sua percepção, sem perder, de início, muito tempo com ondas sonoras e geração de sons. Do ponto de vista prático, essa abordagem tem uma desvantagem: os experimentos que vamos apresentar e analisar neste capítulo exigem necessariamente a geração eletrônica de som, em vez da produção natural com instrumentos musicais de verdade.

Todavia, sempre que possível, vamos indicar de que forma um dado experimento pode ser realizado com instrumentos reais.

2.1 MOVIMENTO E VIBRAÇÃO

Movimento significa mudança de posição de certo corpo em relação a um corpo de referência. Se o corpo em movimento é muito pequeno em relação ao corpo de referência, ou em relação às dimensões do âmbito espacial coberto por seu movimento, de modo que a sua forma é praticamente irrelevante, o problema fica reduzido à descrição do movimento de um ponto no espaço. Eis por que um corpo assim pequeno é chamado frequentemente de ponto material ou partícula. Por outro lado, se o corpo não é pequeno, mas sabemos de antemão, por circunstâncias particulares, que todos os pontos do corpo estão confinados a mover-se em linhas retas paralelas uma à outra (“translação retilínea”), também será suficiente especificar o movimento de apenas um ponto do corpo. Esse é um caso “unidimensional” de movimento, e a posição desse ponto do corpo (e, portanto, de todo o corpo) fica completamente especificada com apenas um número: a distância até um ponto fixo de referência.

Neste livro só vamos lidar com movimentos unidimensionais. Vamos assumir que o nosso ponto material se mova ao longo de uma linha vertical (Figura 2.1). Vamos designar o ponto de referência nessa linha com a letra O . Qualquer ponto fixo pode servir como referência, muito embora, por conveniência, às vezes selecionamos um em especial (tal como a posição de equilíbrio para um determinado movimento oscilatório). Indicamos a posição de um ponto material P pela distância y até o ponto de referência O (Figura 2.1). y também é chamado de deslocamento de P em relação a O , ou coordenada de P . Precisamos usar tanto números positivos quanto negativos para distinguir entre os dois lados com relação a O ¹.

1. Na física, o sistema métrico é usado para medir distâncias. A unidade de comprimento é o metro (1 m = 3.28 pés); também são usados vários submúltiplos decimais (p. ex., centí-

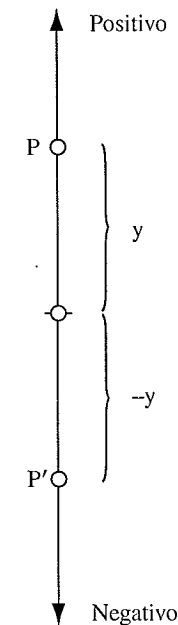


Figura 2.1

O ponto material P estará em movimento em relação a O quando a sua posição y mudar com o tempo. Vamos indicar o tempo com a letra t . Ele é medido com um relógio — e torna-se necessário especificar também um instante de “referência” $t = 0$. O movimento pode ser representado matematicamente de duas maneiras: analiticamente, usando as chamadas relações funcionais, e geometricamente, usando uma representação gráfica. Usaremos apenas o método geométrico. Para representar graficamente um movimento unidimensional, introduzimos dois eixos perpendiculares um ao outro, um representando o tempo t , e o outro a coordenada y (Figura 2.2). Em ambos temos que indicar claramente a escala, *i.e.*, os intervalos unitários (de tempo e deslocamento, respectivamente). Um movimento

metro = 0,01 m = 0,394 polegadas, ou o milímetro = 0,001 m), e também múltiplos (p. ex., o quilômetro = 1000 m = 0,625 milhas).

pode ser representado traçando-se para cada instante t a distância y na qual a partícula está momentaneamente localizada. Cada ponto da curva resultante, como S_1 (Figura 2.2), nos diz que, em $t = t_1$, a partícula P está a uma distância y_1 de O , *i.e.*, na posição P_1 . Observe que, neste gráfico, o ponto material não se move ao longo da curva de pontos S ! Essa curva é apenas um “auxílio” fabricado para nos ajudar a encontrar a posição y da partícula em qualquer instante t .

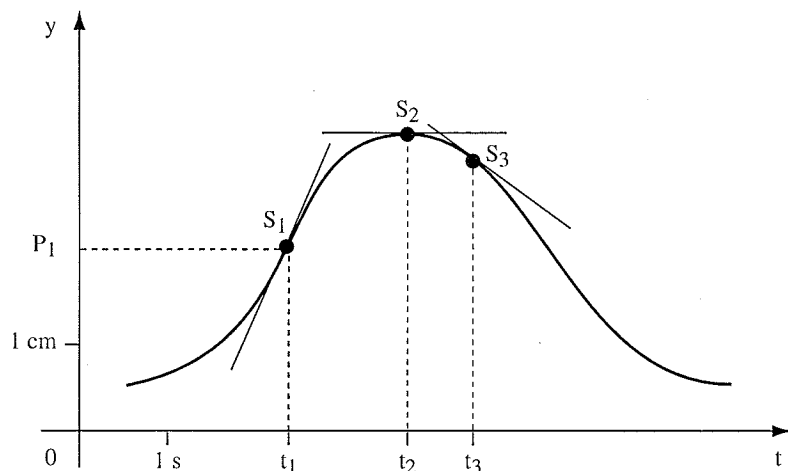


Figura 2.2

O gráfico da Figura 2.2 também fornece informação sobre a velocidade do ponto material, *i.e.*, a rapidez com que ele muda a sua posição. Ela pode ser determinada pela declividade da curva no gráfico: em t_1 a partícula está se movendo para cima com certa velocidade, em t_3 ela está descendo com velocidade menor. Em t_2 , ela está momentaneamente em repouso, mudando a sua direção.

Há um certo tipo de movimento em que o ponto material segue um padrão temporal que se repete sempre. Este é o chamado movimento periódico, ou vibração. É o tipo de movimento de maior importância na física

da música. Para que se tenha um movimento realmente periódico, um corpo não só tem que voltar repetidamente à mesma posição, mas deve fazê-lo em intervalos de tempo exatamente iguais e repetir exatamente o mesmo tipo de movimento enquanto isso. O intervalo de tempo depois do qual o padrão de movimento se repete é chamado de período (Figura 2.3). Nós o denotamos com a letra grega *tau* (τ). Durante um período, o movimento pode ser muito simples (Figura 2.3a) ou bastante complicado (Figura 2.3b). O padrão de movimento que ocorre durante um período e que será sempre repetido é chamado de ciclo².

Existem dispositivos mecânicos e eletrônicos que podem traçar automaticamente o gráfico de um movimento periódico. Num registrador gráfico, a agulha reproduz na direção y o movimento periódico que deve ser descrito, enquanto traça esse movimento numa fita de papel que se move perpendicularmente ao eixo y a uma velocidade constante. Como conhecemos essa velocidade, podemos atribuir uma escala de tempo ao eixo ao longo da fita de papel. A curva obtida é a representação gráfica do movimento. Entretanto, esse método não é prático para o registro de vibrações acústicas. Elas têm períodos tão curtos que seria impossível deslocar uma agulha com a rapidez suficiente para reproduzir este tipo de vibração. Um dispositivo eletrônico chamado osciloscópio serve para esse propósito. Em essência, ele consiste em um feixe de elétrons bem estreito (partículas elementares com carga elétrica negativa) que incide numa tela de TV, gerando um ponto luminoso claramente visível. Esse feixe pode ser defletido tanto na direção vertical como na horizontal. O movimento vertical é controlado por um sinal proporcional à vibração que nós queremos mostrar (por exemplo, a vibração do diafragma de um fone de ouvidos). O movimento horizontal é uma varredura contínua para a direita com velocidade constante, equivalente ao movimento da tira de papel no registrador gráfico, representando assim a escala de tempo. O ponto luminoso na tela, portanto, descreve o gráfico do movimento durante uma varredura. Se a ima-

- Um bom exercício para compreender gráficos como os das Figuras 2.2 ou 2.3 é procurar descrevê-los com um movimento equivalente de sua mão para cima e para baixo, representando a variação de distância (eixo y) com o tempo (eixo x).

gem do ponto luminoso for suficientemente retida, ela aparece como uma curva contínua na tela. Como a tela tem tamanho limitado, o movimento horizontal é instantaneamente reconduzido à origem sempre que o feixe alcança a margem direita da tela, e a varredura começa outra vez.

Para representar um movimento periódico, a varredura deve estar sincronizada com o período τ ou um de seus múltiplos.

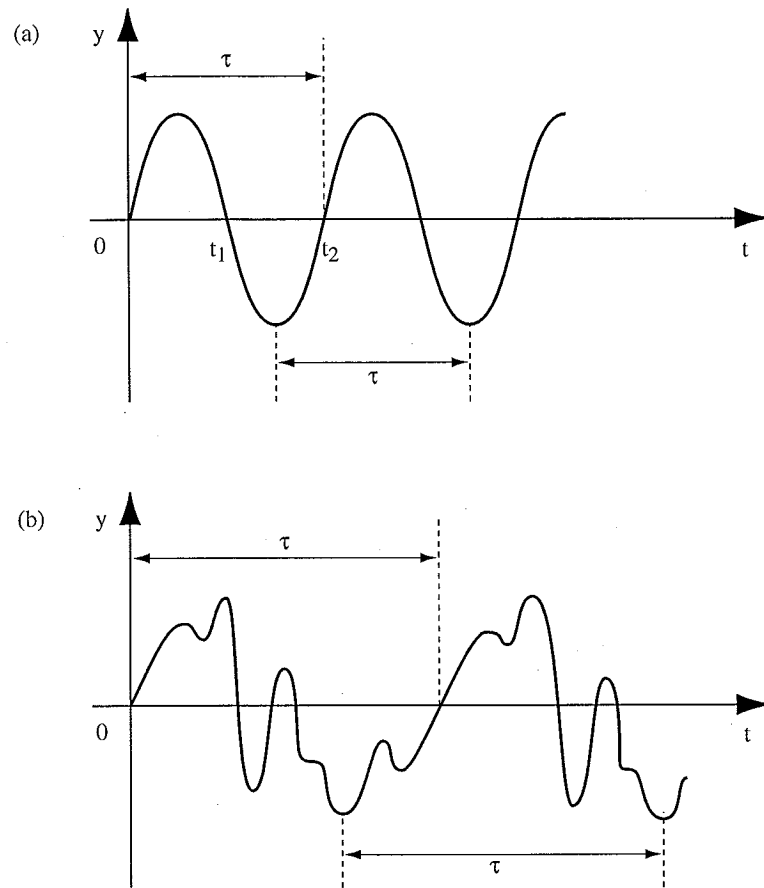


Figura 2.3

2.2 MOVIMENTO HARMÔNICO SIMPLES

A questão que surge agora é saber qual o tipo mais “simples” de movimento periódico. Existem muitos exemplos na natureza: as oscilações de um pêndulo para a frente e para trás, o movimento para cima e para baixo de uma mola, as oscilações das moléculas etc. Esses movimentos têm algo importante em comum: todos podem ser representados como a projeção de um movimento circular uniforme sobre um diâmetro do círculo (Figura 2.4)³. Quando o ponto R gira uniformemente (com período τ , *i.e.*, um giro a cada τ segundos), o ponto de projeção P move-se para cima e para baixo ao longo do eixo y com o que se chama de movimento harmônico simples (veja o gráfico à direita da Figura 2.4). Ele é chamado também de movimento senoidal (porque o y pode ser representado analiticamente por uma função trigonométrica chamada *seno*).

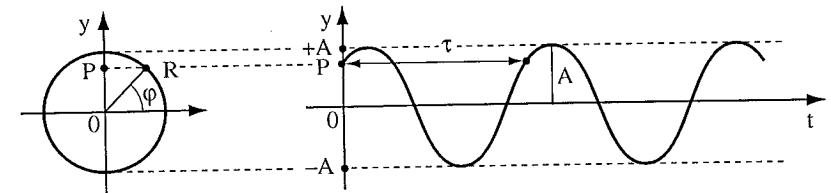


Figura 2.4

Observe que um movimento harmônico simples representa uma vibração que é simétrica em relação ao ponto O , chamado de posição de equilíbrio. O deslocamento máximo A (para cima ou para baixo) é chamado de amplitude. τ é o período do movimento harmônico. Existe outro parâ-

3. Observe com atenção que a montagem do lado esquerdo da Figura 2.4 é auxiliar; o único movimento real é o movimento periódico, para cima e para baixo, da partícula P ao longo do eixo y .

metro que descreve o movimento harmônico simples e que é um pouco mais difícil de se entender. Considere a Figura 2.4: no instante inicial $t = 0$, a partícula (projecção do ponto R) encontra-se na posição P . Podemos agora considerar um segundo caso de movimento harmônico com o mesmo período τ e a mesma amplitude A , mas no qual a partícula parte de uma posição diferente Q (Figura 2.5). O movimento resultante, obviamente, será diferente, não na forma, mas no “sincronismo” relativo. Na verdade, como se vê na Figura 2.5, ambas as partículas passarão numa certa posição (p. ex., a origem O) em instantes diferentes (t_1, t_2). De modo oposto, ambas as partículas estarão, em geral, em posições diferentes num dado instante (p. ex. P e Q em $t = 0$). Se novamente imaginarmos o movimento da segunda partícula Q como a projeção de um movimento circular uniforme (Figura 2.5), veremos que cada caso se refere a posições angulares diferentes φ_1, φ_2 dos pontos associados que giram sobre o círculo. O ângulo φ é chamado de fase do movimento harmônico simples; a diferença $\varphi_1 - \varphi_2$ (Figura 2.5), que permanece constante nesse exemplo, é chamada de diferença de fase entre os dois movimentos harmônicos⁴.

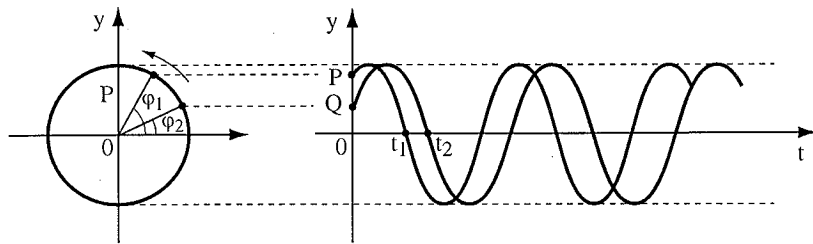


Figura 2.5

Em resumo, certa vibração “pura”, ou harmônica, é especificada pelos valores de três parâmetros: o período τ , a amplitude A e a fase φ (Figura

4. Use as mãos para descrever os dois movimentos representados no lado direito da Figura 2.5. Utilize várias diferenças de fase, por exemplo, 180° (fases opostas), 0° (em fase), 90° etc.

2.4). Esses parâmetros, e em especial os dois primeiros, desempenham um papel muito importante na percepção dos sons musicais.

Os movimentos harmônicos simples ocorrem praticamente em todos os lugares do Universo: as vibrações dos componentes dos átomos, dos átomos como um todo nos cristais, dos corpos elásticos etc. podem ser descritas em termos de movimentos harmônicos simples. Mas existe uma razão ainda mais poderosa para considerarmos o movimento harmônico simples como o mais básico de todos os movimentos periódicos: é possível demonstrar matematicamente que qualquer tipo de movimento periódico, por mais complicado que seja, pode ser descrito como uma soma de vibrações harmônicas simples. Posteriormente (Capítulo 4) vamos lidar mais detalhadamente com essa propriedade fundamental. Ela é de capital importância para a música.



2.3 VIBRAÇÕES ACÚSTICAS E SENSações DE SONS PUROS

Quando o tímpano entra em movimento periódico, as suas vibrações mecânicas são convertidas no ouvido interno em impulsos nervosos elétricos que são enviados ao cérebro e interpretados como som, desde que o período e a amplitude das vibrações se encontrem dentro de certos limites. Em geral, o ouvido é um dispositivo extremamente sensível: ele é capaz de detectar vibrações do tímpano com amplitudes de até 10^{-7} cm, e com períodos de até 7×10^{-5} segundos⁵.

Apresentamos agora a quantidade que é usada com mais constância do que o período τ e que é chamada de frequência:

$$f = \frac{1}{\tau} \tag{2.1}$$

5. Neste livro, usaremos a notação exponencial: $10^{+n} = 1(n \text{ zeros})$. $10^{-n} = 1/10^{+n}$, ou seja, uma fração decimal dada pela unidade da n-ésima posição decimal.

Fisicamente, f representa o número de repetições do padrão de vibração, ou ciclos, por unidade de tempo. A razão para se preferir f a τ é que a frequência aumenta quando a nossa sensação de “altura do som” (graves e agudos) aumenta. Se τ é dado em segundos, f é expressa em ciclos por segundo. Essa unidade é chamada *hertz* (Hz), em homenagem a Heinrich Hertz, um famoso físico alemão. As vibrações no intervalo 20 Hz - 15 000 Hz são percebidas como som por uma pessoa normal. Tanto o limite inferior quanto o superior (e particularmente este) dependem da intensidade do som e variam consideravelmente de pessoa para pessoa e com a idade.

Quando um som causa um movimento harmônico simples do tímpano com características constantes (frequência, amplitude, fase), ouvimos o que é chamado de som puro. Um som puro soa desinteressante, e a música não é feita simplesmente de sons puros. Entretanto, como já foi colocado na introdução deste capítulo, para uma melhor compreensão dos sons complexos é aconselhável lidar primeiro apenas com os sons simples ou puros. Os sons puros precisam ser gerados com osciladores eletrônicos; não há nenhum instrumento musical que os produza (e mesmo quanto aos sons puros gerados eletronicamente não há nenhuma garantia de que eles estejam realmente “puros” quando chegarem ao ouvido). De qualquer modo, como a flauta é o instrumento cujo som mais se aproxima de um som puro e senoidal, principalmente no registro agudo, muitos dos experimentos mencionados neste capítulo (mas não todos) poderão ser realizados em casa, usando uma ou duas flautas — porém tocadas por *experts*!

Quando ouvimos um som puro cuja frequência e amplitude podem ser variadas à vontade, verificamos que há uma correspondência entre altura e frequência e entre volume e amplitude. Temos uma idéia mais ou menos boa de como o mecanismo de detecção primária de frequência e amplitude do ouvido trabalha em relação aos sons puros. Neste capítulo, vamos considerar apenas a altura. As oscilações harmônicas simples do tímpano são transmitidas por uma cadeia de três pequenos ossos chamados martelo, bigorna e estribo (ou, em termos mais eruditos, *malleus*, *incus* e *estapedio*, que são palavras latinas) à entrada (“janela oval”) do ouvido interno propriamente dito (Figura 2.6). A *cóclea*, do tamanho de uma bola de gude, é um túnel em espiral, como uma casca de caracol, através do osso temporal

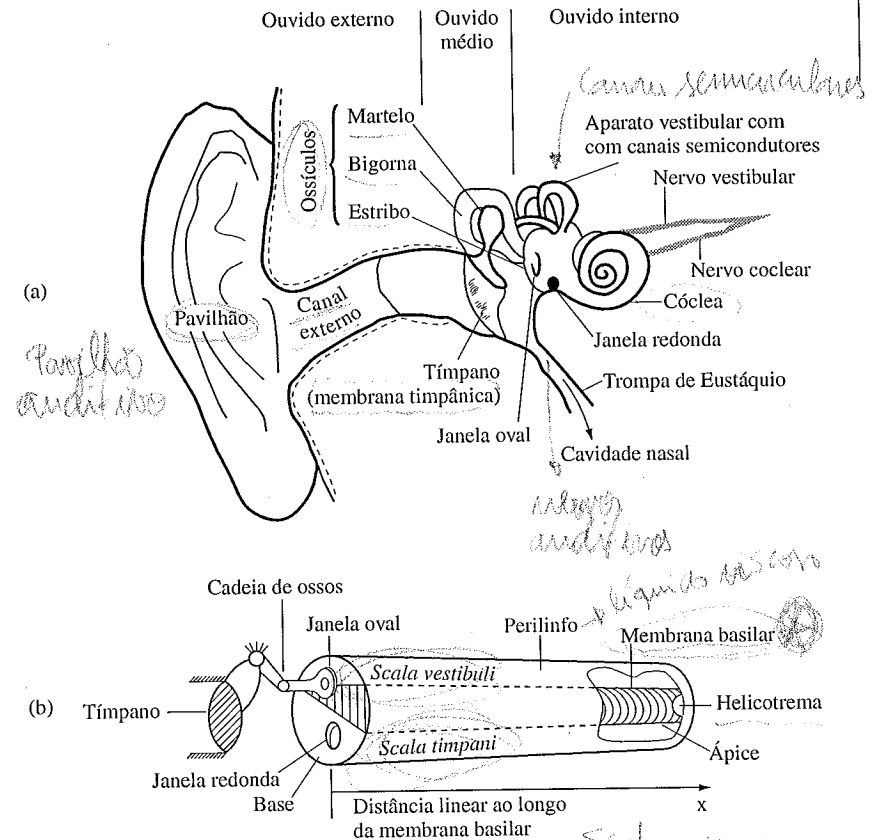


Figura 2.6 (a) vista esquemática do ouvido (Flanagan, 1972, figura 4) (fora de escala); (b) a cóclea desenrolada (bem simplificada).

humano. Essa cavidade, vista na Figura 2.6b numa versão estendida e altamente simplificada, está dividida em dois canais, *scala vestibuli* e *scala timpani*, preenchidos com um líquido incompressível, o *perilinfo* (um líquido filtrado diretamente do fluido cérebro-espinhal). Ambos os canais se comportam como um único sistema hidrodinâmico por estarem conectados numa das extremidades, o ápice, por meio de um pequeno orifício no duto que os separa. Esse orifício se chama *helicotrema*. A seção inferior está isolada por uma membrana elástica na “janela redonda”

Membrana basilar
apice
glóbulo aliaris (16.000 na cóclea)
apice

(Figura 2.6b). A partição que separa ambas as *scalae* é, ela própria, um duto altamente estruturado, de seção transversal triangular (também chamado de *scala media*; Figura 2.7a), preenchido com outro fluido, o *endolinfo*. As suas fronteiras são a *membrana basilar*, que comporta o próprio órgão sensorial (órgão de Corti), a membrana de Reissner, que serve para separar o perilinfo do endolinfo, e a parede lateral rígida da cóclea.

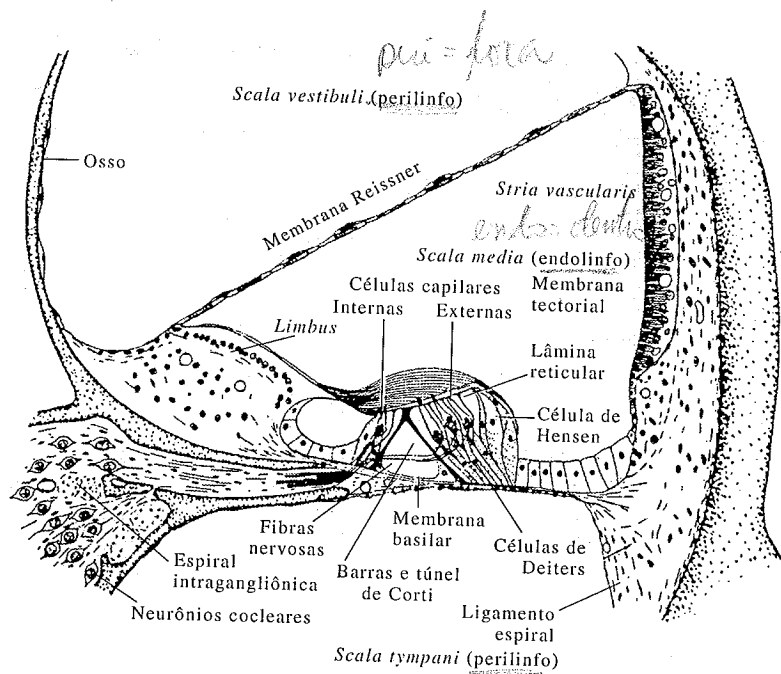


Figura 2.7a Seção transversal do órgão de Corti (de Davis, 1962).

A elasticidade da membrana basilar determina as propriedades hidrodinâmicas básicas da cóclea. Num ser humano adulto, a membrana tem cerca de 34 mm de comprimento desde a base (a extremidade de entrada) até o ápice; por causa da sua variação gradual de largura e espessura, há um decréscimo de 10 000 vezes na rigidez, desde a base até o ápice, o que dá à membrana basilar a sua função fundamental de análise de frequências. As vibrações transmitidas pela cadeia de ossos à janela oval são convertidas em oscilações de pressão do perilinfo na *scala vestibuli*. As dife-

renças de pressão resultantes, através da divisão coclear, entre as duas *scalae* flexibilizam a membrana basilar para cima e para baixo, colocando-a em movimento como uma bandeira tremulante; à medida que essa onda se propaga para o ápice, a sua amplitude chega ao máximo num certo ponto que dependerá da frequência de entrada e será amortecida rapidamente na direção do ápice. Cerca de 16 000 unidades receptoras, chamadas *células capilares*, dispostas em uma fileira “interna” e três fileiras “externas” ao longo da membrana basilar, captam os movimentos desta última e passam os sinais às células nervosas, ou neurônios, que estão em contato com elas. O nome “célula capilar” provém do fato de que na sua extremidade há um feixe de 20 a 300 pequenas apófises chamadas *estereocílios* (Figura 2.7b) projetando-se para dentro do fluido *endolinfático* e cuja deflexão

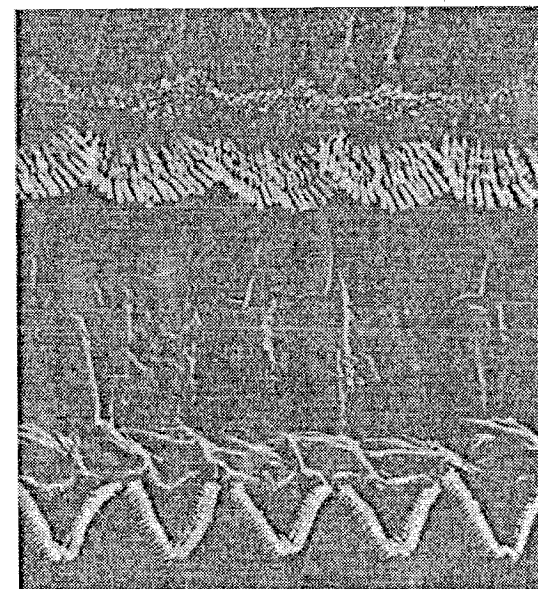


Figura 2.7b Micrografia por varredura eletrônica (Bredberg e outros, 1970) dos *estereocílios* das fileiras interna (acima) e externa (abaixo — vê-se apenas uma de três fileiras) de células capilares na membrana basilar de uma cobaia (esses animais, assim como as chinchilas, têm sistemas acústicos periféricos muito semelhantes aos humanos e são os animais de laboratório mais utilizados em pesquisas sobre audição). © 1970 American Association for the Advancement of Science.

dispara uma cadeia de processos eletroquímicos na célula capilar e em suas vizinhanças, culminando com a geração de sinais elétricos no nervo acústico. A *membrana tectorial* é um tecido gelatinoso suspenso no endolinfa acima do órgão de Corti (Figura 2.7a), no qual estão inseridos os cílios das células capilares externas; ela tem um papel fundamental na recepção e estimulação da ação motora dessas células. Voltaremos à função coclear com mais detalhes nas Seções 2.8 e 3.6.

O interessante é que, para um som puro de certa frequência, as oscilações máximas da membrana basilar ocorrem apenas numa região limitada da membrana, cuja posição depende da frequência do som. Em outras palavras, para cada frequência existe uma região de estimulação máxima, ou “região de ressonância”, na membrana basilar. Quanto menor a frequência do som, mais próxima ao ápice (Figura 2.6b) estará a região das células capilares ativadas (onde a membrana é mais flexível). Quanto maior a frequência, mais próxima à entrada (janela oval) estará essa região (onde a membrana é mais rígida). A posição espacial x ao longo da membrana basilar (Figura 2.6b) das células capilares e neurônios associados que respondem ao estímulo determina a sensação primária de altura (também chamada de altura espectral). Uma mudança na frequência do som puro causa uma mudança na posição da região ativada; essa mudança é então interpretada como uma variação de altura. Dizemos que a informação primária sobre a frequência do som é “codificada” pelo órgão sensorial da membrana basilar na forma de localização espacial dos neurônios ativados. Dependendo de qual grupo de neurônios é ativado, o som nos parecerá grave ou agudo.

A Figura 2.8 mostra como a posição x (medida a partir da base, Figura 2.6b) da região de máxima sensibilidade varia com a frequência de um som puro e senoidal, para uma pessoa normal adulta (Von Béséký, 1960). Várias conclusões importantes podem ser tiradas. Em primeiro lugar, note que a gama de frequências musicalmente mais importantes (aproximadamente 20-4000 Hz) cobre cerca de dois terços da extensão da membrana basilar (12-35 mm a partir da base). A grande porção restante da escala de frequências (4000-16000 Hz; a Figura 2.8 mostra apenas até 5000 Hz) está espremida dentro do terço remanescente. Em segundo lugar, observe o significativo fato de que, sempre que a frequência de um som é dobrada, *i.e.*, a

altura sobe uma oitava, a região de ressonância correspondente é deslocada de uma magnitude mais ou menos constante de 3,5-4 mm, não importando se esse pulso de frequência é de 220 para 440 Hz, de 1760 para 3520 Hz, ou mesmo de 5000 para 10000 Hz. Em geral, sempre que a frequência f é multiplicada por um certo fator, a posição x da região de ressonância não é multiplicada, mas simplesmente deslocada de certa quantidade. Em outras palavras, o que determina o deslocamento da região de ressonância na membrana basilar é a razão entre as frequências, e não a sua diferença. Uma relação desse tipo é chamada de “logarítmica” (Seção 3.4).

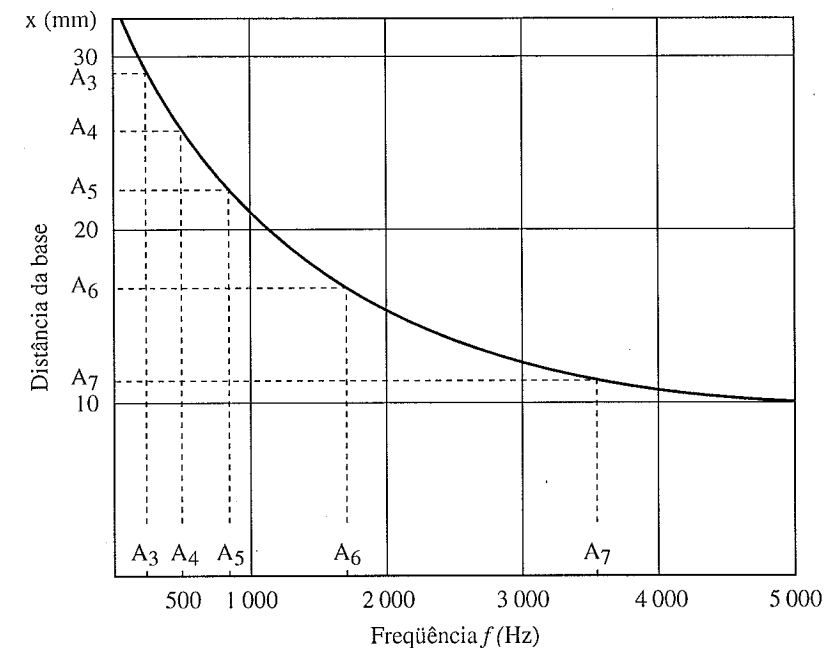


Figura 2.8 Posição do máximo de ressonância da membrana basilar (segundo Von Béséký, 1960) para um som puro de frequência f (escalas lineares).

Os resultados acima provêm de medições fisiológicas (Von Béséký, 1960), que podem ser realizadas em animais anestesiados (*i.e.*, inconscientes), ou mesmo mortos (mas bem conservados). Atualmente essas

medições podem ser feitas em cócleas vivas por efeito Mössbauer. Uma ínfima porção de substância radioativa (cobalto 57) é “implantada” na membrana basilar (o que, até certo ponto, pode afetar o que está sendo medido). Os pequenos deslocamentos da membrana podem ser detectados indiretamente medindo-se o desvio de frequência (efeito Doppler) dos raios gama emitidos pela substância (p. ex., Rhode e Robles, 1974).

Consideremos agora a magnitude psicofísica altura, associada a um som puro de frequência f . Na Seção 1.4, dissemos que uma magnitude psicofísica não pode ser medida da mesma forma quantitativa que uma magnitude física, como a frequência. Apenas certa ordem pode ser estabelecida pela experimentação individual entre duas sensações do mesmo tipo apresentadas numa sucessão imediata. As estimativas quantitativas só são possíveis depois que o cérebro, de alguma forma, está treinado para realizar as operações necessárias (p. ex., uma criança aprendendo a avaliar o tamanho dos objetos que vê) — e os resultados teriam que ser interpretados de forma estatística. Mesmo um sujeito não treinado pode julgar se o “som 2” é mais agudo, igual ou mais grave do que o “som 1”, quando ambos lhe são apresentados em seqüência⁶. Porém, estaria ele apto a julgar quanto — p. ex., o dobro, um terço, dez vezes — um som é mais grave ou mais agudo do que o outro?

Vamos considerar a habilidade individual de estabelecer uma ordem relativa de alturas quando dois sons puros (de mesma intensidade) são ouvidos um após o outro. Existe um limite natural: quando a diferença de frequência entre os dois sons é muito pequena, abaixo de certo valor, ambos os sons serão ouvidos como se tivessem a mesma altura. Isso ocorre com os julgamentos de ordem para todas as magnitudes psicofísicas: sempre que a variação de um estímulo físico original encontrar-se dentro de certo “limiar de diferença” ou “diferença no limite do observável” (abreviada “dlo”), o julgamento será que a sensação associada permanece “a mesma”; tão logo a variação exceda a dlo, uma variação na sensação é detectada. Observe que a dlo se relaciona a uma magnitu-

6. Isso, naturalmente, implica uma concordância mútua anterior sobre o que significa uma altura “mais aguda” ou “mais grave”.

de física (o estímulo), é mensurável no sentido comum e pode ser expressa por um número,

O grau de sensibilidade às variações de frequência do mecanismo primário de percepção de alturas, ou capacidade de resolução de frequências, depende da frequência, intensidade e duração do som em questão — e também da subitaneidade da variação de frequência. Ela varia bastante de pessoa para pessoa, é uma função do treinamento musical e, infelizmente, depende consideravelmente do método de medição empregado. A Figura 2.9 mostra a dlo média de frequência para sons puros de intensidade constante (80 db, Seção 3.4), cuja frequência foi lenta e continuamente modulada para cima e para baixo (Zwicker, Flottorp e Stevens, 1957). O gráfico abaixo mostra, por exemplo, que, para um som de 2000 Hz, pode-se detectar uma variação de 10 Hz — *i.e.*, apenas 0,5%. Essa é uma fração muito pequena de um semitom! Variações súbitas de frequência são detectadas com uma dlo consideravelmente mais baixa — até 30 vezes menor do que os valores vistos na Figura 2.9 (Rakowsky, 1971). A resolução de frequências torna-se pior em baixas frequências (p. ex., 3% em 100 Hz na Figura 2.9). Ela também decresce com a diminuição da duração do som, desde que ele decaia em menos de um décimo de segundo.

IMPORTANTE Um Af 6

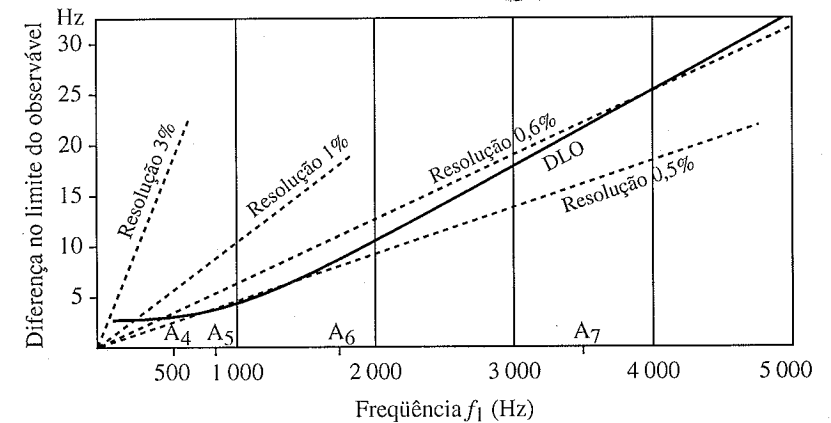


Figura 2.9 Diferença no limite do observável de frequência para um som puro com frequência f_1 (escalas lineares), determinada com um sinal de frequência lentamente modulante (Zwicker e outros, 1957).

Aparentemente, a dlo em frequência (e, na verdade, a dlo em outros estímulos físicos) não está definida “nitidamente”. Experimentos recentes revelaram que existem dois valores críticos de limite:

1. O incremento de frequência abaixo do qual nunca se detecta uma variação de altura que foi chamado de *quantum neural* de detecção de alturas.
2. O incremento de frequência acima do qual uma variação de altura sempre pode ser detectada, e cujo valor é mais ou menos o dobro do *quantum neural*.

As variações de frequência de valores entre esses dois valores-limites às vezes são detectadas, com uma probabilidade que cresce linearmente de 0 a 1. Na discussão que se segue, entretanto, ignoraremos essa “estrutura fina” da dlo. Para uma discussão atual da dlo, veja Zwicker e Fastl (1990).

Desde o advento da psicofísica, os psicólogos vêm tentando considerar a variação mínima perceptível na sensação causada por uma dlo do estímulo como a unidade “natural”, com a qual se pode medir a magnitude psicofísica correspondente. A variação mínima perceptível de altura foi usada para construir uma “escala subjetiva de altura” (Stevens, Volkman e Newman, 1937). Todavia, uma vez que, conforme veremos mais tarde, a oitava desempenha o papel crucial de intervalo de altura “natural”, e como todas as escalas musicais se desenvolveram de forma independente das tentativas de se estabelecer uma escala subjetiva de alturas, esta última tem apenas um interesse secundário em música, e não mais será considerada neste livro.

Em resumo, de acordo com a descrição dada nesta seção, a função primária do ouvido interno (cóclea) é converter um padrão de vibração dependente do tempo (o do tímpano) em um padrão de vibração dependente do espaço (o da membrana basilar), e este, por sua vez, num padrão espacial de atividade neural. A descrição teórica desse mecanismo é chamada de Teoria de Posicionamento da Audição (ou Teoria da Altura com Base Espectral). Voltaremos a ela freqüentemente; veremos que se trata de uma boa teoria, mas não completa. Existe ainda outro mecanismo mais complicado de codificação de alturas operando no sistema nervoso.

2.4 SUPERPOSIÇÃO DE SONS PUROS: BATIMENTOS DE PRIMEIRA ORDEM E A BANDA CRÍTICA

Já dissemos antes que os sons puros individualmente soam desinteressantes. As coisas tornam-se um pouco mais vivas a partir do momento em que superpomos dois sons puros, fazendo-os soar simultaneamente. Nesta seção vamos analisar as características fundamentais da superposição de dois sons puros. Encontraremos alguns conceitos bastante fundamentais da física da música e da psicoacústica.

Existem dois tipos de efeitos de superposição, dependendo de onde eles forem processados no sistema auditivo do ouvinte. Se o processamento for mecânico, ocorrendo no fluido coclear e ao longo da membrana basilar, chamaremos esses efeitos de “efeitos de superposição de primeira ordem”, principalmente porque eles são claramente distinguíveis e de fácil acesso à experimentação psicoacústica. Os efeitos de superposição de “segunda ordem” são conseqüência de processamento neural e são mais difíceis de se detectar, descrever e medir sem ambigüidades. Nesta seção, vamos nos deter apenas nos efeitos de primeira ordem.

Inicialmente vamos discutir o significado físico de “superposição de som”. O tímpano move-se para dentro e para fora segundo as variações de pressão do ar no canal auditivo. Se ele for posto a oscilar num movimento harmônico puro de amplitude e frequência definidas, ouviremos um som puro com intensidade e altura definidas. Mas, se fizermos soar juntos dois sons com características diferentes (p. ex., ouvindo duas fontes independentes ao mesmo tempo), o tímpano reage como se estivesse executando simultaneamente dois comandos independentes, um para cada som puro. O movimento resultante é a soma dos movimentos individuais que ocorreriam se cada som puro soasse sozinho, na ausência do outro. Não só o tímpano se comporta dessa forma, mas também o meio e todos os outros componentes vibrantes (isso, entretanto, não é totalmente verdade se as amplitudes forem muito grandes). Esse efeito é chamado de superposição linear de duas vibrações. A superposição linear de duas vibrações é um termo técnico que significa “coexistência pacífica”: uma

vibração componente não interfere nos negócios da outra, e a superposição resultante simplesmente segue o que dita cada componente simultaneamente. Numa superposição não-linear, a ordem de cada componente dependerá do que o outro tenha a dizer, e vice-versa.

Vamos começar nossa discussão com a análise da superposição de dois movimentos harmônicos simples com frequências iguais e fases iguais (diferença de fase nula, Seção 2.2). Pode-se mostrar graficamente e também analiticamente que, nesse caso, novamente obteremos um movimento harmônico simples com a mesma frequência e a mesma fase, mas com uma amplitude que é a soma das amplitudes das vibrações componentes. Se as duas oscilações componentes de frequência fixada tiverem diferença de fase, a sua superposição ainda será um movimento harmônico simples com a mesma frequência, mas a amplitude já não será mais a soma das amplitudes componentes. Em particular, se as amplitudes das vibrações componentes forem iguais e a diferença de fase ϕ for 180° , as oscilações se aniquilarão uma à outra e não se ouvirá nenhum som. Isso é chamado de interferência destrutiva, que tem importante papel em acústica de ambientes. Em resumo, quando dois sons puros de mesma frequência chegam ao tímpano, percebemos apenas um som de frequência definida (correspondente à frequência dos sons componentes) e volume igualmente definido (controlado pelas amplitudes dos sons superpostos e suas diferenças de fase).

Consideremos agora a superposição de dois sons simples de mesma amplitude, mas com frequências ligeiramente diferentes, f_1 e $f_2 = f_1 + \Delta f$. A diferença de frequência Δf tem um valor pequeno. Vamos assumir que ela seja positiva; o som correspondente a f_2 será, então, ligeiramente mais agudo do que o de f_1 . O padrão de cada tímpano será dado pela soma dos padrões de cada som componente (Figura 2.11). O resultado da superposição (curva mais cheia) é uma oscilação com período e frequência intermediários a f_1 e f_2 , e com uma amplitude lentamente modulada. Observe nesta figura a diferença de fase lentamente variável entre os sons componentes y_1 e y_2 : eles começam em fase (0° de diferença de fase, como na Figura 2.10) no instante $t = 0$, e então y_2 começa a ganhar em fase (à frente de y_1) até que ambos estejam completamente fora de fase (180° de dife-

rença de fase) no instante correspondente a C . A diferença de fase continua crescendo até atingir $360^\circ = 0^\circ$ no instante τ_B . Essa variação de fase contínua e lenta é responsável pela amplitude variável da oscilação resultante: as curvas pontilhadas na Figura 2.11 representam o envelope de amplitude da vibração resultante (veja também a Figura 2.17A).

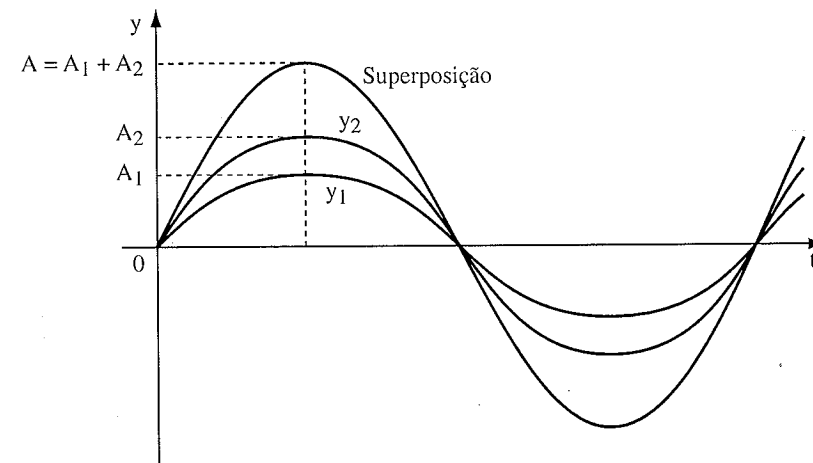


Figura 2.10

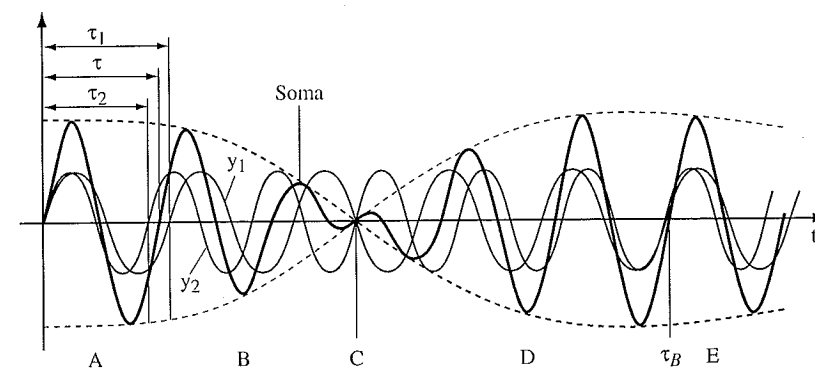


Figura 2.11

Qual será a sensação tonal resultante nesse caso? Em primeiro lugar, observe que cada tímpano seguirá a oscilação correspondente à curva mais cheia da Figura 2.11. O tímpano “não sabe” e “não quer saber” se esse padrão é na verdade resultado da soma de dois outros. Ele tem apenas um padrão de vibração, com amplitude variável. Um fato mais notável ocorre no fluido coclear: esse padrão de vibração único, porém um tanto complicado, dá origem, na janela oval, a dois grupos de ondas propagantes, um para cada som componente original! Se a diferença de frequência Δf for suficientemente grande, as regiões de ressonância correspondentes na membrana basilar estarão suficientemente separadas uma da outra; cada uma oscilará com uma frequência correspondente ao som componente (curvas mais leves na Figura 2.11), e ouviremos dois sons separados com volume constante, com alturas correspondentes às dos sons originais. Essa propriedade da cóclea de desmembrar um padrão de vibração complexo nos seus sons componentes puros é chamada de discriminação de frequência. Trata-se de um processo mecânico, controlado pelas propriedades hidrodinâmicas e elásticas dos constituintes do ouvido interno. Por outro lado, se a diferença de frequência Δf for menor do que certa quantidade, as regiões de ressonância se sobrepõem, e ouviremos apenas um som com altura intermediária e amplitude modulada ou “em batimento”. Nesse caso, a região superposta da membrana basilar seguirá um padrão de vibração essencialmente idêntico ao do tímpano (curva mais cheia da Figura 2.11). A modulação de amplitude do padrão de vibração (curvas envoltórias na Figura 2.11) causa a modulação percebida no volume. Chamamos esse fenômeno de “batimentos de primeira ordem”. Esses são os batimentos comuns, conhecidos por todos os músicos.

A frequência do padrão de vibração resultante de dois sons com frequências muito semelhantes f_1 e f_2 é igual ao valor médio:

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

O intervalo de tempo τ_B (Figura 2.11), após o qual a amplitude resultante atinge o valor inicial, é chamado de período de batimento. A frequência de

batimento (número de variações de amplitude por segundo) $f_B = 1/\tau_B$ será dada pela diferença:

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

Não faz diferença se f_2 é maior que f_1 ou vice-versa. Os batimentos serão ouvidos em ambos os casos, e a sua frequência será sempre dada pela diferença de frequência dos sons componentes (na verdade, a relação 2.3 deve ser tomada sempre em seu valor absoluto, *i.e.*, sempre positivo). Quanto mais próximas estiverem as frequências f_1 e f_2 , tanto mais “lentos” resultarão os batimentos. Se f_2 ficar igual a f_1 , os batimentos desaparecerão completamente: os sons componentes soarão em uníssono.

Vamos resumir as sensações tonais evocadas pela superposição de dois sons puros de mesma amplitude e com frequência f_1 e $f_2 = f_1 + \Delta f$, respectivamente. Para isso, vamos assumir que iremos manter f_1 fixo e aumentarmos f_2 lentamente a partir de f_1 (uníssono, $\Delta f = 0$) para valores mais altos. (Qualitativamente, nada mudaria no que se segue, se diminuíssemos f_2). No uníssono, nós ouvimos apenas um som com altura correspondente a f_1 e um volume que dependerá da diferença de fase particular entre os dois sons. Quando aumentamos ligeiramente a frequência f_2 , continuamos ouvindo um único som, mas com altura ligeiramente maior, correspondendo à frequência média $f = f_1 + \Delta f/2$ (2.2)⁷. O volume desse som sofrerá batimentos com uma frequência Δf (2.3). Esses batimentos aumentarão em frequência à medida que f_2 se afasta de f_1 (Δf aumenta). Desde que f seja menor que uns 10 Hz, esses batimentos serão percebidos claramente. Quando a diferença de frequência Δf exceder, digamos, 15 Hz, a sensação de batimento desaparece, dando lugar a um aspecto rude ou desagradável bem característico da sensação tonal resultante. Quando Δf ultrapassa o chamado limite de discriminação de frequência Δf_D (não confundir com o limiar de resolu-

7. Como podemos verificar que, de fato, a sensação de altura do som resultante corresponde a um som com frequência $f_1 + \Delta f/2$? Isso é conseguido com os experimentos de casamento de alturas: o indivíduo ouve simultaneamente um som de referência com altura controlável e deve ajustar a frequência deste último até que perceba que a altura dele é igual à do som original.

ção de freqüência, ou dlo e freqüência, da Figura 2.9), subitamente distinguimos dois sons separados, com freqüências correspondentes a f_1 e f_2 . Nesse momento, ambas as regiões de ressonância da membrana basilar se separam suficientemente uma da outra para fornecer dois sinais distintos de altura. Entretanto, nesse limite a sensação de rudeza ainda persiste, principalmente na região grave. Apenas depois de ultrapassar uma diferença de freqüência ainda maior Δf_{CB} , chamada banda crítica, a sensação rude desaparece, e ambos os sons puros soam “suaves” e “agradáveis”. Essa transição da “aspereza” à “suavidade” na verdade é mais gradual; a banda crítica definida aqui representa apenas a separação aproximada de freqüência em que ocorre essa transição.

Todos esses resultados podem ser facilmente verificados usando-se dois “geradores de ondas senoidais” eletrônicos de freqüência variável, cujas saídas são combinadas e enviadas de forma monaural aos ouvidos por meio de fones. Mas eles também podem ser verificados, pelo menos qualitativamente, com duas flautas tocadas simultaneamente por *experts* no registro agudo. Enquanto um flautista mantém uma nota fixa (mantendo a altura realmente constante!), o outro toca a mesma nota escrita, mas desafinada (puxando ou empurrando a embocadura passo a passo). Batimentos, aspereza e discriminação de sons podem ser explorados razoavelmente bem.

A Figura 2.12 é uma tentativa (sem escala) de descrever os resultados acima de forma abrangente. As linhas cheias representam as freqüências dos sons (ou batimentos) que são realmente ouvidos. O som f_1 tem freqüência fixa, e f_2 corresponde ao som cuja freqüência é variada gradualmente (aumentada ou diminuída). O som composto corresponde a uma única sensação tonal (de freqüência intermediária), que será percebida desde que f_2 esteja dentro do limite de discriminação de freqüência de f_1 . Observe a extensão da banda crítica de cada lado do unísono ($\Delta f = 0$). Devemos enfatizar mais uma vez que essa transição da aspereza à suavidade não é tão súbita, como seria de esperar segundo a Figura 2.12, mas gradual. Há uma discussão bem mais rigorosa em Zwicker e Fastl (1990).

O limite para a discriminação de altura e a banda crítica dependem bastante da freqüência média $(f_1 + f_2)/2$ dos dois sons (e que é chamada de freqüência central do estímulo de dois sons). Eles são relativamente inde-

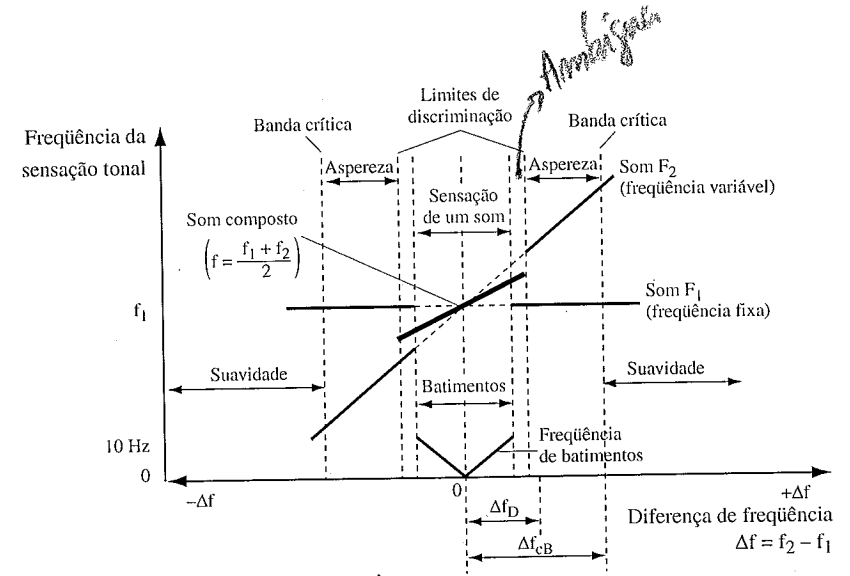


Figura 2.12 Representação esquemática das freqüências (linhas cheias) correspondentes às sensações tonais evocadas pela superposição de dois sons puros com freqüências próximas f_1 e $f_2 = f_1 + \Delta f$.

pendentes da amplitude, mas podem variar consideravelmente de indivíduo para indivíduo. A banda crítica está relacionada com vários outros fenômenos psicoacústicos, e existem muitas maneiras (e, na verdade, bem mais precisas) de defini-la experimentalmente (Seção 3.4). A Figura 2.13 mostra a dependência que a discriminação de altura Δf_D (Plomp, 1964) e a banda crítica Δf_{CB} (Zwicker, Flottorp e Stevens, 1957) têm da freqüência central dos dois sons componentes. Para referência, as diferenças de freqüência que correspondem aos intervalos musicais de um semitom, um tom inteiro e uma terça menor são apresentados com linhas pontilhadas. Por exemplo, dois sons próximos de 2000 Hz devem ter um afastamento de pelo menos 200 Hz para serem discriminados, e de mais de 300 Hz para soarem “suavemente”. Observe o interessante fato de que o limite de discriminação de freqüência é maior que um semitom⁸, e até maior que um tom inteiro nos

8. Isso pode ser uma surpresa para os músicos: eles alegarão que podem ouvir muito bem as duas notas componentes quando tocam uma segunda menor em seus instrumentos! O caso é que os resultados apresentados na Figura 2.13 só se aplicam a superposições de sons puros soando com intensidade constante. Quando um intervalo musical é tocado em instrumentos

extremos de frequências muito altas ou muito baixas. Observe também o fato de que, na parte aguda, a banda crítica localiza-se entre a diferença de frequência que corresponde a um intervalo de tom inteiro (qualificado como “dissonância”) e a que corresponde a um intervalo de terça menor (chamado de “consonância”) — *i.e.*, ela se estende mais ou menos por um terço de uma oitava. Na região de baixas frequências, existe um afastamento importante; a discriminação de frequências e a banda crítica são maiores do que uma terça menor (e até que uma terça maior). Eis por que as terças em geral não são usadas nos registros mais graves!

Compare a Figura 2.13 com a Figura 2.9: o limite para discriminação de frequência Δf_D é cerca de 30 vezes maior do que a dlo para resolução de frequências. Em outras palavras, nós podemos detectar variações de frequência bastante diminutas e um único som puro, mas é necessário uma diferença de frequências apreciável entre dois sons puros simultâneos para que possamos ouvir cada um deles separadamente⁹.

Quais são as implicações desses resultados para a teoria da audição? A existência de um limite finito para a discriminação de sons é uma indicação de que a região ativada da membrana basilar correspondente a um som puro

reais, os sons não são sons simples nem constantes, além do que um efeito estéreo está presente. Tudo isso fornece indícios adicionais para o sistema auditivo, que os utiliza eficientemente na discriminação dos sons.

- Existe um experimento equivalente que pode ser realizado com o sentido do tato para ressaltar a diferença entre “resolução” e “discriminação”. Peça a alguém que toque a pele do seu braço por cerca de um segundo, num ponto fixo e com a ponta de um lápis, enquanto você olha para outro lado. Peça então à pessoa que repita isso em outros pontos. É necessário que exista uma distância mínima finita para que você consiga afirmar que a posição de contato mudou — esta é a “dlo” para localização de uma sensação de toque, ou “resolução de toque”. Peça agora à pessoa para que use dois lápis e determine a que distância devem estar um do outro os dois pontos de contato para que você consiga identificar duas sensações de toque. Essa é a distância mínima para a “discriminação de toque”, e é consideravelmente maior do que a dlo. Tanto a resolução como a discriminação de toque variam para diferentes partes do corpo. A equivalência entre os experimentos de tato e audição não é nada casual: do ponto de vista da evolução biológica, a membrana basilar é um pedaço de tecido epitelial (pele) com uma sensibilidade de toque altamente ampliada! Essa analogia foi profusamente utilizada por Von Békésy (1960) em seus magníficos experimentos.



IMPORTANTE

Banda crítica

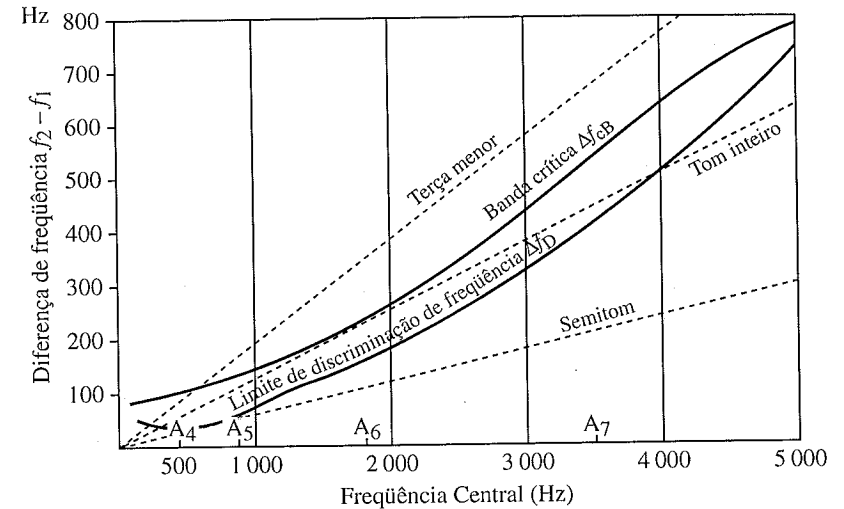


Figura 2.13 Largura de banda crítica Δf_{CB} (segundo Zwicker, Flottorp e Stevens, 1957) e o limite de discriminação de frequência Δf_D (Plomp, 1964) em função da frequência central de um estímulo de dois sons (escalas lineares). A diferença de frequência correspondente a três intervalos musicais é mostrada para comparação.

deve ter uma extensão espacial finita. Por outro lado, se ela fosse perfeitamente “nítida”, dois sons superpostos sempre seriam ouvidos como dois sons separados, contanto que suas frequências diferissem uma da outra — não importando quão pequena fosse essa diferença — e nunca surgiria nenhuma sensação de batimento. Na verdade, o fato de que a sensação de rudeza persiste mesmo além do limite de discriminação é uma indicação de que as duas regiões ativadas ainda se superpõem ou interagem em certo grau, pelo menos até que seja atingida a diferença de frequência da banda crítica. Um experimento ilustrativo seria o seguinte: enviando cada um dos sons f_1 e f_2 de maneira binaural em cada ouvido, a sensação de batimento primário ou de aspereza desaparece imediatamente, ambos os sons podem ser discriminados mesmo que a sua diferença de frequência esteja abaixo de Δf_D , e o seu efeito combinado soa “suave” sempre! No instante em que tornamos a entrada monaural, os batimentos ou a aspereza retornam. Naturalmente, o que acontece no caso binaural é que existe apenas uma região ativada em cada mem-

brana basilar, sem que haja possibilidade de os sinais se superporem na cóclea¹⁰; daí a inexistência de batimentos ou aspereza.

A esta altura, o leitor poderá se perguntar: se a região ativada da membrana basilar devido a um único som puro com uma única frequência está espacialmente distribuída, cobrindo certa extensão finita Δx ao longo da membrana, por que é que ouvimos apenas uma altura, e não uma pequena “mancha”^{*} envolvendo todas aquelas alturas que corresponderiam a diferentes posições ativadas dentro do Δx ? Lamentavelmente, teremos que adiar a resposta para seções posteriores (p. ex., Seção 3.6). Vamos apenas antecipar aqui que o chamado processo de depuração ocorre quando a atividade coletada ao longo de toda a região Δx é “focalizada” ou “afunilada” para um número bastante limitado de neurônios que respondem.

O fenômeno dos batimentos tem um papel muito importante na música. Sempre que ocorrem batimentos, eles são processados pelo cérebro, dando-nos sensações que podem ir desde o desagradável ou irritante até o agradável ou suave, dependendo da frequência dos batimentos e das circunstâncias musicais em que eles ocorrem. O som desagradável e peculiar de um instrumento desafinado em relação ao acompanhamento é causado pelos batimentos. O som horrível das cordas desafinadas de uma orquestra medíocre de escola é horrível em parte por causa dos batimentos, e o som “engraçado” de um piano de salão é causado pelos batimentos que ocorrem entre os pares ou trios de cordas desafinadas (deliberadamente) entre si nos registros médio e agudo. O fato de os batimentos desaparecerem completamente quando duas notas têm exatamente a mesma frequência ($f_2 = f_1$) assume um papel muito importante no processo de se afinar um instrumento. Se queremos ajustar a frequência de certa nota para que seja igual à frequência de certo padrão (p. ex. um diapasão), fazemos isso ouvindo os batimentos e ajustando a frequência até que os batimentos desapareçam completamente¹¹.

10. Existe, todavia, uma superposição dos sinais neurais nos estágios superiores do percurso neural, dando origem a efeitos de “segunda ordem” que serão discutidos nas Seções 2.6-8.

* Um pequeno *cluster* (N. da T.).

11. Neste capítulo, discutimos apenas o caso de batimentos entre sons puros. Como veremos mais tarde, eles ocorrem semelhantemente para os sons complexos dos instrumentos musicais reais.

A banda crítica também desempenha um papel importante na percepção musical. Vamos discutir esse conceito mais detalhadamente em futuras seções. Por enquanto, apenas ressaltamos que a banda crítica representa uma espécie de “unidade de coleta e integração de informação” na membrana basilar. O fato experimental de que a extensão de frequência da banda crítica Δf_{CB} é mais ou menos independente da amplitude sonora ou volume é um forte indício de que ela deve estar relacionada a alguma propriedade inerente da estrutura do órgão sensorial na membrana basilar, e não à forma de onda no fluido coclear. Na verdade, se convertermos a extensão de frequência Δf_{CB} vista na Figura 2.13 em extensão espacial ao longo da membrana basilar usando a Figura 2.8, podemos obter um valor quase constante de cerca de 1,2 mm para a banda crítica. Um resultado ainda mais significativo é o seguinte: verifica-se que a banda crítica é correspondida na membrana basilar por uma extensão “servida” pelo número mais ou menos constante de cerca de 1 300 células receptoras (de um total de mais ou menos 16 000 na membrana) (Zwislocki, 1965), independentemente do centro particular de frequência envolvido (*i.e.*, a posição na membrana).

Um estímulo auditivo complexo (p. ex., dois sons puros), cujos componentes estejam distribuídos por uma extensão de frequência que se encontre dentro da banda crítica, causa uma sensação subjetiva (p. ex., aspereza, neste caso) que em geral é bem diferente do caso em que a extensão excede a da banda crítica (suavidade no exemplo dos dois sons). Isso é verdade para uma grande variedade de fenômenos, tem um papel importante na percepção da qualidade do som (Seção 4.8) e fornece a base para uma teoria sobre consonâncias e dissonâncias de intervalos musicais (Seção 5.2).

2.5 OUTROS EFEITOS DE PRIMEIRA ORDEM: SONS RESULTANTES E HARMÔNICOS AURALS (Pseudo-Son Resultante)

Até agora estivemos analisando os efeitos de superposição de dois sons puros cujas frequências não diferissem muito entre si (Figura 2.12). O que

acontece com nossas sensações tonais quando a frequência da nota variável, f_2 , aumenta além da banda crítica, enquanto mantemos f_1 constante? Os efeitos decorrentes podem ser classificados em duas categorias, dependendo se eles se originam no sistema neural ou no ouvido. Nesta seção vamos nos dedicar a um fenômeno pertencente à primeira categoria, a percepção de sons resultantes. Esses sons são sensações adicionais de altura que aparecem quando dois sons puros de frequências f_1 e f_2 soam simultaneamente; eles são percebidos mais facilmente se tiverem um nível alto de intensidade. Essas sensações adicionais de altura correspondem a frequências que diferem tanto de f_1 como de f_2 , conforme se pode estabelecer facilmente por meio de experimentos de casamento de alturas ou anulação de alturas (Goldstein, 1970). Esses sons não estão presentes no estímulo sonoro original — eles surgem em consequência do que se chama de distorção não-linear do sinal acústico no ouvido.

Vamos repetir o experimento correspondente à Figura 2.12, relativo à superposição de dois sons puros, mas dessa vez aumentando consideravelmente o volume e variando a frequência f_2 lentamente para cima e para baixo entre o uníssono f_1 e a oitava, com frequência $2f_1$. Enquanto fazemos isso, prestamos atenção cuidadosamente nas sensações de altura evocadas. Obviamente, ouviremos ambos os sons, o de altura constante correspondente à frequência f_1 e o som variável f_2 . Mas, além disso, podemos ouvir claramente um ou mais sons de altura menor variando para cima e para baixo, dependendo de como variamos a frequência f_2 . Em particular, quando f_2 sobe, afastando-se de f_1 , ouvimos um som com altura crescente começando numa região bem grave. Quando f_2 desce, começando na oitava $f_2 = 2f_1$, também ouvimos um som de altura crescente começando numa região bem grave. E, prestando ainda mais atenção, é possível ouvir mais de uma nota grave ao mesmo tempo. Esses sons, que absolutamente não existem no som original, são os sons resultantes.

O som resultante que talvez seja mais facilmente identificável a um alto nível de intensidade é aquele cuja frequência é dada pela diferença das frequências componentes:

$$f_{c1} = f_2 - f_1 \quad (2.4)$$

Ele também é chamado de som de diferença. Observe que, para valores de f_2 muito próximos de f_1 , f_{c1} nada mais é que a frequência de batimento (2.3). f_{c1} deve ter pelo menos 20-30 Hz para ser ouvido como nota. Quando f_2 cresce, f_{c1} cresce também. Quando f_2 está uma oitava acima de f_1 , $f_{c1} = 2f_1 - f_1 = f_1$, ou seja, o som resultante é igual ao do componente inferior f_1 . Quando f_2 está a meio caminho entre f_1 e $2f_1$ — i.e., $f_2 = 3/2 f_1$ (um intervalo musical chamado quinta) — o som resultante tem uma frequência $f_{c1} = 3/2 f_1 - f_1 = 1/2 f_1$, correspondendo a uma altura de uma oitava abaixo de f_1 .

Os dois outros sons resultantes que são facilmente identificáveis (Plomp, 1965), mesmo com níveis baixos de intensidade dos sons originais, correspondem às frequências:

$$f_{c2} = 2f_1 - f_2 \quad (2.5)$$

$$f_{c3} = 3f_1 - 2f_2 \quad (2.6)$$

Ambos os sons decrescem em altura quando f_2 aumenta do uníssono até a quinta, e são ouvidos mais facilmente quando f_2 se encontra entre mais ou menos $1,1f_1$ e $1,3f_1$. Com maior intensidade dos sons originais, eles também podem ser percebidos perfeitamente como sensações graves perto da oitava e da quinta, respectivamente. Observe que os sons f_{c2} e f_{c1} coincidem em frequência ($= 1/2 f_1$) quando f_2 está na quinta $3/2 f_1$. Na Figura 2.14 resumimos as sensações tonais de primeira ordem evocadas pela superposição de dois sons puros de frequências f_1 e f_2 . Observe que a Figura 2.12 nada mais é que um retrato em *close* do que acontece quando a frequência f_2 está muito próxima de f_1 (área hachurada na Figura 2.14). As porções dos sons resultantes vistos com linhas cheias são as mais fáceis de serem ouvidas (a extensão real depende da intensidade).

Como são geradas essas sensações tonais extras? Conforme foi apontado acima, elas não estão presentes na vibração do som original que chega ao tímpano. Experimentos cuidadosos realizados com animais mostraram que as frequências dos sons resultantes não estão presentes nem mesmo na entrada da cóclea (membrana da janela oval, Figura 2.6); por outro lado, a

partir de medições diretas de pulso neural (Goldstein, 1970) pode-se concluir que existem na verdade regiões ativadas na membrana basilar nas posições correspondentes às frequências dos sons resultantes. Acredita-se que isso se deva a uma distorção “não-linear” do estímulo primário em forma de onda na cóclea. Pode-se mostrar matematicamente que, na verdade, quando duas oscilações harmônicas (senoidais) de frequências diferentes f_1 e f_2 entram num transdutor com resposta distorcida (não-linear), a saída conterà, além das frequências originais f_1 e f_2 , todas as combinações lineares do tipo $f_2 - f_1, 2f_1 - f_2, 3f_1 - 2f_2, f_2 + f_1, 2f_2 + f_1$ etc.¹². Experimentos mais recentes (Smootenburg, 1972) indicam que o som de diferença (2.4) e os dois outros sons resultantes (2.5 e 2.6) devem se originar em mecanismos cocleares mutuamente independentes, respectivamente. O limite de intensidade para a geração do primeiro é consideravelmente mais alto do que os dos outros e razoavelmente independente da razão de frequências f_2/f_1 .

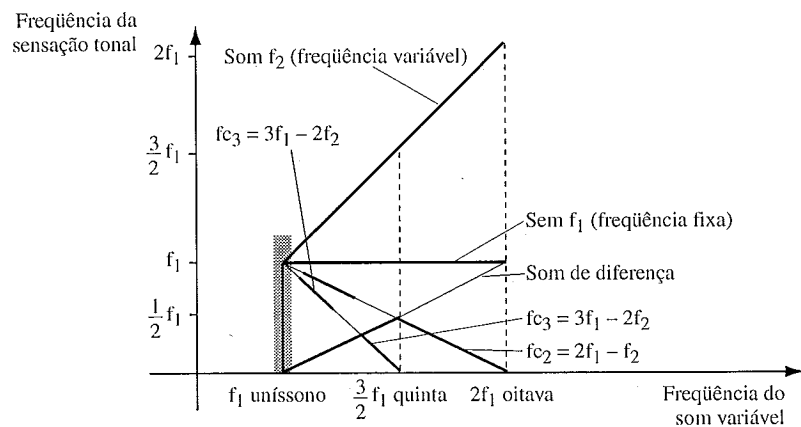


Figura 2.14 Frequências dos sons resultantes f_{c1}, f_{c2}, f_{c3} , evocadas por uma superposição de dois sons (f_1, f_2). Linhas cheias — regiões mais facilmente detectáveis de sons resultantes.

É interessante observar que, devido à distorção não-linear, mesmo um único som com frequência f_1 dará origem a sensações adicionais de altura

12. Ainda não se sabe por que as duas últimas combinações não aparecem como sensações de som.

quando o volume for muito alto. Esses sons adicionais, chamados harmônicos aurais, correspondem a frequências que são múltiplos inteiros da frequência original: $2f_1, 3f_1, 4f_1$ etc.

Embora todos os experimentos pertinentes a esta seção sejam mais apropriadamente realizados com geradores eletrônicos de som, é possível, pelo menos qualitativamente, explorar sons resultantes e harmônicos aurais usando alguns instrumentos musicais que sejam capazes de emitir sons uniformes num alto nível de intensidade. Na verdade, talvez o “instrumento” mais adequado para essa finalidade seja um apito para cães (agudo) cuja altura pode ser variada. Um simples experimento caseiro com sons resultantes pode ser realizado soprando-se dois desses apitos ao mesmo tempo — um numa altura constante e o outro com a frequência variando em torno do uníssonos — e ouvindo-se as sensações tonais na região grave. As notas resultantes não exercem um papel importante em música porque é necessária uma intensidade muito grande dos sons originais para que elas se produzam. Todavia, elas devem aparecer sempre quando se ouve *rock* nos níveis de intensidade normalmente exagerados (e fisiologicamente prejudiciais), especialmente com fones de ouvido.

Pseudo-sons resultantes podem ser facilmente gerados em órgãos eletrônicos e amplificadores e alto-falantes de baixa qualidade. Nesses casos, o que gera essas frequências parasitas é uma distorção não-linear no circuito eletrônico e no sistema mecânico do alto-falante. Em particular, o som de diferença pode ser produzido e ouvido claramente com um órgão eletrônico: aumente o volume, puxe os registros de flauta 8', toque para a frente e para trás a seqüência vista no pentagrama superior da Figura 2.15, e procure ouvir as notas graves indicadas no pentagrama inferior.

Alguns dos sons resultantes assim gerados estão desafinados por causa da afinação em temperamento equivalente do instrumento (Seção 5.5). Devemos ressaltar mais uma vez que o que é ouvido neste experimento é um som resultante falso, uma vez que a sensação de altura grave é produzida no alto-falante, e não no ouvido. Fica bem claro com esse exemplo por que os circuitos eletrônicos e os alto-falantes dos sistemas *hi-fi* e dos órgãos eletrônicos devem ter uma resposta “bastante” linear.

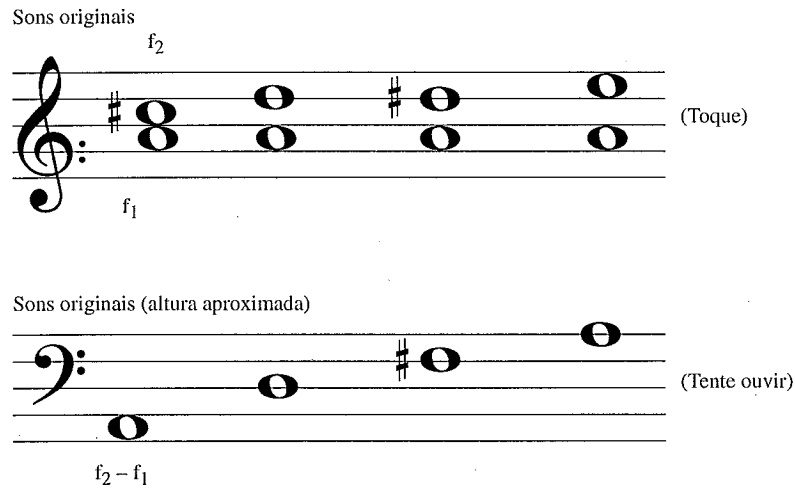


Figura 2.15

2.6 EFEITOS DE SEGUNDA ORDEM: BATIMENTOS DE CONSONÂNCIAS DESAFINADAS

Vamos agora repetir o experimento da seção anterior com dois sons puros gerados eletronicamente, mas dessa vez ignorando as possíveis sensações de sons resultantes. Fazemos soar ambos os sons a um baixo nível de intensidade no mesmo ouvido; o som f_1 é mantido a uma frequência constante, enquanto f_2 pode ser variado com liberdade. As amplitudes de ambos os sons são mantidas constantes durante o experimento. Quando variamos f_2 lentamente para cima, observamos algo peculiar ao passarmos nas vizinhanças da oitava $f_2 = 2f_1$: uma sensação distinta de batimento, bem diferente da dos batimentos de primeira ordem perto do uníssono, mas claramente perceptível. Quando f_2 é exatamente igual a $2f_1$ essa sensação de batimento desaparece. Ela reaparece tão logo nós desafinemos a oitava, *i.e.*, quando f_2 ficar $f_2 = 2f_1 + \varepsilon$, onde ε (*épsilon*) representa apenas alguns hertz. A frequência de batimentos resulta igual a ε . É difícil descrever o que está batendo. Muita gente descreve o fenômeno como um bati-

mento da “qualidade tonal”. Chamamos isso de batimentos de segunda ordem; alguns preferem chamá-los de “batimentos subjetivos”. Eles são resultantes de processamento neural.

É muito instrutivo observar os padrões de vibração no osciloscópio enquanto se ouvem os batimentos de segunda ordem. Observamos que esse padrão muda em exato sincronismo com as sensações de batimento. Obviamente, nosso sistema auditivo deve ter de alguma forma a capacidade de detectar essas variações na forma do padrão de vibração. A Figura 8.2.16 mostra vários padrões de vibração correspondentes à superposição de um som fundamental de frequência f_1 e de sua oitava $f_2 = 2f_1$ (com amplitude menor), para quatro valores diferentes da diferença de fase. Desde que a oitava esteja perfeitamente afinada, a diferença de fase permanece constante e a imagem no osciloscópio ficará estática, todas as quatro superposições soam idênticas — nosso ouvido não distingue um caso do outro. Mas quando tiramos f_2 ligeiramente de afinação: $f_2 = 2f_1 + \varepsilon$, a relação mútua de fase mudará continuamente com o tempo, e o padrão de vibração resultante sofrerá gradualmente um deslocamento de uma das formas vistas na Figura 2.16 para a seguinte. Pode-se mostrar matematicamente que esse ciclo de padrão de vibração mutante se repete com uma frequência ε , que é a quantidade em que a nota superior se desafina em relação à oitava. Isso obviamente significa que o ouvido é sensível a uma diferença de fase lentamente mutante entre dois sons¹³. Uma afirmação equivalente é: o sistema auditivo é capaz de detectar variações cíclicas na forma dos padrões de vibração. Observe cuidadosamente que não há uma variação real de amplitude de um padrão para outro na Figura 2.16 — exatamente o contrário do que acontece com os batimentos de primeira ordem, que são variações cíclicas da amplitude dos padrões de vibração (Figura 2.11). A Figura 2.17 mostra duas fotos reais de osciloscópio confrontando batimentos de primeira ordem nas proximidades do uníssono e batimentos de segunda ordem para uma oitava desafinada. Observe a

13. Variações súbitas de fase (p. ex., apresentações dos estímulos de oitava vistos na Figura 2.16, alternadas com um estímulo de oitava de referência) também são detectadas. O grau de detectabilidade tem um máximo para uma diferença de fase de 180° em relação ao sinal de referência (Raiford e Schubert, 1971).

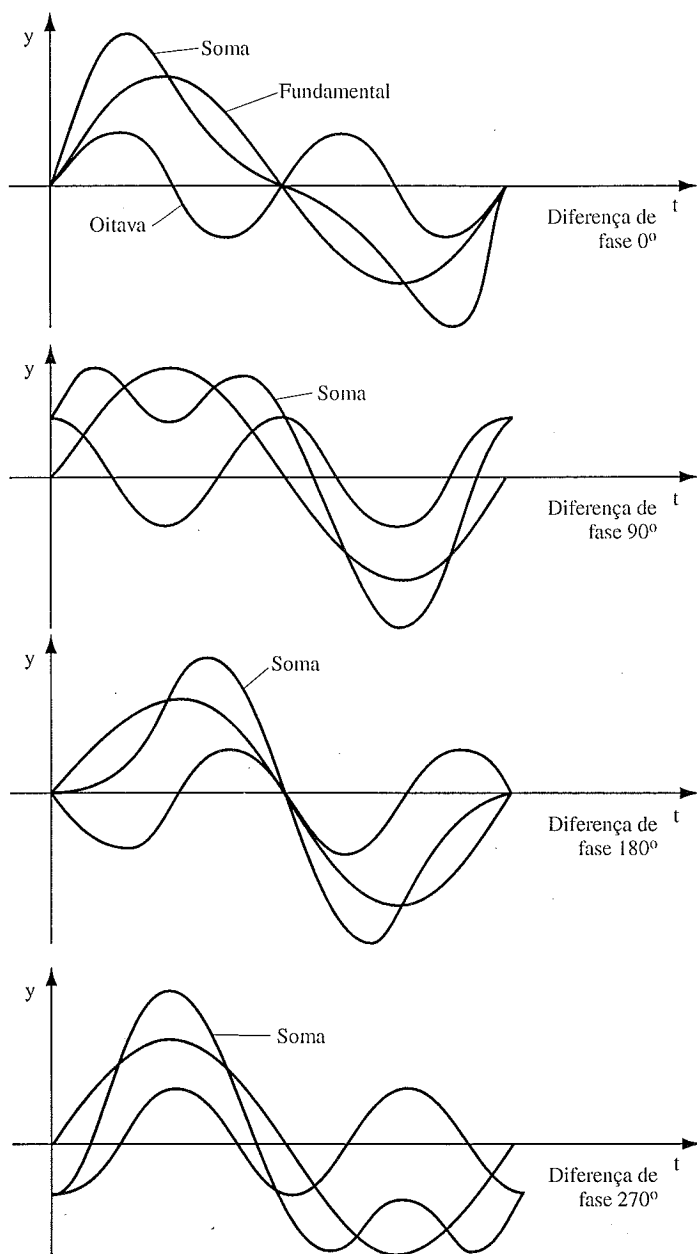
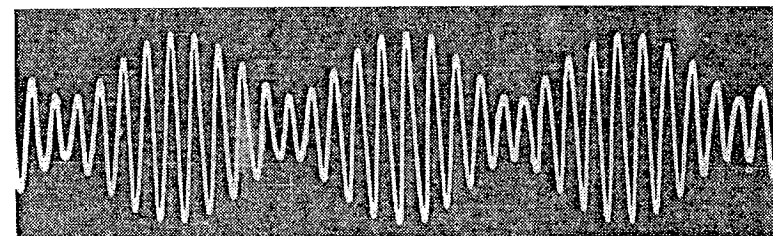
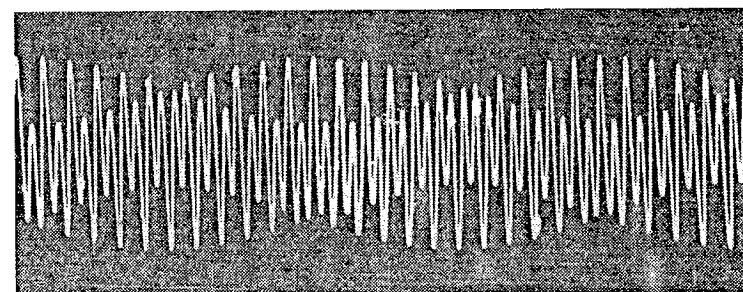


Figura 2.16



(a)



(b)

Figura 2.17 Comparação dos batimentos de primeira e segunda ordem. *Acima*: batimentos de primeira ordem (unísono desafinado); modulação de amplitude sem mudança na forma do padrão de vibração. *Abaixo*: batimentos de segunda ordem (oitava de desafinada); modulação do padrão sem variação na amplitude total.

ondulação na amplitude do primeiro e a ondulação no padrão de vibração do segundo. É importante observar que a sensação de batimento de segunda ordem só aparece na região de baixa frequência do estímulo inicial de dois sons. Quando f_1 (e f_2) excede uns 1 500 Hz, os batimentos de segunda ordem não podem mais ser percebidos (Plomp, 1967).

Vamos agora voltar ao nosso tratamento experimental e explorar toda a gama de frequências entre o unísono e a oitava. Descobrimos que existem outros pares de valores para f_2 e f_1 , *i.e.*, outros intervalos musicais, em cujas vizinhanças aparecem sensações de batimento, embora muito menos perceptíveis do que na oitava. Dois desses “focos de batimento”, como podemos chamá-los, podem ser encontrados centrados nas frequências $f_2 = 3/2 f_1$ e $f_2 = 4/3 f_1$, o que corresponde aos interva-

los musicais de quinta e quarta, respectivamente. Observando novamente o padrão de vibração no osciloscópio ao mesmo tempo que o ouvimos, verificamos que, para uma quinta desafinada ($f_2 = 3/2f_1 + \epsilon$) e para uma quarta desafinada ($f = 4/3f + \epsilon$), a forma do padrão de vibração não é estática (como acontece com uma quinta ou quarta justas, *i.e.*, $\epsilon = 0$); essa forma varia periodicamente (não em amplitude). Os batimentos de segunda ordem da quinta são mais “rápidos” do que os da oitava (para a quinta, a frequência de batimento é $f_B = 2\epsilon$, e para a quarta, $f_B = 3\epsilon$). Essa não é a única razão pela qual eles são mais difíceis de serem percebidos: o próprio padrão de vibração vai-se tornando mais e mais complicado (*i.e.*, afasta-se mais e mais do padrão de um movimento harmônico simples) à medida que nos afastamos da oitava em direção a quinta e à quarta (Figura 2.16). Quanto mais complexo o padrão de vibração, mais difícil será para o sistema auditivo detectar a sua variação gradual. Para uma discussão detalhada sobre batimentos de consonâncias desafinadas, veja Plomp (1976).

Existe uma relação ótima entre as intensidades dos sons componentes para a qual os batimentos de segunda ordem estão mais pronunciados e que sempre coloca o som de maior altura numa intensidade menor (Plomp, 1967). Finalmente, é importante observar que os batimentos de segunda ordem também são percebidos quando os sons componentes são enviados separadamente, um para cada ouvido.

Os batimentos de segunda ordem de consonâncias desafinadas de sons puros não têm um papel importante na música (principalmente porque os sons puros não o têm). Mas eles são um ingrediente crucial para a compreensão do mecanismo de processamento do som musical (Seção 2.8).

2.7 RASTREAMENTO DE FUNDAMENTAIS

Vamos agora introduzir outra série de experimentos psicoacústicos que têm sido de crucial importância para as teorias de percepção acústica. Vamos considerar dois sons puros, a uma distância de quinta justa, com fre-

quências f_1 e $f_2 = 3/2f_1$. A Figura 2.18 mostra a vibração resultante (soma) para uma relação particular de fase. Observe que o padrão repete a forma exatamente após um tempo τ_0 , que é o dobro de τ_1 , o período do som mais grave. Isso significa que a frequência de repetição $f_0 = 1/\tau_0$ do padrão de vibração de uma quinta é igual à metade da frequência da nota mais grave:

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

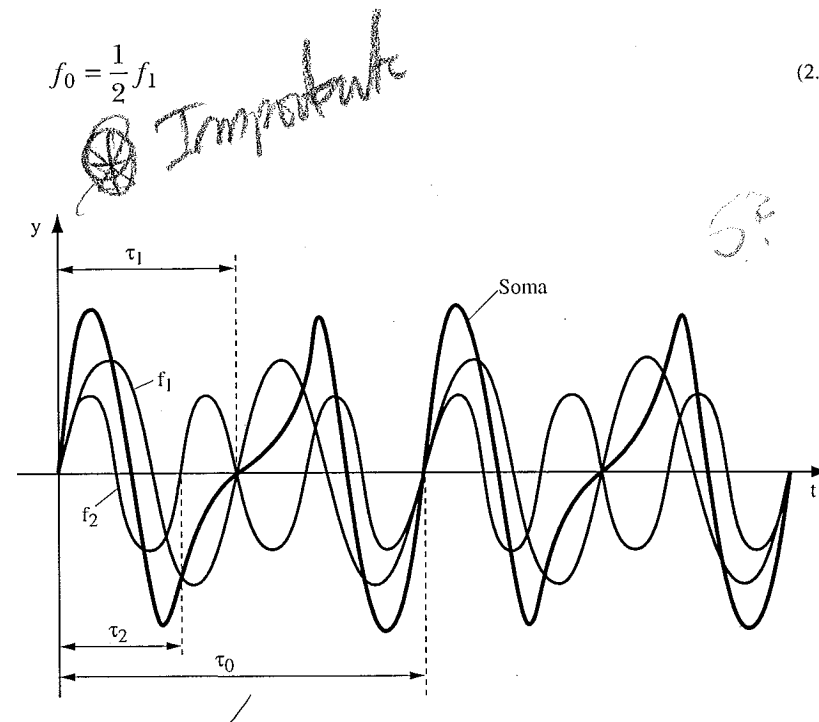


Figura 2.18

Chamamos essa frequência de repetição de “frequência fundamental” de vibração. Nesse caso, ela está uma oitava abaixo de f_1 . Se considerarmos duas notas formando uma quarta ($f_2 = 4/3f_1$), podemos traçar o padrão de vibração resultante, da mesma forma que fizemos para a quinta. A velocidade de repetição resultante será

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

i.e. uma décima segunda abaixo da nota mais grave. Para uma terça maior ($f_2 = 5/4 f_1$), a velocidade de repetição está duas oitavas abaixo de f_1 :

$$f_0 = \frac{1}{4} f_1 \quad (2.7c)$$

Nosso sistema auditivo se revela sensível a essas velocidades de repetição. Na verdade, realizaram-se experimentos cuidadosos nos quais indivíduos eram expostos a seqüências curtas de estímulos constituídos por pares de sons puros soando simultaneamente, formando intervalos de quinta, quarta, terça etc. (Houtsma e Goldstein, 1972). Pediu-se a esses indivíduos que identificassem uma única altura básica da “melodia”. A maior parte deles realmente isolou uma freqüência única que coincide com as freqüências dadas pelas relações (2.7a), (2.7b) ou (2.7c), respectivamente!¹⁴ É importante ressaltar que esse experimento de identificação de alturas exige que os complexos de duas notas sejam apresentados numa seqüência temporal, ou melodia. (Quando nosso sistema auditivo se encontra diante de pares de sons puros soando indefinidamente, ele não é capaz de buscar uma sensação de altura única; ele rapidamente redireciona sua atenção para discriminar as alturas espectrais dos dois componentes puros, como explica a Seção 2.4.).

Observe que as velocidades de repetição (2.7a-c) dos complexos de dois sons acima são idênticas às freqüências dos sons de diferença correspondentes (p. ex., veja o quarto, o segundo e o primeiro casos da Figura 2.15). Contudo, os experimentos mostram que a detecção de velocidades de repetição pode ocorrer mesmo que as intensidades dos dois sons sejam mínimas, bem abaixo do limite para a produção de sons resultantes — assim, está excluída a possibilidade de um som de diferença (que, nesse caso, coincidentemente, tem a mesma freqüência $1/2 f_1$) (Plomp, 1967). Na verdade, esse efeito tem sido usado em música por muitos séculos (e erradamente atribuído a um efeito de sons resultantes). Por exemplo,

14. Observe que esse experimento *deve* ser realizado com pares de sons senoidais, gerados eletronicamente — ele não dará certo se usarmos um piano ou outro instrumento musical. Veja adiante as observações sobre órgãos.

desde o final do século XVI, muitos órgãos incluíam um registro (a “quinta de $5 \frac{1}{3}$ pés”) formado por tubos que soavam uma quinta acima da nota escrita e que realmente era tocada. O propósito é simular ou reforçar a base uma oitava abaixo (2.7a) da nota escrita (*i.e.*, para reforçar o som de $16'$ do órgão). De uso ainda mais antigo é a quinta de $10 \frac{2}{3}$ pés nos pedais, a qual, em combinação com registros de $16'$, estimula ou reforça o baixo de $32'$ (duas oitavas abaixo da nota escrita).

O som de freqüência f_0 (2.7) não está presente como componente original. Esse som também é chamado de “fundamental ausente” (por razões que se tornarão evidentes a seguir); a sensação de altura correspondente é chamada de altura de periodicidade ou altura subjetiva, som residual ou altura virtual. Devemos diferenciar claramente essa sensação de altura da altura primária ou espectral de cada um dos dois sons puros individuais. Os experimentos têm mostrado que, para níveis normais de volume, essa freqüência também não está presente nas oscilações do fluido coclear (enquanto os sons resultantes estão). Na verdade, a região da membrana basilar que corresponde à freqüência f_0 (Figura 2.8) pode estar saturada (mascarada) com uma faixa de ruído (som com uma quantidade infinita de freqüências componentes estendendo-se por uma certa faixa), de modo que qualquer excitação adicional dessa região passaria despercebida — mas mesmo assim a fundamental ausente ainda seria ouvida (Small, 1970). Podemos introduzir uma nota extra e que esteja ligeiramente desafinada em relação a f_0 ; deverão surgir batimentos de primeira ordem se o som da fundamental ausente f_0 realmente existir na cóclea — nenhuma sensação de batimento será sentida, podemos adiantar. Um efeito mais drástico é que a fundamental ausente é percebida mesmo que os dois sons componentes sejam ouvidos binauralmente em cada ouvido (Houtsma e Goldstein, 1971, 1972). Tudo isso é uma indicação de que a fundamental ausente, ou altura de periodicidade, deve ser o resultado de processamento neural. A detecção de altura subjetiva, *i.e.*, a capacidade do nosso sistema auditivo de identificar a velocidade de repetição do padrão de vibração não analisado de entrada, só funciona na faixa de baixas freqüências (musicalmente mais importantes), abaixo de uns 1 500 Hz. Quanto mais complexo for o padrão de vibração, *i.e.*, quanto menor for o intervalo entre os sons componentes, maior será a dificuldade

encontrada pelo sistema auditivo para identificar a velocidade de repetição, *i.e.*, maior dificuldade ele encontrará para perceber a altura subjetiva.

Vamos agora inverter as relações 2.7-9 e descobrir quais os pares de frequências de sons puros que dão origem à mesma velocidade de repetição ou frequência fundamental f_0 . Obtemos:

$$\underbrace{2f_0 \text{ e } 3f_0}_{\text{quinta}} \quad \underbrace{3f_0 \text{ e } 4f_0}_{\text{quarta}} \quad \underbrace{4f_0 \text{ e } 5f_0}_{\text{terça maior}} \quad \underbrace{5f_0 \text{ e } 6f_0}_{\text{terça menor}}$$

Em outras palavras, se f_0 corresponder à nota escrita no pentagrama inferior da Figura 2.19, os intervalos musicais escritos no pentagrama superior proporcionarão essa mesma nota como sensação e altura subjetiva. É importante pensar que as notas da Figura 2.19 representam sons puros com apenas uma frequência cada um, e não como notas de instrumentos musicais reais.

Os componentes individuais de frequências $2f_0, 3f_0, 4f_0, 5f_0, \dots$ são chamados harmônicos superiores da frequência fundamental f_0 . As frequências dos harmônicos superiores são múltiplos inteiros da frequência fundamental. Quaisquer dois sons sucessivos da série de harmônicos superiores formam um par com a mesma velocidade de repetição ou frequência fundamental f_0 . Portanto, se todos os harmônicos superiores soarem juntos, eles produzirão uma única sensação de altura subjetiva correspondente a f_0 — mesmo que essa frequência esteja totalmente ausente no estímulo de múltiplos sons! Essa é a razão pela qual a frequência fundamental correspondente à altura subjetiva, dada pela velocidade de repetição, também é chamada de “fundamental ausente”, e por que a percepção dessa velocidade de repetição é chamada de rastreamento de fundamental. Observe ainda uma interessante propriedade desse conjunto de sons puros com frequências $2f_0, 3f_0, 4f_0, \dots, nf_0, \dots$: dentre uma variedade infinita de superposições imagináveis de sons puros, esta é a única cujos componentes, tomados em pares de frequências contíguas, proporcionam sempre a mesma velocidade de repetição. Por outro lado, esse é o motivo pelo qual todo som periódico com um padrão de vibração complexo mais repetitivo (com velocidade de repetição f_0) é formado por uma superposição de sons puros com frequências nf_0 ($n =$ número inteiro) (ver 4.2).

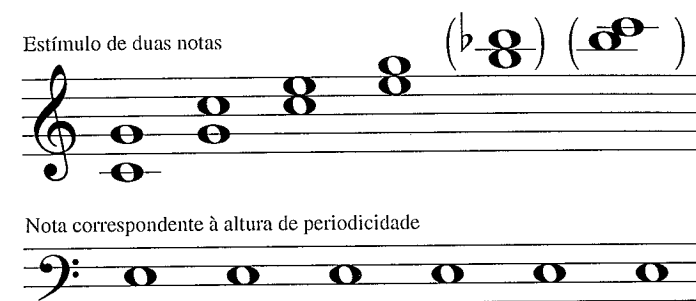


Figura 2.19 Estímulos de duas notas (pentagrama superior) que originam a mesma altura de periodicidade (pentagrama inferior). O si-bemol marcado entre parênteses deve ser afinado um pouco mais baixo para gerar um dó como altura de periodicidade.

Os experimentos psicoacústicos mencionados acima, usando complexos de dois sons, foram ampliados para incluir melodias ou seqüências compostas por complexos de vários sons, começando no n -ésimo harmônico (*i.e.*, superposições de sons puros com frequências $nf_0, (n+1)f_0, (n+2)f_0$ etc.). Embora o som de frequência fundamental não esteja presente, a altura subjetiva atribuída a esses complexos tonais sempre corresponde a f_0 . Aliás, quanto mais harmônicos forem incluídos, mais claramente se ouvirá a altura de periodicidade (a menos que a seqüência de harmônicos de partida seja alta, *i.e.*, n seja grande). Os pares de harmônicos vizinhos mais cruciais na determinação da altura de periodicidade são aqueles perto de $n = 4$ (Ritsma, 1967). Como acontece de os sons musicais “reais” serem constituídos por uma superposição de harmônicos (Capítulo 4), o rastreamento de fundamental é o mecanismo auditivo que nos possibilita atribuir uma única sensação de altura ao som complexo de um instrumento musical¹⁵.

É importante compreender a plena implicação do rastreamento de fundamental para a teoria da audição. Uma breve análise dos clássicos experi-

15. Talvez o exemplo mais convincente de rastreamento de fundamentais de sons complexos seja o fato de que ainda somos capazes de perceber a altura correta das notas mais graves que ouvimos num rádio portátil barato, embora todas as frequências abaixo de 100-150 Hz sejam cortadas pela circuitagem inadequada e pelo alto-falante!

mentos de casamento de alturas de Smoorenburg (1970) servirá a esse propósito¹⁶. Considere um estímulo de dois sons, de curta duração, cujas frequências componentes f_a e f_b apresentem entre si uma diferença de $\Delta f = f_b - f_a$. Quando esse estímulo é apresentado num certo contexto específico, cerca de metade dos indivíduos sob teste percebe uma altura “residual” subjetiva baixa, claramente identificável (os outros parecem ser capazes de apenas ouvir uma ou ambas as alturas do estímulo original). Os experimentos mostram que, se f_a e f_b correspondem a dois harmônicos vizinhos de um som complexo (p. ex., de ordem n e $n + 1$), a altura subjetiva, quando percebida, é a da fundamental ausente $f_1 (= f_a/n = \Delta f)$. Por exemplo, se $f_a = 800$ Hz e $f_b = 1\ 000$ Hz ($n = 4$ e $\Delta f = 200$ Hz), a altura residual ouvida é a de uma nota de 200 Hz. A Figura 2.20a mostra o padrão de vibração do estímulo de dois sons, remanescente do fenômeno dos batimentos de primeira ordem (Seção 2.4). Nesse caso, porém, a modulação de amplitude (variação da envoltória da curva) é muito rápida (200 vezes por segundo) e não é percebida como batimento. Ao contrário, o que é percebido (por 50% dos indivíduos em teste) é uma altura correspondente à velocidade de repetição do padrão de vibração, que é exatamente 200 Hz. O período de repetição correspondente é $\tau_1 = 1/\Delta f = 1/(f_b - f_a)$ (o período da fundamental ausente), indicado na figura (também marcamos o outro período, bem mais curto, na estrutura fina temporal do padrão de vibração, que corresponde à chamada frequência central do estímulo de dois sons, $f_c = (f_a + f_b)/2$). Obviamente, nosso sistema auditivo extrai informação da variação periódica do padrão de vibração, assim como ocorre quando batimentos (lentos) de consonâncias desafinadas são percebidos (Figura 2.17). Para um outro par de harmônicos vizinhos, como, por exemplo $f_a = 2\ 000$ Hz e $f_b = 2\ 200$ Hz ($n = 10$), a mesma altura subjetiva é percebida (com dificuldade crescente à medida que n aumenta — Houtsma, 1970); o padrão de vibração tem exatamente a mesma envoltória que a da Figura 2.20a, mas, como a frequência central é agora muito mais alta, a curva de vibração tem mais oscilações (10, a ordem harmônica de f_a) dentro de um período de repetição.

16. Infelizmente, esses interessantes experimentos não podem ser facilmente demonstrados, mesmo num laboratório para o ensino de física bem equipado!

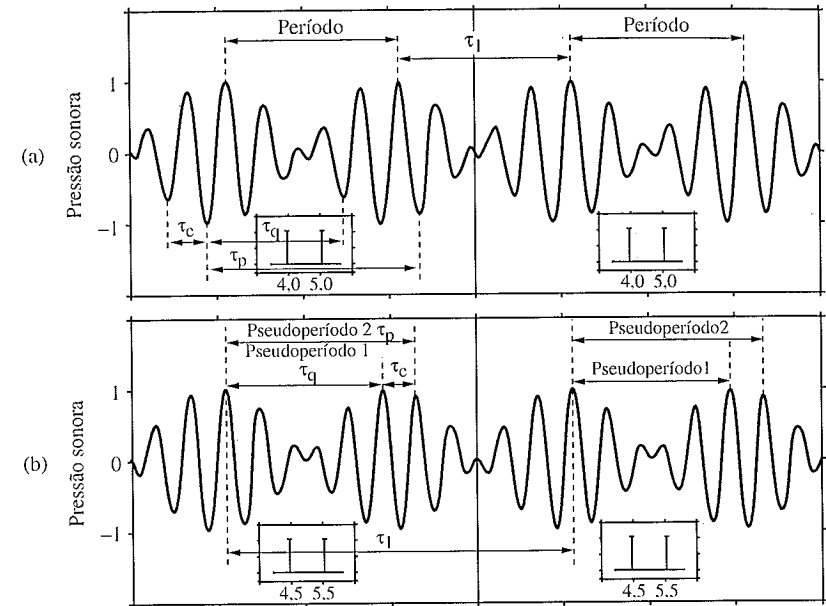


Figura 2.20 Padrão de vibração de dois sons puros simultâneos.

(a) Os sons são harmônicos vizinhos ($n = 4$).

(b) Os sons têm a mesma diferença de frequência do caso (a), mas não são harmônicos vizinhos.

τ_1 = período de repetição exata do padrão de vibração

τ_c = frequência central

τ_p, τ_q = pseudoperíodos (ver o texto)

Uma complicação interessante surge quando o par f_a, f_b não corresponde a dois harmônicos vizinhos de alguma fundamental. Consideremos um estímulo de dois sons em que $f_a = 900$ Hz e $f_b = 1\ 100$ Hz. Não existe nota musical que apresente esses harmônicos como vizinhos; eles são antes o 9º e 11º harmônicos de uma nota com frequência fundamental $f_1 = 100$ Hz. Será essa a altura percebida? Não! A percepção se revela ambígua: duas alturas possíveis podem casar (dependendo do contexto em que o estímulo é apresentado), correspondendo a cerca de 180 Hz ou 220 Hz! A Figura 2.20b mostra o padrão de vibração nesse caso. Primeiro, note que o padrão

de vibração apresenta uma envoltória com o mesmo período de modulação que o da Figura 2.20a (*i.e.*, correspondendo a uma frequência de 200 Hz). Segundo, examine cuidadosamente o padrão dos picos e vales e observe que a estrutura fina *exata* se repete com um período *duas vezes* mais longo (*i.e.*, correspondendo a 100 Hz). A velocidade de repetição exata é, assim, 100 Hz — mas nem esta nem a velocidade de modulação de 200 Hz são percebidas. Na verdade, o que ocorre é que as duas alturas possíveis ouvidas correspondem exatamente aos dois “pseudoperíodos” τ_p e τ_q indicados na figura! Além disso, os experimentos de Smoorenburg mostram que, mesmo no primeiro exemplo (quando o estímulo consiste em dois harmônicos vizinhos), podem-se ouvir outras alturas ambíguas correspondendo a “pseudoperíodos” definidos pelos intervalos de tempo entre o pico central e os picos secundários no período de modulação subsequente (Figura 2.20a). Tudo isso é uma indicação de que um processo de extração de alturas muito mais sofisticado do que a detecção da velocidade de repetição ou de modulação de amplitude está atuando: o mecanismo de audição é capaz de se fixar em aspectos muito detalhados do padrão de vibração.

Contudo, existe uma alternativa a esse mecanismo de extração de alturas com “base temporal” à qual podemos recorrer para explicar os resultados acima. Trata-se de um processo com “base espectral”, em que o padrão *espacial* de excitação gerado ao longo da membrana basilar por um estímulo de dois sons é analisado em detalhes pelo processador de alturas e casado o mais próximo possível a configurações “familiares” (p. ex., a posição das regiões de ressonância de harmônicos vizinhos). Sempre que um casamento é conseguido, uma sensação de altura é gerada; como mais de um casamento “aceitável” é possível, o resultado poderá ser alturas ambíguas. Pode-se provar matematicamente que esse processo leva aos mesmos resultados quantitativos que o mecanismo de análise de sugestão temporal acima. Abordaremos novamente esse assunto na Seção 4.10 e no Apêndice II.

Finalmente, a disparidade entre os indivíduos no que se refere à habilidade de perceber alturas residuais subjetivas de estímulos de dois sons levou alguns psicoacústicos a lançarem um alerta quanto à interpretação dos experimentos de Smoorenburg. Todavia, a consistência dos resultados

quantitativos (verificada por vários grupos de pesquisa independentes) para os indivíduos que de fato ouvem a altura residual é tão marcante que tal disparidade pode ser apenas uma indicação de uma diferença na estratégia auditiva, sem maiores conseqüências para as conclusões tiradas desses experimentos sobre o mecanismo de extração de alturas em si. Para uma excelente resenha histórica dos mais importantes experimentos de percepção de altura, veja Plomp (1976).

2.8 CODIFICAÇÃO AUDITIVA NO SISTEMA NERVOSO PERIFÉRICO

A descoberta de efeitos de segunda ordem no processamento auditivo, tais como a percepção de batimentos de consonâncias desafinadas e o rastreamento de fundamental, teve um grande impacto na teoria da audição. Por um lado, a percepção de batimentos de consonâncias desafinadas (Seção 2.6) é uma indicação de que o sistema auditivo de alguma forma obtém e utiliza informação sobre a estrutura temporal detalhada do padrão de vibração acústica. Por outro lado, o rastreamento de fundamental (Seção 2.7) poderia, a princípio, implicar duas alternativas:

1. Um mecanismo realizando uma análise detalhada do padrão de vibração *temporal*, com a instrução de proceder ao ajuste (zerar) sobre detalhes repetitivos cuja frequência leva a uma sensação de altura (Figura 2.20); ou
 2. Um mecanismo que analisa a informação em detalhes do padrão de excitação *espacial* induzido ao longo da membrana basilar, com a instrução de fornecer uma única sensação de altura se esse padrão corresponder, pelo menos em parte, à excitação característica gerada por uma nota musical¹⁷.
17. Comparativamente, um processo equivalente de casamento espacial nos permite interpretar corretamente os seguintes símbolos incompletos como letras do alfabeto: $A \curvearrowright J H$. Mas um símbolo como $A \curvearrowright$ pode levar a diferentes interpretações, dependendo do contexto em que ele aparece: $A \curvearrowright$ métrica ou $A \curvearrowright$ ungría!

Devemos antecipar que a segunda alternativa funcionaria melhor na região dos harmônicos inferiores (valores baixos de n), em que os máximos de excitação espacial são mais distintos. Qualquer alternativa implica que a informação acústica detalhada, decodificada na periferia, deve ser analisada num nível superior do sistema nervoso central.

Para compreendermos os mecanismos subjacentes que foram propostos para explicar esses fenômenos psicoacústicos, é necessário, primeiro, analisar alguns aspectos operacionais do sistema nervoso auditivo.

A unidade fundamental de processamento e transmissão é a célula nervosa ou neurônio. Um exemplo de neurônio é visto na Figura 2.21. Podemos distinguir o corpo celular, uma série de elementos ramificados chamados dendritos e uma fibra longada, o axônio, que também pode se abrir em múltiplas ramificações. Os dendritos e o corpo celular são os receptores dos sinais neurais de entrada, e o axônio é o transmissor, passando-os para os outros neurônios. Esses sinais neurais consistem em impulsos elétricos (variações de voltagem elétrica) da ordem de dezenas de milivolts, que se originam no fluxo de átomos eletricamente carregados (íons) pela membrana da célula. Esses impulsos podem ser registrados pela implantação de microeletrodos dentro do neurônio (um procedimento que não afeta a função normal da célula). No axônio, todos os impulsos, chamados de potenciais de ação, têm aproximadamente a mesma forma e duração (dezenas de milissegundos), e se propagam do corpo da célula para as terminações do axônio. Uma mensagem neural “integrada” é dada pela velocidade e distribuição no tempo com que ocorrem os disparos de impulsos individuais ao longo do axônio.

O axônio está “ligado” aos dendritos ou corpos celulares de outros neurônios. Os pontos de contato são chamados de sinapses. Um determinado axônio pode estar em contato sináptico com muitas outras células; por outro lado, uma determinada célula pode estar conectada aos axônios provenientes de centenas de outras células¹⁸. Um impulso que chega a um contato sináptico causa a liberação de uma substância química pela

18. As exceções são os neurônios unipolares cujos dendritos únicos formam as fibras nervosas aferentes.

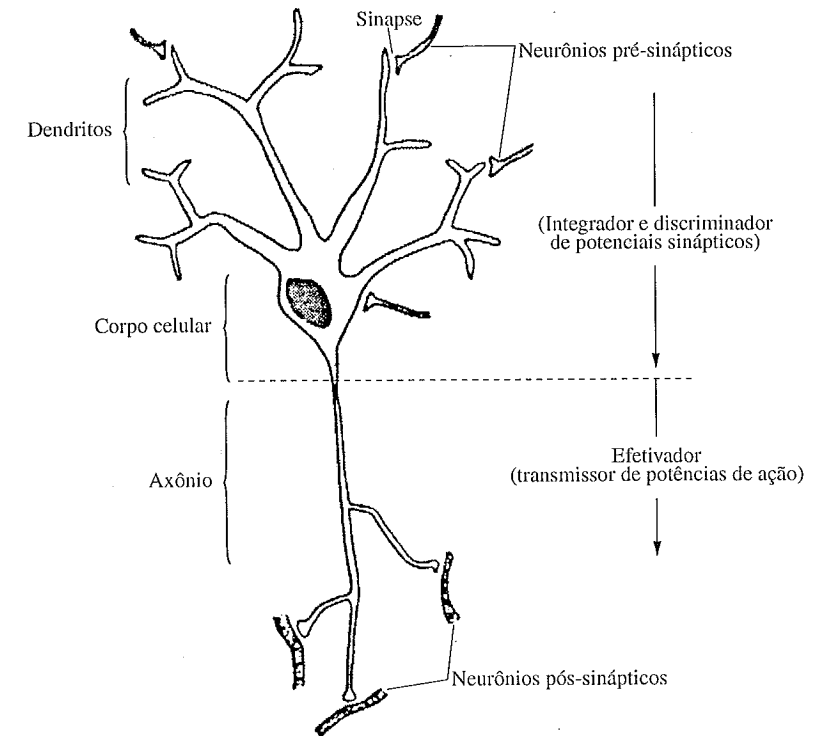


Figura 2.21

célula pré-sináptica no espaço entre as membranas celulares de ambas as células (fenda sináptica). A presença dessa substância dispara um impulso elétrico na célula pós-sináptica, o “potencial pós-sináptico”. Os potenciais sinápticos são sinais analógicos com formas variadas e duração maior do que os potenciais digitais de ação padronizados, que se propagam pelo axônio. Existem dois tipos distintos de sinapse, a excitadora e a inibidora, evocando potenciais sinápticos de polaridades opostas. Se, num certo instante, um neurônio recebe uma quantidade de estímulos excitadores que exceda o número de sinais inibidores que chegam simultaneamente em um certo valor-limite, o neurônio responderá, disparando um impulso no seu axônio. Caso contrário, ele permanecerá em silêncio. Concluímos que os dendritos e suas sinapses funcionam como

o sistema de coleta e integração de informação do neurônio — representando assim a unidade fundamental de processamento de informação do sistema nervoso.

É importante observar que o que irá determinar se um sinal de saída será ou não disparado por um neurônio é a distribuição espacial e temporal dos sinais de entrada provenientes dos neurônios pré-sinápticos. Por outro lado, muitos neurônios disparam potenciais de ação espontaneamente a uma razão característica. Um neurônio específico pode transmitir apenas ordens excitadoras ou inibidoras aos outros neurônios. Quando um neurônio inibidor dispara um pulso para outro neurônio inibidor, ele cancela o efeito inibidor deste último. Existem indícios de que a resposta dos neurônios ligados às células capilares da fileira interna é excitadora, sendo inibidora a da fileira externa (Sokolich e Zwislocki, 1974).

Existe um atraso de tempo característico (normalmente menos que um milissegundo) entre a chegada de um impulso numa sinapse e a formação da resposta na célula pós-sináptica. Esse atraso de tempo possibilita que a ativação neural disparada por um único estímulo externo possa subsistir ou “reverberar” por um tempo considerável, quando ela se propaga por uma série de milhares de passos sinápticos sucessivos no tecido cerebral. Isso pode ser um processo-chave na elaboração de padrões dependentes do tempo na atividade neural (p. ex., Seção 4.9). Depois de cada ativação, um neurônio tem um período refratário, durante o qual ele não pode ser reexcitado, ou durante o qual o seu limite de disparo é elevado. Um único neurônio é apenas um componente “microscópico” de um conglomerado de cerca de 10 bilhões de neurônios que existem no cérebro humano. É no sistema de ligação entre esses 10 bilhões de neurônios que se encontra escondido o “segredo” da inteligência, do comportamento, dos sentimentos e os medos dos seres humanos. Num sentido amplamente macroscópico, o desenvolvimento desse sistema de ligação é comandado pelo código genético da espécie.

Entretanto, numa parte importante do sistema neural, especialmente no córtex cerebral, a distribuição particular de limites de disparo é o resultado da ação repetitiva de padrões de estimulação — em outras palavras, da experiência e do aprendizado (ver Seção 4.10).

Quando um dado neurônio do sistema neural pertencente a um certo órgão de sentido é observado com um microeletrodo, encontra-se em geral uma correlação entre a frequência com que ocorrem os disparos e as magnitudes de certos parâmetros físicos do estímulo original. Normalmente, os disparos individuais não ocorrem espaçados regularmente no tempo — eles aparecem distribuídos estatisticamente de forma aleatória. É a sua frequência média ou o fato de que um neurônio está disparando que importa (todavia, veja mais adiante). Verifica-se que um neurônio pode disparar espontaneamente em frequências médias de até dezenas de hertz. Para esses neurônios, é a variação na frequência dos disparos (aumento ou inibição) que constitui a mensagem neural elementar. Um estímulo constante e persistente provocará em geral uma frequência de disparo que decrescerá gradualmente com o tempo até nivelar-se num certo patamar inferior. Esse fenômeno é chamado de adaptação. Por outro lado, sabe-se que alguns neurônios respondem apenas a variações dos estímulos sensoriais no tempo, e outros a todo um complexo de configurações espaciais particulares ou padrões temporais do estímulo (detectores de detalhes). Numa regra geral, quanto mais avançarmos no caminho neural desde os receptores até o córtex cerebral, mais complexos e elaborados serão os detalhes dos estímulos aos quais um certo neurônio responderá (ver Figura 2.26).

Estamos agora numa posição melhor para discutir como o sistema neural pode coletar e codificar informação sobre o padrão de vibração. Quando o sinal acústico de um único som puro de certa frequência chega ao ouvido, as oscilações da membrana basilar estimulam as células capilares que se encontram na região de ressonância correspondente àquela frequência (Seção 2.3). Nos seres humanos, essas unidades sensoriais estão agrupadas numa fileira de aproximadamente 4 000 células capilares “internas” (distribuídas ao longo da membrana basilar, da base ao ápice, do lado do modíolo, ou região central do “caracol” coclear) e em três fileiras de um total de aproximadamente 12 000 células capilares “externas” (vistas na Figura 2.7, no caso de um porquinho-da-índia). Quando os *estereocílios* de uma célula capilar são defletidos, impulsos elétricos são disparados nos neurônios gangliônicos que têm contato sináptico com a célula sensorial (ou, no caso de uma sinapse inibidora, a razão de disparos

espontâneos do neurônio poderia ser inibida). Os axônios desses neurônios formam as fibras aferentes do nervo auditivo; os seus potenciais de ação transportam coletivamente para o sistema nervoso central a informação sobre o movimento da membrana basilar decodificada digitalmente.

Um aspecto importante é a disposição das extremidades nervosas aferentes. Enquanto uma única fibra nervosa está conectada a apenas uma célula capilar interna, recebendo assim mensagens de uma região extremamente limitada da membrana basilar, cada fibra nervosa aferente ligada às fileiras externas faz contato com 10 a 50 unidades sensoriais espalhadas por vários milímetros, estando em condições de coletar informações de um domínio de frequência de ressonância muito mais amplo. As células capilares internas respondem à *velocidade* do movimento da membrana basilar porque a deflexão dos seus cílios é proporcional à velocidade do fluido endolinfático circunvizinho (a força sobre um obstáculo imerso num fluido viscoso é proporcional à velocidade do fluxo). As células capilares externas, por outro lado, reagem de acordo com o *deslocamento*, provavelmente porque os seus cílios estão ligados à membrana tectorial (as forças mecânicas interativas dependem da distorção momentânea da partição coclear); a resposta delas fica saturada para altos níveis de intensidade. O fato de que aproximadamente 95% das fibras aferentes do nervo acústico estão ligadas a células capilares internas, com 10-50 fibras individuais fazendo contato sináptico com cada célula, confere à fileira interna o papel de receptor sensorial primário. Mas, por outro lado, as células capilares externas recebem terminações de fibras eferentes que liberam impulsos neurais provenientes do sistema nervoso central¹⁹; esse fato, mais a motilidade acentuada das células capilares externas, descoberta recentemente, confere a essas células um papel duplo especial de receptoras e causadoras (a ser discutido em detalhes na Seção 3.6). Embora esteja claro que cada tipo de célula capilar desempenha um papel muito distinto na transdução do sinal acústico, os dois tipos devem trabalhar juntos: qualquer dano às células capilares externas reduz severamente a audição,

19. Uma pequena proporção de fibras eferentes também atua sobre a saída das células capilares internas, mas apenas indiretamente: elas provocam sinapse nas fibras aferentes que estão em contato com as células capilares internas.

mesmo que a fileira interna permaneça plenamente ativa. Continuaremos discutindo esse assunto na Seção 3.6; para informações atuais mais detalhadas, veja, por exemplo, Gelfand (1990), e Zwicker e Fastl (1990).

Verificou-se, pela implantação de microeletrodos em fibras nervosas cocleares acusticamente ativadas, que cada fibra tem um limite inferior de disparo para a frequência acústica f que provoca uma oscilação máxima na região x da membrana basilar (Figura 2.8) ligada a essa fibra. Essa frequência de resposta máxima é chamada de *frequência característica* do neurônio, ou “frequência ideal” (Kiang e outros, 1965). Voltando agora à distribuição temporal real dos pulsos individuais, as pesquisas (Zwislocki e Sokolich, 1973) mostraram que uma frequência máxima de disparos está associada à velocidade máxima da membrana basilar quando ela está se movendo em direção à *scala timpani*; durante o movimento em direção oposta, à *scala vestibuli*, ocorre uma inibição da frequência de disparos. Além disso, a posição momentânea da membrana basilar tem um efeito excitador ou inibidor (menos pronunciado), dependendo se a membrana está momentaneamente distorcida em direção à *scala timpani* ou na direção oposta, respectivamente. Ambos os efeitos somam-se para determinar uma resposta total. A Figura 2.22 mostra uma distribuição temporal hipotética de impulsos neurais numa fibra nervosa do ouvido interno conectada à região apropriada de ressonância da membrana basilar, quando ela é excitada por um som de baixa frequência com padrão de vibração trapezoidal (segundo Zwislocki e Sokolich, 1973).

Um exame dessa figura revela como a informação sobre a frequência de repetição (na verdade, sobre o período de repetição) do sinal acústico original pode ser codificada na forma de grupos de impulsos nervosos. A Figura 2.22 corresponde a um caso ideal de frequência extremamente baixa. Na verdade, as frequências acústicas são em geral mais altas do que as frequências com que ocorrem os disparos neurais, e a situação real corresponde melhor ao caso em os intervalos “ligado” e “desligado” estão delineados de uma maneira bem menos clara devido à sua curta duração (em comparação com um período refratário típico de neurônio) e ao caráter aleatório da distribuição de impulsos. A única propriedade estatisticamente importante é que haverá mais impulsos caindo nos intervalos “ligado” do que nos intervalos

“desligado”. Em conseqüência, para os sons puros, o intervalo de tempo entre impulsos sucessivos tenderá a ser um múltiplo inteiro do período de vibração sonora τ (Kiang e outros, 1965). É claro que, quanto mais alta for a freqüência do som, *i.e.*, quanto menor a duração dos intervalos “ligado”, menos definido será este grupo. Para freqüências acima de alguns milhares de Hz, isso absolutamente não é válido. Quando várias fibras que recebem estímulos de uma região estreita da membrana basilar estão amontoadas (como acontece no nervo auditivo), a soma dos seus impulsos (que podem ser detectados por um microeletrodo que faz contato simultaneamente com muitas fibras) estará em sincronia com o estímulo auditivo. Esses sinais nervosos coletivos sincronizados foram chamados de rajadas.

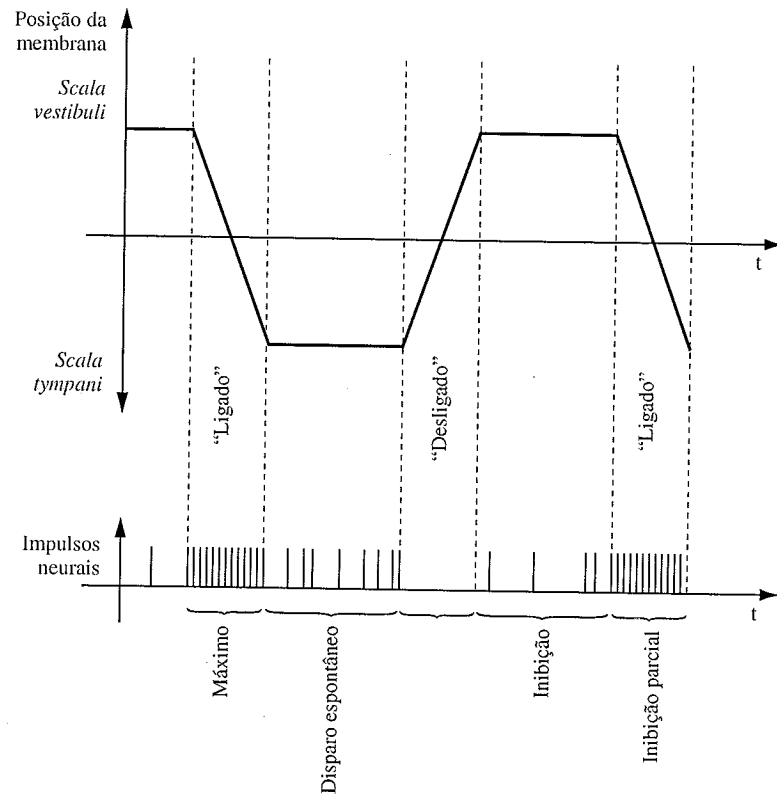


Figura 2.22

2.9 A ALTURA SUBJETIVA E O PAPEL DO SISTEMA NERVOSO CENTRAL

Pela seção anterior, conclui-se que cada fibra neural do nervo auditivo é capaz de portar dois tipos de informação.

1. O simples fato de que ela está disparando (ou que sua freqüência espontânea tenha sido inibida) diz ao sistema auditivo que a membrana basilar foi ativada (próximo ou) na região ligada àquela fibra — a distribuição espacial (ou organização “tonotópica”) das fibras disparando codifica a informação sobre a altura primária (esse processo é válido para toda a gama de freqüências).
2. A distribuição temporal real de impulsos leva a informação sobre a velocidade de repetição ou periodicidade, e também sobre detalhes do próprio padrão de vibração (veja abaixo). Isso só funciona na região de baixas freqüências.

Não há dúvida que a informação sobre a região de excitação é usada pelo sistema auditivo em todos os níveis. Mas será que esse sistema realmente utiliza a informação contida na distribuição temporal de pulsos neurais, vista esquematicamente na Figura 2.22?

Primeiro, vamos voltar um pouco à percepção de sons individuais puros (senoidais). Vários argumentos apontam para o fato de que a distribuição temporal de pulsos neurais não é utilizada na percepção da altura de um som puro. Por exemplo, cálculos teóricos (Siebert, 1970) prevêm que, se a altura primária fosse mediada por sugestões de tempo, a dlo de resolução de freqüência (p. ex., veja a Figura 2.9) seria independente da freqüência (e não é!), e ainda deveria decrescer com uma crescente amplitude de estímulo (o que não acontece!).

Não deve ser nenhuma surpresa que as sugestões de tempo sejam bastante ignoradas na percepção de altura primária de sons puros. Mas, e quanto à percepção de batimentos de consonâncias desafinadas e altura de periodicidade de complexos harmônicos? É difícil encontrar uma explicação para os batimentos de consonâncias desafinadas e outros efeitos sen-

síveis à fase sem assumir que, em algum estágio, um mecanismo analisa a estrutura fina temporal do padrão de vibração do estímulo.

Podemos lançar mão do efeito visto na Figura 2.21 para tentar uma explicação sobre como a informação acerca do padrão de vibração e suas variações (batimentos de segunda ordem) é codificada. Considere a superposição de dois sons a uma oitava de distância. Suponha que o padrão de vibração resultante seja aquele visto na parte inferior da Figura 2.16 e suponha que os movimentos descendentes da membrana basilar correspondam à direção negativa de y . Surgirão na membrana basilar duas regiões de ressonância, centradas nas posições x_1 e x_2 , correspondentes às duas frequências componentes $f_1, f_2 = 2f_1$ (Figura 2.8). No pacote de nervos cocleares, teremos focos principais de atividade centrados nas fibras cujas “frequências características” são f_1 e f_2 , levando a duas sensações primárias de altura a uma oitava de distância. Entretanto, as regiões de ressonância da membrana basilar são amplas o bastante para que haja uma superposição na região entre x_1 e x_2 , onde os pontos da membrana vibram de acordo com um padrão superposto relacionado ao movimento original do tímpano²⁰. Assim, as fibras conectadas a essa região de superposição responderão com disparos que estão agrupados em intervalos “ligados” que correspondam, digamos, às porções descendentes (declividade negativa) do segundo gráfico da Figura 2.16. Observe que, nesse caso, esses intervalos “ligados” não têm a mesma duração, porém formam uma seqüência alternada “curto-longo-curto-longo”. Se as duas notas estiverem a uma quinta de distância, o padrão de vibração da região superposta pode ser aquele que se vê na Figura 2.18, levando a um tipo diferente de seqüência de intervalos “ligados”. A periodicidade dessa seqüência representaria então a informação sobre a frequência de repetição, enquanto a estrutura da seqüência (uma espécie de “código Morse”) daria informação sobre o padrão de vibração. Essa estrutura fina tem sido identificada estatisticamente por meio de medições eletrofisiológicas. A Figura 2.23 mos-

20. As ondas que se propagam no fluido coclear mudam as suas relações de fase e, obviamente, a amplitude, enquanto se propagam, alterando assim a forma real do padrão de vibração em diferentes pontos da membrana basilar.

tra um exemplo (chamado histograma) da distribuição de intervalos de tempo entre pulsos neurais numa fibra nervosa auditiva (Rose e outros, 1969), para um estímulo que corresponde a uma quinta musical com uma certa relação de fase (veja padrão de vibração). Observe a diferença no número relativo de vezes (eixo vertical) em que aparece um certo intervalo entre pulsos sucessivos (eixo horizontal). Isso representa o “código Morse” (estatístico) mencionado acima, transportando informações sobre o padrão de vibração. Quanto maior for a complexidade do padrão de vibração original e maiores as frequências dos sons componentes, mais “nebulosa” será a informação transmitida pela seqüência de pulsos, *i.e.*, mais difícil de ser interpretada pelos centros cerebrais superiores.

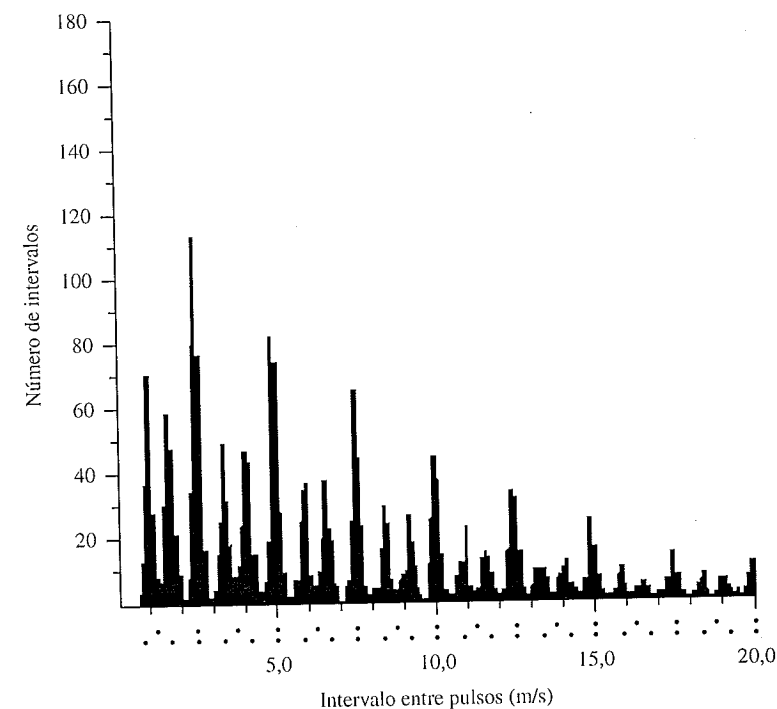


Figura 2.23 Um histograma mostrando o número de vezes (eixo vertical) em que um certo intervalo entre pulsos neurais ocorre (eixo horizontal), numa fibra nervosa auditiva estimulada com uma superposição de duas notas (quinta), (Rose e outros, 1969).

Uma análise detalhada da distribuição temporal de pulsos neurais implicaria a atuação, em certo grau, daquilo que se chama mecanismo de *autocorrelação temporal* (proposto inicialmente por Licklider, 1959), em que um “trem” de pulsos é comparado com trens anteriores para que detalhes repetitivos semelhantes (como os períodos indicados na Figura 2.20b) sejam realçados e os outros (não-periódicos) sejam suprimidos. Todavia, ainda não há evidência anatômica da existência de tal mecanismo no sistema nervoso auditivo aferente.

As sugestões de tempo também atuam no mecanismo responsável pela sensação de localização sonora espacial²¹ (estereofonia) (p. ex., Molino, 1974). Essa audição binaural deve envolver um processo chamado *transcorrelação* temporal dos sinais neurais de ambas as cócleas, segundo o qual a diferença de tempo entre os sinais das cócleas é determinado. Há evidência fisiológica da existência de tal mecanismo (na oliva superior média, Figura 2.26). Um modelo neural de transcorrelator foi proposto por Licklider (1959). Nesse modelo (Figura 2.24) assume-se que um neurônio ascendente só poderá disparar se for simultaneamente excitado por ambas as fibras de entrada. Como o sinal neural se propaga ao longo da fibra com velocidade finita, a chegada simultânea num determinado neurônio ascendente implica uma certa diferença de tempo entre os sinais originais de ambas as cócleas. Por exemplo, a simultaneidade exata (diferença de tempo nula) dos dois sinais cocleares dispararia um neurônio ascendente situado exatamente no centro, porque esse é o ponto onde os sinais da direita e da esquerda se encontram. Porém, se o sinal original for detectado primeiro no ouvido direito, o seu pulso passará do ponto central até que encontre o pulso atrasado do ouvido esquerdo. É fácil verificar que a posição y (Figura 2.24) do neurônio ascendente ativado dependerá do atraso de tempo interaural, que, por sua vez, depende da direção do som que chega. O diagrama da Figura 2.24 é um modelo hipersimplificado. Estudos detalhados da neuroanatomia do complexo olivar superior (Figura 2.26) reve-

21. Sugestões dadas pela intensidade (diferença de amplitude entre as ondas sonoras que chegam a cada um dos ouvidos) e pelo espectro (diferenças de timbre) também contribuem para a localização sonora, especialmente para frequências altas e em ambientes fechados.

lam uma estrutura mais complicada. Na verdade, a informação binaural é codificada por meio de uma complexa interação de entradas excitadoras e inibidoras resultante de diferenças de tempo (fase) e de intensidade entre os estímulos que chegam aos dois ouvidos (Goldberg e Brown, 1969). Qualquer que seja o verdadeiro mecanismo do transcorrelator, as suas habilidades de resolução temporal são impressionantes: os seres humanos podem localizar fontes sonoras no espaço com base em diferenças de tempo intra-aural de menos de 20 microssegundos! Esta é apenas uma pequena fração da duração de um potencial de ação. A enervação incomum das células capilares internas, com mais de 20 fibras fazendo contato com uma única célula sensorial, pode ser requisitada para assegurar uma coerente transmissão de informação neural com a rapidez necessária (Hudspeth, 1989).

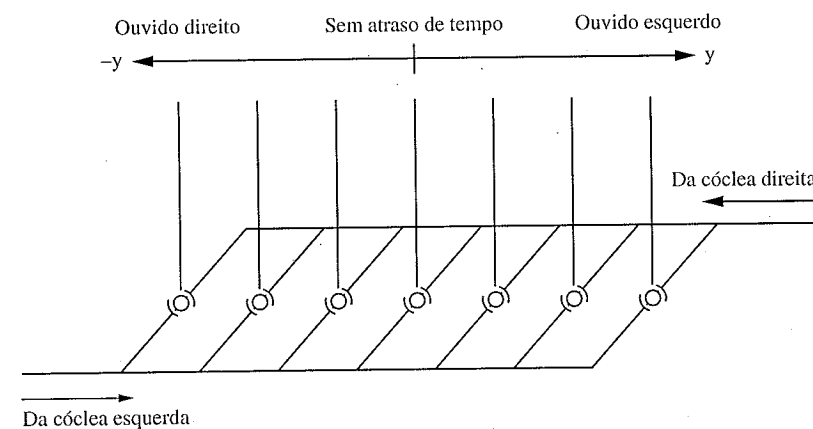


Figura 2.24 Modelo de um mecanismo de transcorrelação neural (detector de diferenças de tempo interaurais) (segundo Licklider, 1959).

Dois sons formando um intervalo desafinado, chegando cada um em um ouvido, podem “sujar” o transcorrelator: a diferença de fase entre os dois sons muda gradualmente (p. ex., Figura 2.16) e será interpretada por esse mecanismo como uma diferença variável entre os tempos de chegada

dos sinais auditivos direito e esquerdo. O mecanismo então avisará o cérebro com uma sensação (físicamente não existente) de direção sonora ciclicamente variável! É por isso que dois sons puros formando um intervalo dissonante desafinado, apresentados dicoticamente com fones de ouvido, dão a estranha sensação de uma imagem sonora que parece estar “rodando dentro da cabeça”.

Um assunto não resolvido nas últimas décadas tem sido a questão de se as análises seqüenciais temporais de pulsos neurais são ou não uma hipótese necessária para explicar a percepção da altura de periodicidade (p. ex., Yost e Watson, 1987). Um mecanismo de autocorrelação temporal, com sua capacidade potencial de detectar a velocidade de repetição dos sinais neurais, pode explicar alguns resultados psicoacústicos importantes do rastreamento de fundamentais (mas não todos). Se não for uma análise das sugestões de tempo, qual será o mecanismo que nos possibilita atribuir uma única altura a um complexo tonal harmônico — mesmo que a fundamental não esteja presente no estímulo original? Por que percebemos as alturas correspondentes às freqüências dadas pelas relações (2.7a)-(2.7c) quando uma melodia é tocada com os complexos harmônicos de dois sons?

Algumas idéias que podem levar a uma explicação desses efeitos (p. ex., Terhardt, 1972; Wightman, 1973; Goldstein, 1973) serão apresentadas aqui de uma forma bem simplificada. Os sons “naturais” presentes nas comunicações acústicas humanas e dos outros animais contêm uma proporção importante de sons harmônicos (vogais, cantos de pássaros, vozes dos animais). Esses sons apresentam uma propriedade em comum — são formados por uma superposição de harmônicos de freqüências nf_1 , múltiplos inteiros de uma freqüência fundamental f_1 . Esses sons geram um complicado padrão de ressonância na membrana basilar, com vários picos de amplitude, um para cada harmônico (Figura 2.25). A despeito de sua complexidade, esse padrão pode comportar algumas características invariantes. Uma dessas invariâncias é a relação de distância entre máximos de ressonância vizinhos²². Para harmônicos de ordens superiores (n maior

22. Outra característica invariante é a alta coerência das variações temporais macroscópicas desse complicado padrão de excitação sobre todo o âmbito espacial da membrana basilar.

que 7 ou 8), essa relação perde a sua definição física por causa das superposições mútuas de ressonância?²³

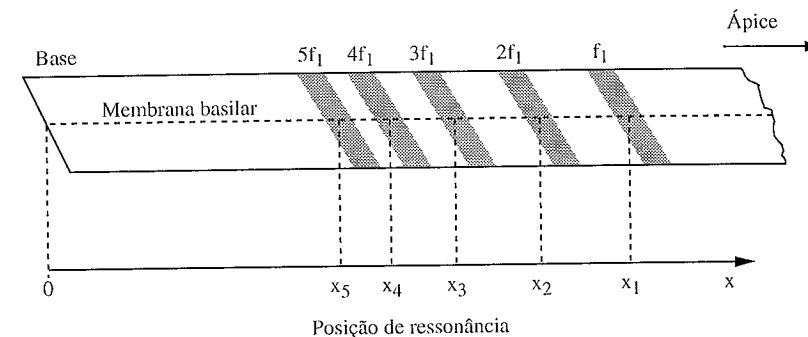


Figura 2.25

Podemos ter aprendido, em tenra idade, a reconhecer essa característica invariante como “uma mesma coisa” (Terhardt, 1972, 1974), ou talvez tenhamos um mecanismo inato para isso (Wightman, 1973; Goldstein, 1973). Vamos chamar esse mecanismo de reconhecimento de *processador central de alturas*. A principal função dessa unidade neural é transformar o padrão de atividade periférica em outro padrão, de tal modo que todos os estímulos com a mesma periodicidade sejam representados de maneira semelhante. O resultado é uma única sensação de altura — a despeito da contribuição de muitos harmônicos e da resultante complexidade do padrão de excitação primária. Essa sensação de altura única corresponde à da componente fundamental f_1 , que nos sons “naturais” é normalmente a mais intensa. Tudo isso deve funcionar em analogia com o reconhecimento de padrões visuais. Por exemplo, quando você vê o símbolo μ , ele pode não transmitir nenhum significado “único” (a sua interpretação dependerá provavelmente da orientação espacial do símbolo e do contex-

23. Devido à relação particular entre o lugar da ressonância e a freqüência (Figura 2.8), as regiões de ressonância vão se aglomerando cada vez mais quando subimos a série harmônica (quando n aumenta).

to em que ele é visto). Mas qualquer um que esteja familiarizado com o alfabeto cirílico irá percebê-lo claramente como “uma única coisa” (a letra “ch”), não importando onde ele esteja projetado no campo visual nem com que orientação.

Supõe-se que temos, edificados em nosso sistema de processamento central, modelos básicos com os quais podemos comparar as complexas estruturas do padrão de excitação espacial da membrana basilar. Sempre que uma coincidência é conseguida, uma única sensação de altura é gerada. Esse processo de casamento funciona mesmo que apenas uma seção parcial do padrão de excitação esteja disponível. Se, ao invés de um som complexo “natural”, formos expostos a um onde alguns elementos normalmente esperados sejam suprimidos (p. ex., uma fundamental ausente), o padrão de excitação parcialmente truncado da membrana basilar, quando chega ao mecanismo de reconhecimento do processador de alturas, ainda pode apresentar coincidências, dentro de certas limitações. Mais uma vez, encontramos muitas analogias com o reconhecimento de padrões visuais. Um exemplo notável é a percepção aparente de contornos não-existentes — mas esperados, que experimentamos ao olharmos as letras do título impressas na página de rosto deste livro,

O processo de casamento descrito acima funciona mesmo que os harmônicos de um som sejam fornecidos alternada mas simultaneamente a cada ouvido (p. ex., Houtsma e Goldstein, 1972). Isso obviamente significa que o processador central de alturas deve estar localizado em algum estágio superior do caminho auditivo, depois que as entradas de ambas as cócleas se combinaram. Além disso, o processo de casamento funciona mesmo que apenas dois harmônicos vizinhos de um som complexo sejam apresentados, como se vê na Seção 2.8. Nesse caso, porém, o mecanismo de casamento pode cometer erros — e se fixar numa das muitas posições “aceitáveis”.

Três teorias prevêm corretamente muitos dos resultados quantitativos de medições psicoacústicas relativas à altura de periodicidade. Uma delas (Goldstein, 1973) assume que a informação neural sobre as posições espaciais dos máximos de ressonância na membrana basilar não é nitidamente definida, flutuando estatisticamente dentro de certos limites.

Espera-se que o casamento do modelo seja tal que minimize as diferenças em relação ao sinal real dentro das flutuações esperadas (uma estimativa de semelhança máxima da frequência fundamental e da ordem harmônica — veja mais sobre este assunto na Seção 2.8 e Apêndice II). Uma teoria quase equivalente (Wightman, 1973) supõe a operação de um mecanismo de autocorrelação atuando no domínio espacial (diferentemente da autocorrelação temporal mencionada na página 90). A distribuição de atividade espacial induzida num grupo de fibras nervosas recebendo informação de ambas as cócleas é informada a uma rede de neurônios onde a ativação de uma dada área espacial é comparada quantitativamente com a existente em áreas distantes num certo instante. Nesse processo, certos detalhes de entrada são convertidos em um foco de atividade de saída cuja localização espacial codifica *uma* qualidade (*i.e.*, a altura) que está relacionada a uma certa característica invariante da entrada (*i.e.*, as relações de distância entre os máximos de excitação da membrana basilar), não importando outras possíveis variáveis de entrada (p. ex., as intensidades ou as fases dos diversos harmônicos). Devemos ressaltar que esse mecanismo de autocorrelação “espacial” muito provavelmente envolve também um processo de aprendizagem. Na verdade, pode ser necessário um condicionamento para fazer com que o sistema responda às constelações de máximos de atividade espacial relevantes que ocorrem com mais frequência (Figura 2.25), provocados por sons naturais (harmônicos)²⁴.

Nenhuma dessas teorias propõe de que maneira algoritmos-chave (casamento de modelos ou autocorrelação espacial, respectivamente) são efetivamente realizados pelo processador de alturas no sistema nervoso. Porém, acredita-se que existam redes neuronais perfeitamente capazes de realizar as operações de soma e multiplicação de pulsos neurais necessárias na execução desses algoritmos. A teoria de Terhardt (1974) chega mais perto na proposição de redes neurais. Na verdade, seus cálculos estão baseados numa *matriz de aprendizagem*, um circuito analógico que

24. É improvável que a “métrica” necessária (quasilogarítmica e com dependência coclear) esteja geneticamente incorporada.

“aprende a responder”²⁵ a detalhes característicos das configurações de entrada que ocorrem com mais frequência (*i.e.*, às relações de distância entre os máximos de excitação de entrada provocados por um som complexo). Voltaremos a essas teorias mais tarde, quando discutirmos explicitamente a percepção de sons musicais complexos (Seção 4.8 e Apêndice II) e a consonância e dissonância (Seção 5.2).

Finalmente, afirmamos que não se pode excluir a possibilidade de que seja feito pelo menos um uso parcial da distribuição temporal de pulsos neurais na percepção da altura de sons complexos. É difícil acreditar que o sistema nervoso, sempre equipado para trabalhar com uma eficiência espantosa, com tantos sistemas de apoio, não tire vantagem da conveniente informação em “código Morse” (Figura 2.21), que de fato existe nos canais periféricos de transmissão auditiva! É de esperar, como já mencionamos, que o mecanismo de alturas com base espectral funcione melhor para números baixos de harmônicos (onde as regiões de ressonância vizinhas estão bem separadas umas das outras), enquanto um mecanismo com base temporal deverá operar otimamente na região dos harmônicos superiores. Alguns experimentos psicoacústicos requerem, a princípio, uma explicação via análise de sugestões temporais. Por exemplo, sons puros de baixa frequência com duração muito curta (2-3 ciclos de vibração real) podem originar uma clara sensação de altura (Moore, 1973). Ou, se um sinal acústico (ruído branco) for apresentado a um ouvido apenas, com o outro ouvido recebendo o mesmo sinal com um atraso de um intervalo de tempo τ (alguns milissegundos), percebe-se uma altura muito fraca, correspondendo à frequência $1/\tau$ (Bilsen e Goldstein, 1974). Nenhum desses resultados pode ser explicado satisfatoriamente por uma teoria de “posicionamento” (análise de indícios espaciais).

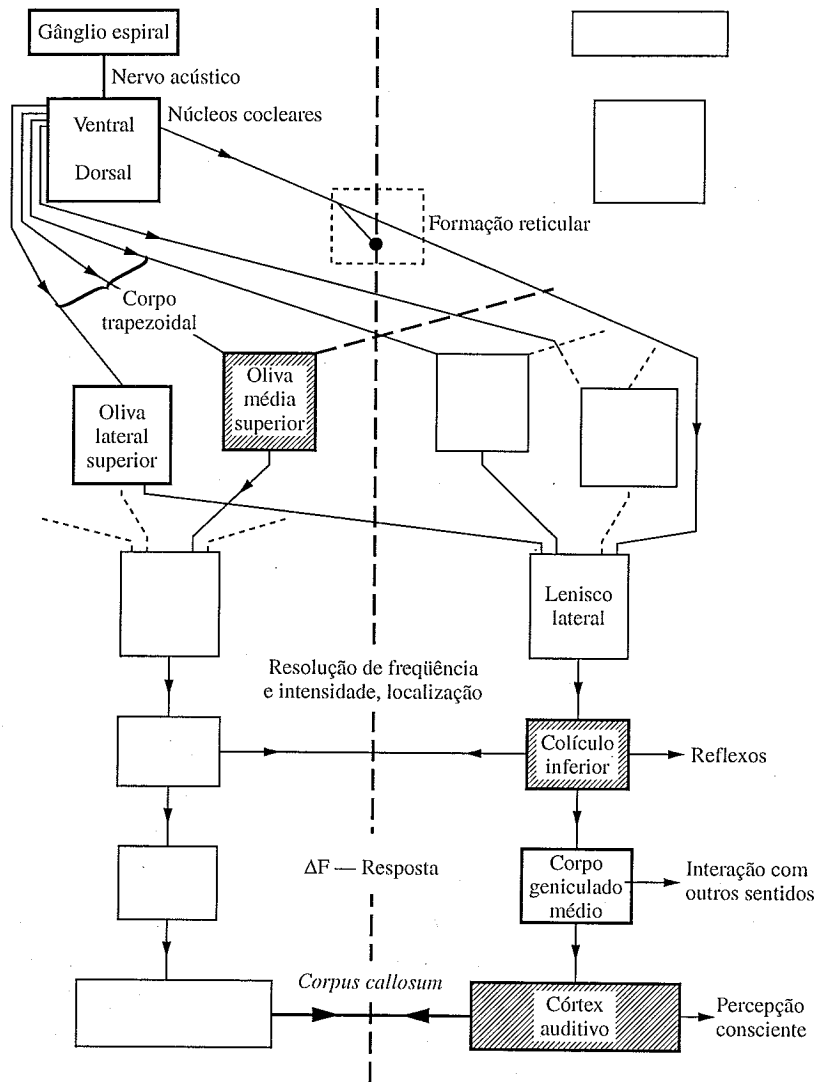
Com a discussão precedente, uma coisa fica clara: a percepção de altura subjetiva exige que se realizem operações “superiores” de extração de altura no sistema nervoso central, depois que as entradas das cócleas estiverem combinadas. Por essa razão, concluímos este capítulo com uma descrição resumida de alguns dos aspectos mais relevantes do caminho

25. Isso é conseguido no modelo do laboratório diminuindo-se adequadamente as resistências elétricas entre as linhas de transmissão (as linhas e colunas da matriz) ativadas simultaneamente (conduzindo corrente) por uma certa configuração de entrada repetidamente apresentada.

auditivo (Whitfield, 1967; Brodal, 1969; Gelfand, 1990). Isso servirá também como referência para discussões a serem apresentadas nos próximos capítulos. A exploração anatômica dos caminhos neurais e de suas interconexões é uma tarefa experimental extremamente difícil. Os neurônios são células com apófises (axônios e dendritos) que podem medir vários centímetros; cada neurônio, especialmente do tecido cerebral, pode receber informação de milhares de células, enquanto passa informação para centenas de outras. É quase impossível seguir microscopicamente a rede de interconexões no caso de uma célula. Podem-se fazer apenas estimativas grosseiras dos aspectos gerais desses caminhos com a utilização de várias técnicas de coloração e degeneração de células, ou seguindo pelo tecido o curso de determinados padrões específicos de atividade neural.

A Figura 2.26 apresenta um diagrama do caminho auditivo desde a cóclea até a área de recepção auditiva do córtex cerebral, na forma de diagrama de blocos. Esse diagrama descreve os canais de transmissão de informação e estações de processamento de informação repassada, não mantendo nenhuma relação de escala com a realidade neuroarquitetônica. O *gânglio espiral* é o circuito neural da cóclea, um primeiro estágio de processamento nesse caminho. É aqui que os neurônios em contato com as células capilares internas e externas têm uma primeira oportunidade de interação, determinando a distribuição espaço-temporal específica da atividade do nervo acústico (o 8º nervo cranial), que transmite essa informação ao cérebro. O próximo estágio de processamento, localizado na medula oblonga, compreende os núcleos cocleares, compostos por três subdivisões cuja elaborada estrutura é responsável pelos primeiros passos nas tarefas de resolução e discriminação de sons. A partir daqui, a informação neural é canalizada para três feixes principais. Um cruza diretamente para o lado contralateral oposto e entra no lenis-co lateral, o principal canal que passa pela raiz do cérebro (ponte). Algumas fibras terminam na *formação reticular*, uma rede difusa na raiz do cérebro que desempenha a função de principal “painel de controle” cerebral²⁶. Um

26. Essa estrutura, que recebe dados básicos dos sentidos e do corpo e também elabora a informação do córtex, é responsável por ativar ou inibir o processamento cerebral de acordo com as necessidades do momento, controla o sono, o despertar e a consciência e influi em muitas funções viscerais.



outro feixe, o *corpo trapezoidal*, tem fibras que vão do núcleo coclear ventral até importantes estações de processamento e retransmissão, as *olivas lateral superior e média superior*. Destas, a oliva média superior é o primeiro centro intra-aural de mixagem de sinais. É nesse local que um transcorrelator (Figura 2.23) pode fornecer a informação necessária para a localização da fonte sonora. Finalmente, um terceiro feixe intermediário leva do núcleo coclear ventral ao complexo olivar contralateral.

Os três estágios superiores envolvem o *colículo inferior*, o *corpo geniculado médio* e, finalmente, o *córtex auditivo* (Figura 2.25). Algumas fibras estão conectadas ao colículo superior, também ligado a caminhos visuais. Observe as interconexões, nesses vários estágios, com o caminho contralateral e com outros caminhos sensoriais e centros cerebrais.

A Figura 2.26 não mostra uma rede de *fibras eferentes*, que leva informação dos estágios superiores para os inferiores e termina na cóclea. Esse sistema tem um papel no controle da informação aferente de entrada. O sistema inferior da rede eferente, o *feixe olivococlear*, pode participar de uma forma importante no processo de depuração (p. 62). Embora existam apenas umas 1 600 fibras eferentes que chegam a cada cóclea, as fibras maiores enervam profusamente as células capilares externas e podem, portanto, exercer um controle fundamental sobre o funcionamento mecânico e/ou elétrico destas últimas (ver Seção 3.6).

Finalmente, vamos apresentar algumas generalidades que serão úteis nos próximos capítulos. No estágio inicial, há uma correspondência geométrica bem específica entre as fibras neurais ativadas e a posição espacial do estímulo da fonte na membrana basilar. Na verdade, a distribuição espacial de excitação ao longo da membrana basilar é mapeada continuamente numa distribuição espacial de atividade neural através dos feixes de fibras. Isso é particularmente evidente em cada um dos núcleos cocleares. Contudo, quando subimos para estágios superiores, essa correspondência vai se perdendo gradualmente (exceto num estado anestesiado). O número de neurônios participando cresce dramaticamente e a resposta neural torna-se cada vez mais representativa de detalhes complexos do sinal sonoro, sendo mais e mais influenciada pelo estado e desempenho comportamentais do indivíduo. Os canais contralaterais (*i.e.*, cruzados) são

Figura 2.26 Fluxograma dos sinais neurais no caminho auditivo, desde o ouvido, passando pela raiz do cérebro, até os córtices auditivos.

“melhores” portadores de informação do que os canais ipsilaterais (do mesmo lado) — se uma informação conflitante for apresentada a ambos os ouvidos, o canal contralateral tende a ignorar a informação que está sendo enviada a um determinado hemisfério pelo canal ipsilateral (Milner e outros, 1968).

No mesmo estágio do colículo inferior também já existe uma boa resolução da frequência, intensidade e direção do som; há também uma resposta seletiva para varreduras de frequência. Os reflexos funcionam, mas nesse estágio não há evidência de uma percepção consciente do som, como foi mostrado por experimentos de ablação. Alguma capacidade de reconhecimento de padrões já está atuante no corpo geniculado médio (e provavelmente no colículo superior). Nesse estágio, existe informação sobre *onde* se localiza um determinado estímulo sonoro e *para onde* ele vai no espaço e no tempo. Ocorre a primeira integração com as informações dos outros sentidos.

O último estágio de processamento da informação que chega acontece na área de recepção do córtex auditivo. A função primária é identificar o estímulo, integrá-lo à imagem do ambiente normalmente configurada e colocá-lo à disposição do estado consciente do cérebro. Na verdade, daqui para a frente, a informação é distribuída aos outros centros cerebrais, onde ela é armazenada, analisada, integrada à totalidade das funções cognitivas do cérebro, ou descartada como irrelevante. O *corpus callosum* (Figura 2.26), uma gigantesca junção de aproximadamente 200 milhões de fibras ligando ambos os hemisférios cerebrais, desempenha um papel decisivo no processamento global de informações, especialmente em vista da notável especialização dos dois hemisférios, já mencionada na Seção 1.5. Voltaremos a esse assunto na Seção 5.6.

ONDAS SONORAS, ENERGIA ACÚSTICA E A PERCEPÇÃO DO VOLUME

No capítulo anterior, estudamos vibrações sonoras simples e os seus efeitos subjetivos, sem investigar como elas realmente chegam ao ouvido. Mencionamos experimentos onde a fonte sonora (fones de ouvido) estava situada muito próximo dos tímpanos. Neste capítulo, vamos discutir o processo de propagação da energia sonora a partir de uma fonte distante até o ouvinte e analisar como esse fluxo de energia acústica determina a sensação de volume.

Terminaremos o capítulo com outro exame mais detalhado dessa maravilha eletromecânica que é a cóclea.

3.1 ONDAS ELÁSTICAS, FORÇA, ENERGIA E POTÊNCIA

Quando o som se propaga por um meio, os pontos desse meio vibram. Se não houver nenhum som, e se não houver nenhum outro tipo de pertur-

bação, cada ponto do meio¹ estará em repouso e assim permanecerá até que façamos alguma coisa ao meio. A posição espacial de um certo ponto do meio, quando este não apresenta nenhuma perturbação, é chamada de posição de equilíbrio desse ponto.

As ondas sonoras são uma forma particular das chamadas ondas elásticas. Sempre que produzimos uma súbita deformação num certo ponto do meio (p. ex., quando percutimos uma corda do piano com o martelo, ou quando subitamente deslocamos a coluna de ar ao colocarmos em movimento a palheta de uma clarineta), as forças elásticas farão com que os pontos próximos à deformação inicial comecem a se mover. Esses pontos, por sua vez, atuarão, por meio de forças elásticas, sobre outros pontos vizinhos, passando-lhes a ordem de começar a se mover, e assim por diante. Essa “reação em cadeia” representa uma onda elástica afastando-se da região de perturbação inicial. O que se afasta com esta onda não é matéria, mas energia: aquela energia necessária para pôr em movimento cada ponto alcançado pela onda. As ondas sonoras de interesse em música são ondas elásticas nas quais os pontos executam movimentos periódicos. Em sua vibração, cada ponto do meio sempre permanece muito, muito próximo da posição de equilíbrio. Uma onda sonora afasta-se da fonte com uma velocidade bem definida e em linha reta, até que seja absorvida ou refletida. A maneira como as ondas sonoras se propagam, são refletidas ou absorvidas determina as qualidades acústicas de um aposento ou sala de concerto.

Mencionamos acima os conceitos de força e energia. É necessário agora especificar o seu significado físico preciso. Todo mundo tem uma noção intuitiva de força: o quanto precisamos empurrar ou puxar um objeto para alterar-lhe a forma, para colocá-lo em movimento, o esforço que precisamos fazer para erguê-lo, para reduzir-lhe o movimento etc. Mas a física não fica satisfeita com meros conceitos intuitivos. Precisamos apresentar uma definição clara de força, assim como a “receita” de como medi-la. Tanto a definição quanto a receita devem estar baseadas em cer-

1. Um “ponto” do meio é usado aqui no sentido macroscópico, envolvendo bilhões de moléculas!

tos experimentos cujos resultados são condensados ou resumidos na formulação de uma lei física.

É de nossa experiência diária que, a fim de mudar a forma de um corpo, precisamos fazer-lhe algo específico: temos que “aplicar-lhe uma força”. A deformação, *i.e.*, a variação na forma, não é o único efeito possível de uma força que age sobre um corpo. Na verdade, a nossa experiência diária também nos diz que, se quisermos afetar o movimento de um corpo, também temos que lhe aplicar uma força. Em termos gerais, verifica-se que a aceleração a de um corpo, que representa a rapidez de variação de sua velocidade, causada por uma força F , é proporcional a esta. Ou, inversamente, a força é proporcional à aceleração produzida: $F = ma$. Esta é a chamada equação de Newton. A constante de proporcionalidade m é a massa do corpo. Ela representa a sua “inércia” ou “resistência” a uma variação de movimento. Se mais de uma força estiver agindo sobre o corpo, a aceleração resultante será dada pela soma de todas as forças. Essa soma pode ser zero; neste caso, as forças atuantes estarão em equilíbrio.

A unidade de força é definida como a força necessária para acelerar um corpo de 1 kg numa razão de 1 m/s². Essa unidade de força é chamada *newton*. Um newton é igual a 0,225 libras-força. A libra é uma unidade de força (peso) mais familiar entre os povos dos países de língua inglesa. Como a aceleração da gravidade é 9,8 m/s², o peso de um corpo de 1 kg de massa será 9,8 N (= 2.2 lbs). Podemos medir uma força medindo a aceleração que ela impõe a um corpo de massa definida, ou equilibrando-a (*i.e.*, cancelando o seu efeito) com uma força conhecida — p. ex., a tensão numa mola calibrada².

Em muitas situações físicas, uma força dada é aplicada ou “espalhada” sobre uma região da superfície do corpo. Por exemplo, num avião a grande altitude com cabine pressurizada, o ar de dentro exerce uma considerável força para o exterior, F , em cada janela (e em qualquer outra parte da fuselagem), força esta que é proporcional à superfície S da janela. A rela-

2. “Calibrada” significa que já determinamos previamente o quanto a mola se deforma sob determinada força, p. ex., um certo peso.

ção $p = F/S$ representa a pressão do ar dentro da cabine. Em geral, definimos a pressão do ar como a razão entre a força que age sobre uma superfície S que separa o ar do vácuo. Se, em vez de vácuo, tivermos simplesmente uma pressão diferente p' do outro da superfície, a força F agindo sobre S será dada por

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

Tudo isso é muito importante em música. As ondas sonoras no ar são oscilações de pressão do ar. Assim, se na relação (3.1) S corresponder à superfície do tímpano, sendo p' a pressão (constante) no ouvido médio e p a pressão oscilante no meato (Figura 2.6), F será a força oscilante que age sobre o tímpano, responsável pelo seu movimento e pelo da cadeia de ossos no ouvido médio.

A pressão é expressa em newtons por metro quadrado ou pascais. A pressão atmosférica normal ao nível do mar é de aproximadamente $100\,000 \text{ N/m}^2$ (= $1\,000$ hectopascals). Uma unidade mais familiar nos EUA é a libra por polegada quadrada (por exemplo, o nível de pressão dentro das câmaras dos pneus dos automóveis é normalmente medido nessa unidade). Convertendo N em lbs e m^2 em pol^2 encontramos que $1 \text{ N/m}^2 = 0,00015 \text{ lbs/pol}^2$. A pressão normal do ar ao nível do mar portanto é de 15 lbs/pol^2 .

Vamos agora tratar do conceito de energia. Aqui também temos uma idéia intuitiva sobre isso — mas nossa intuição pode muito bem nos enganar neste caso. Por exemplo, algumas pessoas dizem que “é necessária uma grande quantidade de energia para segurar uma mala pesada por muito tempo” — embora para o físico não haja nenhuma energia envolvida (exceto durante o ato de levantar a mala ou colocá-la no chão). Para o fisiologista, por outro lado, é necessário um fluxo contínuo de energia química para os músculos para manter um estado contínuo de contração muscular. Para evitar confusão, é necessário introduzir o conceito de energia de uma forma mais quantitativa e precisa.

Apenas o conceito de força não é suficiente para a solução dos problemas práticos da física. Por exemplo, precisamos saber por quanto tempo,

ou até que distância, uma força constante está agindo se quisermos determinar, digamos, a velocidade final adquirida por um corpo acelerado por essa força (mesmo a força mais intensa pode ter apenas um pequeno efeito final, se a duração ou o alcance de sua ação forem muito curtos). Na verdade, o que realmente conta para se determinar uma certa variação de velocidade, digamos de 0 a v , é o produto da força vezes a distância percorrida na direção da força. Se chamamos de x essa distância, pode-se mostrar matematicamente, baseado na equação de Newton, que $F \cdot x = 1/2 mv^2$. O produto $F \cdot x$ é chamado de trabalho e será positivo se o deslocamento x for na mesma direção da força F . O produto $1/2 mv^2$ é chamado de energia cinética do corpo de massa m . Se $F \cdot x$ for positivo, interpretamos a relação acima dizendo que o trabalho da força aumentou a energia cinética do corpo ou, equivalentemente, que o “sistema liberou trabalho”, aumentando a sua energia cinética de 0 até $1/2 mv^2$.

O trabalho e a energia cinética são medidos em $\text{N} \cdot \text{m}$. Essa unidade é chamada *joule*, em homenagem a um físico e engenheiro britânico de nome Joule. Um corpo de 1 kg (unidade de massa) movendo-se a uma velocidade de 1 m/s tem uma energia cinética de $0,5 \text{ joule}$. Se ele se mover com o dobro dessa velocidade, a sua energia cinética será quatro vezes maior: 2 joules . Uma pessoa normal (70 kg), correndo a uma velocidade de 3 m/s ($10,8 \text{ km/h}$) tem uma energia cinética de 315 J ; a de um carro de $2\,000 \text{ kg}$ viajando a 30 m/s (108 km/h) é $900\,000 \text{ J}$.

A energia pode aparecer em outras formas além da cinética. Considere um corpo preso a uma mola. Precisamos fornecer uma certa quantidade de trabalho para podermos comprimir a mola. Se fizermos isso muito lentamente, praticamente não haverá energia cinética envolvida. Mas o trabalho fornecido será convertido em energia potencial; nesse caso, a energia potencial elástica do corpo preso a uma mola comprimida. Soltando a mola, o corpo será acelerado pela força da mola em expansão, e a energia potencial será convertida em energia cinética. Podemos dizer que a energia potencial é a energia de posição de um corpo, e a energia cinética é a sua energia de movimento.

A soma das energias cinética e potencial de um corpo é chamada de energia mecânica total (existem muitas outras formas de energia que não

vamos considerar: térmica, química, eletromagnética etc.). Existem casos importantes em que a energia mecânica de um corpo permanece constante. Um caso “musicalmente” importante é o do exemplo prévio de um corpo preso a uma mola, oscilando sob a ação da força elástica da mola. Pode-se demonstrar que a vibração resultante em torno da posição de equilíbrio é harmônica (desde que a amplitude permaneça pequena). Quando o corpo é liberado a partir de uma posição tensionada, a sua energia cinética inicial é zero. Mas ele possui uma energia potencial elástica inicial que, quando as oscilações começam, é transformada em energia cinética. Sempre que o corpo passa pela posição de equilíbrio, a energia potencial elástica é instantaneamente nula, enquanto a sua energia cinética é máxima. Durante a oscilação harmônica, existe uma conversão de energia potencial em cinética e vice-versa.

A energia mecânica total permanece constante desde que se verifique a ausência de forças dissipativas. O atrito causa uma diminuição contínua da energia total e, conseqüentemente, uma diminuição da amplitude de oscilação. O movimento resultante é chamado de oscilação amortecida. Ele é extremamente importante em música. Na verdade, muitos instrumentos musicais envolvem oscilações amortecidas; uma corda de piano é um exemplo típico. Outras forças externas podem agir de tal modo que incrementem gradualmente a energia mecânica. Elas podem ser usadas para compensar perdas dissipativas e, assim, manter uma oscilação com amplitude constante. A corda de um violino é um exemplo típico: as forças que aparecem ao se friccionar a corda com o arco fornecem energia para a corda vibrante numa proporção que é igual à proporção de energia perdida através do atrito e da radiação acústica (Seção 4.2).

Chegamos agora ao último, porém crucial, ponto relativo à energia. As máquinas (e os seres humanos) liberam energia a uma certa velocidade. Qualquer máquina (ou ser humano) pode realizar quase arbitrariamente uma grande quantidade de trabalho — mas para isso seria necessário muito tempo. O que realmente define a “qualidade” ou “potência” de uma máquina é a velocidade com que ela pode liberar energia (*i.e.*, realizar trabalho). Essa velocidade, se for constante, é dada por

$$P = \frac{\text{trabalho realizado}}{\text{tempo gasto}} = \frac{W}{(t_2 - t_1)} \quad (3.2)$$

W é o trabalho liberado entre os tempos t_1 e t_2 . P é chamado de potência mecânica. Sua unidade de medida é joule/seg., também chamada *watt* (nome de outro engenheiro britânico). Se você subir uma escada, seu corpo libera uma potência de aproximadamente 600 watts; a energia elétrica consumida por segundo por um ferro elétrico é de cerca de 1000 watts, e a potência máxima liberada pelo motor de um carro pequeno é 30 kW (1 HP = 0,735 kW). Um trombone tocando fortíssimo emite uma energia acústica total de aproximadamente 6 watts.

O conceito de potência é dos mais importantes para a física da música. Na verdade, o nosso ouvido não está absolutamente interessado na energia acústica total que chega ao tímpano — antes, ele é sensível à rapidez com que chega esta energia, *i.e.*, à potência acústica. Essa rapidez é o que determina a sensação de volume.

3.2 VELOCIDADE DE PROPAGAÇÃO, COMPRIMENTO DE ONDA E POTÊNCIA ACÚSTICA

Depois da nossa excursão no campo da física “pura” na seção anterior, já temos melhores condições de compreender o fenômeno da propagação de ondas. Para esse efeito fazemos uso de um modelo do meio. Nós o imaginamos formado por pequenos corpos de massa determinada, ligados uns aos outros por molas comprimidas (representando as forças elásticas). Inicialmente, as forças das molas estão em equilíbrio e todos os pontos estão em repouso. A Figura 3.1 mostra a situação quando o ponto P é subitamente deslocado para a direita numa quantidade x_1 .

Considerando as forças vistas na Figura 3.1, observamos que os pontos Q e R — que inicialmente estão em repouso nas suas respectivas posições de equilíbrio — ficam sujeitos a uma força resultante que age para a direita. Em outras palavras, de acordo com a equação de Newton,

eles serão acelerados para a direita e iniciarão um movimento na mesma direção em que o ponto P havia sido inicialmente deslocado. Esse ponto P , por outro lado, estará caminhando de volta à sua posição de equilíbrio, acelerado por uma força resultante que age sobre ele para a esquerda (Figura 3.1). Pouco tempo depois, quando os pontos Q e R estiverem seguindo em direção à direita, a compressão da mola entre R e T começa a aumentar, enquanto a da mola entre S e Q diminui. É fácil ver que os pontos S e T inicialmente estarão sujeitos a uma força dirigida para a direita e que os fará se moverem nessa direção, enquanto Q e R podem estar já se movendo de volta para a esquerda. Esse processo vai continuando de ponto para ponto — representando uma onda que se propaga a partir de P para ambos os lados. A “frente” da onda nada mais é do que uma “ordem” que vai de ponto para ponto dizendo: “Comece a se mover para a sua direita”. A “ordem” é dada pelas molas comprimidas (as forças elásticas). Em nenhum momento existe qualquer transporte de matéria envolvido. Chamamos este caso de onda longitudinal, porque os deslocamentos dos pontos são paralelos à direção de propagação da onda. No caso real de uma onda sonora que se propaga pelo ar, a ação conjunta das forças das molas sobre os pontos P, Q, R, \dots corresponde aproximadamente à pressão do ar; as variações dessas forças (p. ex., variações das distâncias entre os pontos) correspondem às variações de pressão do ar nas ondas sonoras.

O modelo unidimensional da Figura 3.1 também mostra como o transporte de energia está envolvido numa onda elástica. Em primeiro lugar,

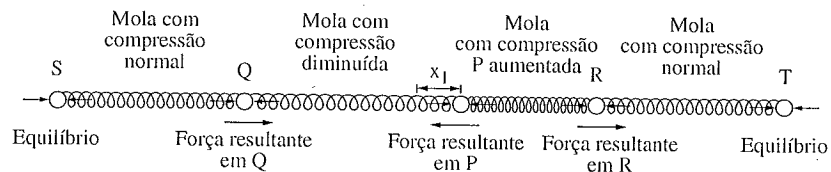


Figura 3.1 Modelo unidimensional para um meio elástico (molas em compressão), onde o ponto P foi deslocado longitudinalmente.

temos que realizar um trabalho “externo” para produzir o deslocamento inicial x_1 no ponto P , porque precisamos alterar os comprimentos das duas molas PQ e PR . Em outras palavras, é necessário uma fonte de energia. Nesse caso, a energia inicial é convertida a forma de energia potencial (de posição) do ponto P . Então, com o passar do tempo, os pontos à direita e à esquerda de P começam a se mover e os comprimentos das molas se alteram. Todos esses processos envolvem energia cinética (movimento dos pontos) e potencial (compressão ou expansão das molas). A energia inicialmente fornecida ao ponto P é transferida de ponto a ponto enquanto a onda se propaga: temos um fluxo ou transporte de energia que se afasta da fonte.

Vamos agora voltar ao caso em que as molas do modelo estão sob tensão (expandidas), em vez de estarem comprimidas, com os pontos vizinhos puxando-se uns aos outros. Fisicamente, isso corresponde a uma corda de violino tensa. Para deslocamentos longitudinais (na direção ao longo das molas) obtemos uma situação qualitativamente semelhante à anterior em termos de propagação de ondas, apenas com a diferença de que todas as forças da Figura 3.1 estão agora invertidas. Mas, além disso, temos uma possibilidade totalmente nova que não existe no caso das molas comprimidas: podemos deslocar o ponto P perpendicularmente à direção x (Figura 3.2) e obter um tipo diferente de onda. Como todas as forças das molas agora puxam os pontos, de acordo com a Figura 3.2, a força resultante F_P irá acelerar P de volta à sua posição de equilíbrio O . Os pontos Q e R , por sua vez, estariam sujeitos a forças resultantes que os acelerariam para cima, numa direção essencialmente perpendicular a x .

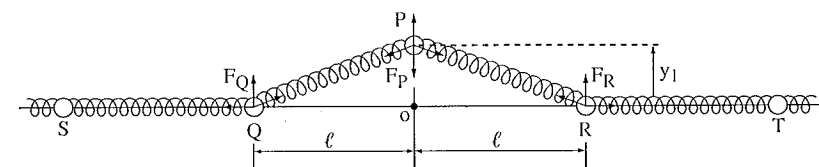


Figura 3.2 Modelo unidimensional de um meio elástico (molas em expansão), onde o ponto P foi deslocado transversalmente.

Isso representa uma onda elástica transversal propagando-se para a direita e para a esquerda de P . Numa onda transversal, os deslocamentos dos pontos são perpendiculares à direção de propagação. Num meio sob tensão, como uma corda de violino, dois modos de propagação de ondas elásticas podem ocorrer simultaneamente: o transversal e o longitudinal.

Vamos agora voltar à expressão da velocidade de propagação de ondas transversais. Aplicando-se a lei de Newton aos pontos individuais do modelo unidimensional da Figura 3.2, pode-se mostrar que, para uma corda sob uma tensão T (em newtons), a velocidade V_T das ondas elásticas transversais é dada por:

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

d é a “densidade linear” do meio, *i.e.*, massa por unidade de comprimento (kg/m). Observe que, quanto mais tensa estiver a corda, mais rapidamente as ondas transversais se propagarão. Da mesma forma, uma maior densidade da corda implica uma menor velocidade de propagação das ondas.

Existe uma relação fisicamente equivalente para a velocidade de propagação de ondas longitudinais num meio de densidades δ (em kg/m³), e onde a pressão seja p (em newtons/m²):

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

Para um gás ideal, entretanto, resulta que a razão p/δ é proporcional à temperatura “absoluta” t_A , definida em termos das temperaturas centígrada ou Fahrenheit t_C e t_F pela simples transformação:

$$t_A = 273 + t_C = 273 + \frac{5}{9}(t_F - 32) \quad (\text{graus Kelvin}) \quad (3.5)$$

Observe que, no ponto de congelamento ($t_C = 0^\circ\text{C}$, $t_F = 32^\circ\text{F}$), a temperatura absoluta é $t_A = 273^\circ$. Embora o ar comum não seja 100% um “gás

ideal”, ele se comporta aproximadamente como tal, e a velocidade das ondas sonoras pode ser expressa por

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

Isto passa a ser 331,5 m/s (= 1 087 pés/s) a 0°C (32°F) e 334 m/s (= 1 130 pés/s) a 21°C (70°F). O fator numérico em (3.6) é válido apenas para o ar. Em geral, o seu valor depende da composição do meio através do qual o som se propaga. Para o hidrogênio puro, por exemplo, ele é igual a 74.0. Assim, as ondas sonoras se propagam quase quatro vezes mais rápido no hidrogênio do que no ar. Isso leva a efeitos acústicos muito divertidos se uma pessoa fala ou canta após ter inalado hidrogênio (NÃO FUME!).

O som é rápido, mas não infinitamente rápido. Isso, por exemplo, leva a pequenas mas perceptíveis diferenças de tempo de chegada entre as ondas sonoras provenientes de diferentes instrumentos numa grande orquestra e pode causar sérios problemas de sincronização rítmica. Um pianista que, pela primeira vez, toca num grande órgão, onde o teclado está muito longe dos tubos, pode inicialmente se sentir confuso pela chegada retardada do som, fora de sincronismo com seus dedos. A reverberação de uma sala está baseada na superposição de ondas sonoras atrasadas que sofreram múltiplas reflexões nas paredes (Seção 4.7).

Vamos agora considerar uma corda muito longa onde o ponto inicial é posto em vibração com um movimento harmônico simples e continua a vibrar indefinidamente, sob a ação de uma força externa. Após certo tempo, verifica-se que todos os pontos da corda vibram com o mesmo movimento harmônico simples. Se, num certo instante, o ponto inicial estiver, digamos, no seu deslocamento máximo, os seus vizinhos ou já terão estado lá ou estarão se encaminhando para estar. A Figura 3.3 mostra os deslocamentos transversais de todos os pontos da corda num certo instante. Esta curva é um “instantâneo” da forma da corda durante a passagem de uma onda senoidal transversal. O gráfico da Figura 3.3 não deve ser confundido com a curva vista na Figura 2.4, que representa a história temporal de apenas um certo ponto. Esta última mostra um padrão de vibração no tempo, a outra mostra um padrão de onda no

espaço. A menor distância entre dois pontos quaisquer da corda que estejam vibrando paralelamente (vibrando “em fase”, *i.e.*, tendo deslocamentos idênticos y em qualquer instante) é chamada de comprimento de onda. Ele em geral é designado pela letra grega λ . Alternativamente, o comprimento de onda pode ser definido como a distância mínima depois da qual o padrão espacial de onda se repete. Compare isso com a definição de período, que representa o intervalo de tempo mínimo depois do qual o padrão de vibração de um certo ponto se repete. (Figura 2.3b).

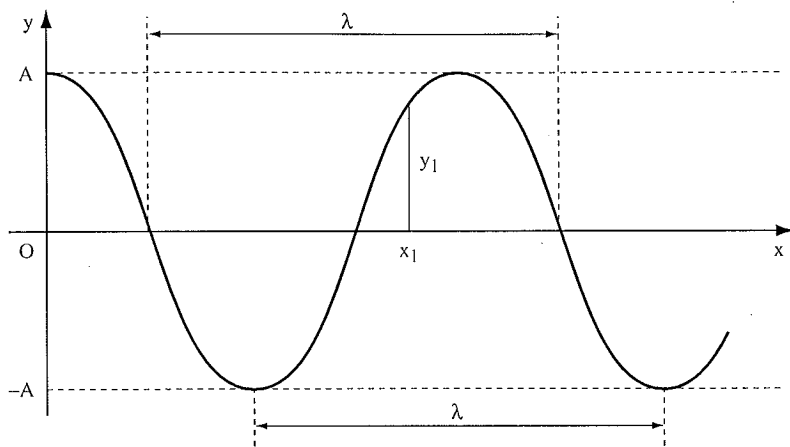


Figura 3.3

Com o passar do tempo, o instantâneo parece se mover com a velocidade da onda para a direita (Figura 3.4) — embora cada ponto da corda se mova apenas para cima e para baixo (por exemplo, considere o ponto x_1 na Figura 3.4). O que se move para a direita é a configuração, *i.e.*, a forma real da corda, mas não a corda propriamente. Em outras palavras, o que se move para a direita é uma qualidade; por exemplo, a qualidade de estar na posição de “deslocamento máximo” (p. ex., os pontos P, Q, R , na Figura 3.4), ou a qualidade de estar passando pela posição de “equilíbrio” (pon-

tos S, T, U). E, obviamente, o que se move para a direita, além disso, é a energia, a energia potencial e cinética envolvida nas oscilações para cima e para baixo dos pontos da corda.

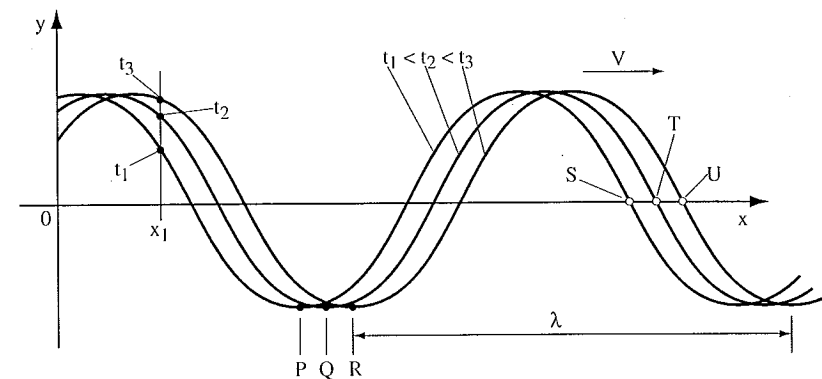


Figura 3.4

Existe uma relação importante entre a velocidade V de uma onda senoidal, o seu comprimento de onda λ e a frequência f de oscilação dos pontos individuais. Considerando a Figura 3.3, vemos que a onda terá se movido exatamente um comprimento de onda λ durante o tempo que o ponto inicial (ou qualquer outro) leva para realizar uma oscilação completa, *i.e.*, durante um período τ . Assim, podemos escrever a velocidade da onda da seguinte forma:

$$V = \frac{\text{distância percorrida}}{\text{tempo gasto}} = \frac{\lambda}{\tau}$$

Como o inverso do período é igual à frequência f (relação 2.1), também podemos escrever

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

Essa relação constitui-se no elo quantitativo entre a “representação espacial” da Figura 3.3 e a “representação temporal” da Figura 2.4. A relação (3.7) nos possibilita exprimir o comprimento de onda de uma onda transversal numa corda em termos da frequência de oscilação dos pontos individuais e da velocidade de propagação (3.3):

$$\lambda = \frac{1}{f} \sqrt{\frac{T}{d}} \quad (3.8)$$

É interessante observar que as relações (3.3) e (3.8) também podem ser aplicadas, até um certo grau, à membrana basilar, substituindo-se a tensão T por um parâmetro de rigidez apropriado. Como a rigidez diminui da base para o ápice, segundo um fator de aproximadamente 10 000, de acordo com (3.3) e (3.8), a velocidade local de propagação e o comprimento de onda das ondas na membrana basilar que têm uma certa frequência diminuirão segundo um fator de 100 enquanto se propagam para o ápice. Também a frequência de ressonância da membrana basilar é proporcional à raiz quadrada do parâmetro de rigidez. Considerações de energia mostram que, quando uma onda se propaga, a sua amplitude aumenta (a energia “se acumula” porque a onda se atrasa). Quando a onda atinge a região de ressonância, a amplitude chega ao máximo e a dissipação de energia estará num pico, sendo a onda rapidamente amortecida adiante desse ponto. A Figura 3.5 mostra esquematicamente como a onda gerada por um som de frequência única se propaga ao longo da membrana basilar. Podemos antecipar que, quando dois ou mais sons puros chegam ao ouvido (como os sons musicais reais), surgirão “pacotes de ondas” separados, como o que se vê na figura, definindo diferentes regiões de ressonância, uma para cada componente de frequência (veja Figura 2.25). Em termos hipersimplificados, eis como funciona o mecanismo de análise de frequências hidromecânico. Lembremo-nos que a Figura 3.5 representa uma onda transversal: os pontos individuais vibram para cima e para baixo, mas o padrão da onda (e a energia associada) propaga-se da esquerda para a direita, com a amplitude das oscilações individuais permanecendo dentro do “envelope” da onda. Observe a diminuição do comprimento de onda quando a onda avança na

direção de ressonância máxima. Todos os pontos da membrana basilar, inclusive os que estão bem fora da região de ressonância, oscilam com a mesma frequência do som puro original. Finalmente, quando mudamos a frequência desse som para cima ou para baixo, todo o quadro da Figura 3.5 será deslocado em direção à base ou ao ápice, respectivamente³.

No caso de ondas longitudinais, como as ondas sonoras no ar, os pontos vibram numa direção paralela à direção de propagação, e não é tão fácil retratar a sua posição real numa forma visual. Por essa razão, as ondas sonoras são representadas mais convenientemente como oscilações de pressão. A Figura 3.6 (parte inferior) mostra os deslocamentos dos pontos de um modelo unidimensional do meio, quando este é percorrido por uma onda longitudinal. Observe que os pontos se encontram em acumulação máxima (*i.e.*, pressão máxima) e em rarefação máxima (*i.e.*, pressão mínima) em posições onde o seu deslocamento é zero (pontos P e Q , respectivamente). Por outro lado, em lugares onde os deslocamentos são máximos, as variações de pressão são zero. Isso significa que as variações de pressão de uma onda sonora estão defasadas em 90° em relação à oscilação dos pontos: as variações máximas de pressão (tanto aumentos como diminuições) ocorrem em lugares onde o deslocamento dos pontos é zero; de modo semelhante, os deslocamentos máximos dos pontos ocorrem em lugares onde as variações de pressão são zero.

Uma onda sonora senoidal é tal que a pressão em cada ponto oscila harmonicamente em torno do valor “normal” (não-perturbado) (Figura 3.7). Num ponto como A , todos os pontos do meio estão mais próximos uns dos outros (aumento máximo de pressão, pontos P na Figura 3.6); num ponto como B , eles se afastaram uns dos outros (diminuição máxima de pressão, pontos Q na Figura 3.6). A variação média de pressão, Δp , é igual à

3. É assim que surge a “frequência característica” de uma fibra neural no nervo acústico: considere um neurônio ligado às células acústicas localizadas na posição A da Figura 3.5. A sua resposta estará relacionada à amplitude da oscilação local da membrana basilar. À medida que a frequência do som de teste cresce gradualmente, desde um valor bem baixo, todo o padrão visto na Figura 3.5 se deslocará da direita para a esquerda: quando a envoltória da oscilação passar pelo ponto A , a frequência de disparos desse neurônio chegará ao máximo (a frequência característica) e subitamente decairá quando o padrão se afastar para a esquerda.

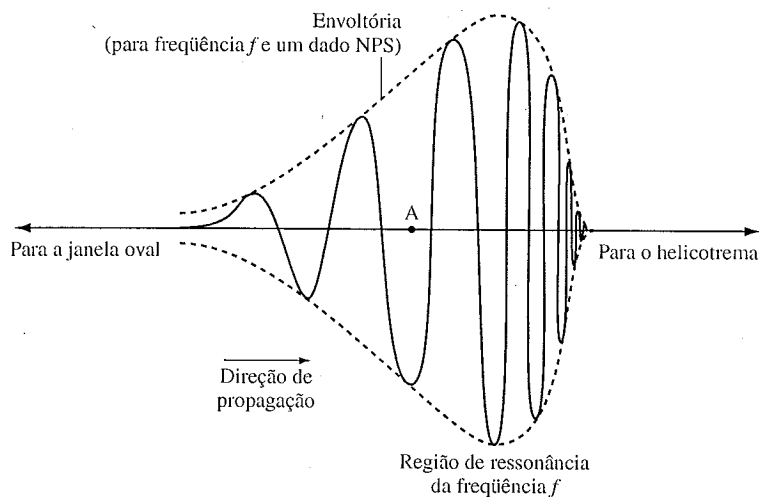


Figura 3.5 Esquema de uma onda progressiva na membrana basilar, gerada por um som de frequência única. Curva cheia: instantâneo dos deslocamentos transversais da membrana (sem escala!). Imagine em sua mente esta curva propagando-se dentro do limite das linhas pontilhadas, da esquerda para a direita, e se retardando à medida que sua amplitude decai na direita. Linha pontilhada: envelope (ou envoltória) de amplitude, que permanece fixo, a menos que haja alguma alteração na frequência ou amplitude do som.

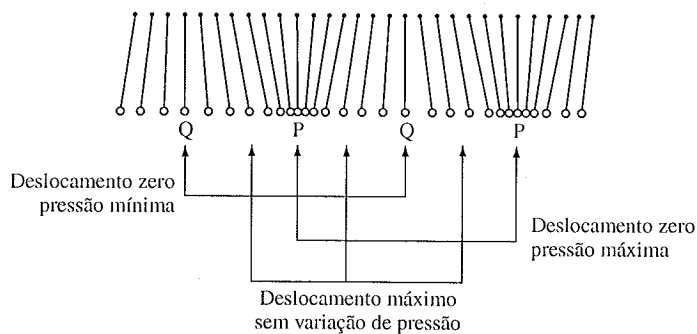


Figura 3.6 Onda longitudinal num meio unidimensional. Para mostrar os deslocamentos reais, cada ponto está representado como um pequeno pêndulo.

amplitude de variação de pressão dividida por $\sqrt{2}$ ($\approx 1,41$). Levando em conta as relações (3.6) e (3.7), obtemos, para o comprimento de onda de ondas senoidais no ar:

$$\lambda = \frac{20,1}{f} \sqrt{t_A} \quad (\text{em m}) \quad (3.9)$$

t_A é a temperatura “absoluta” dada por (3.5). Valores típicos de comprimentos de onda em temperatura normal são vistos na Figura 3.8.

As ondas elásticas podem ser transmitidas de um meio para outro — por exemplo, do ar para a água, do ar para uma parede e de novo para o ar, de uma corda para uma placa de madeira, e daí para o ar que a circunda. A natureza da onda pode mudar em cada transição (p. ex., a transição de uma onda transversal na corda e na placa para uma onda longitudinal no ar). Entretanto, em cada transição, a frequência permanece invariável. O comprimento de onda, por outro lado, variará de acordo com a relação (3.7): $\lambda = V/f$. Nessa relação, V muda de meio para meio, enquanto f é regida exclusivamente pela vibração inicial (fonte).

Quando uma onda elástica atinge a fronteira entre dois meios, parte dela é refletida de volta ao meio inicial. Algumas fronteiras são refletores quase perfeitos (p. ex., paredes de massa fina, para ondas sonoras; as extremidades fixas de uma corda tensa, para ondas transversais). Esse fenômeno é controlado pelo fato de que, na fronteira refletora, os pontos ao meio são forçados a permanecer em repouso, perturbando assim o balanço de forças elásticas que “comandam” a propagação das ondas. Numa reflexão, a frequência permanece inalterada, enquanto a direção de propagação se inverte, no caso de uma incidência perpendicular (ou, em geral, a direção é tal que o ângulo de reflexão é igual ao ângulo de incidência). Também a amplitude deveria permanecer a mesma se não houvesse absorção.

Finalmente, consideramos o fluxo de energia associado a uma onda sonora. Nós o definimos como a quantidade de energia mecânica total (potencial e cinética, associadas com as oscilações elásticas dos pontos do meio) que é transferida a cada segundo através de uma superfície de área unitária (1 m^2) perpendicular à direção de propagação (Figura 3.9). Esse

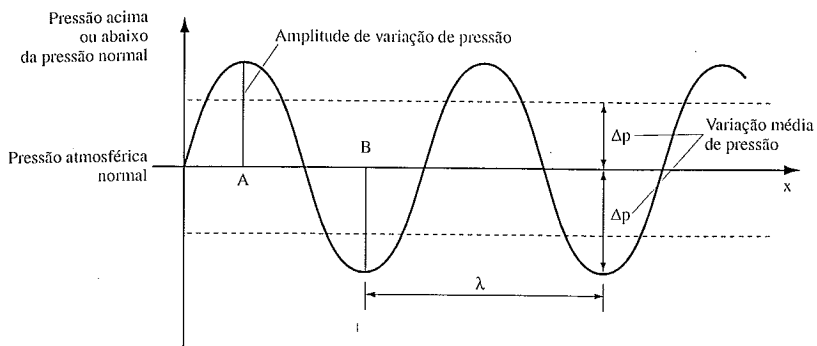


Figura 3.7

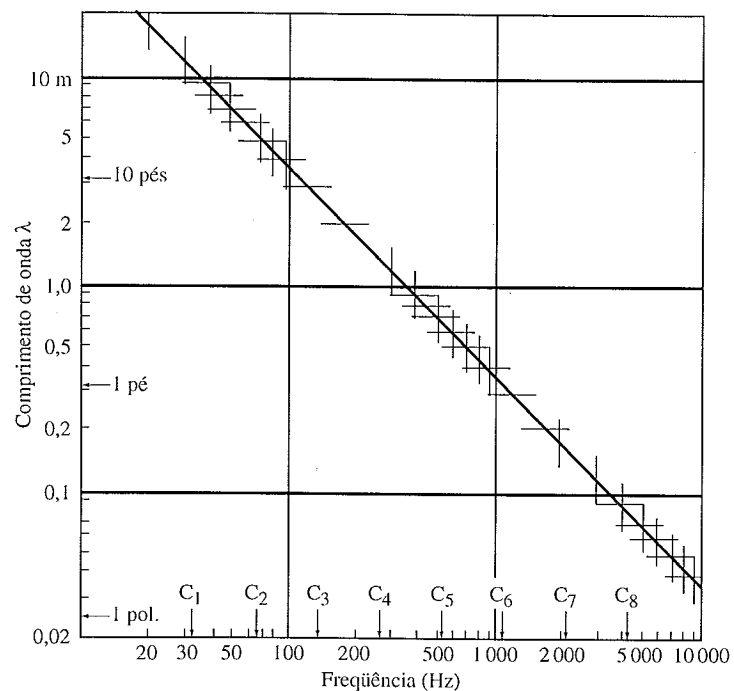


Figura 3.8 Comprimento de onda de uma onda sonora no ar em temperatura normal, em função da frequência (escala logarítmica).

fluxo de energia é expresso em joule/m²s, ou, levando em conta a definição e as unidades de potência (3.2), em watt/m². Ele é mais comumente chamado de intensidade da onda, e representado pela letra *I*. Pode-se mostrar que existe uma relação entre a intensidade de uma onda sonora senoidal e o valor da oscilação média de pressão associada à onda (ver Figura 3.7), que representamos por Δp (igual à amplitude de variação de pressão dividida $\sqrt{2}$):

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

Nessa relação, *V* é a velocidade da onda sonora (3.6) e δ é a densidade do ar. Para condições normais de temperatura e pressão, temos a seguinte relação numérica:

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

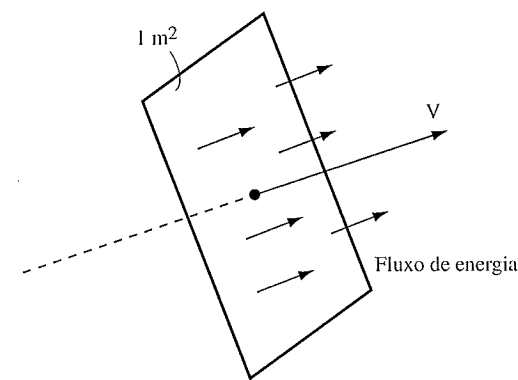


Figura 3.9

Δp deve ser expresso em N/m². Conforme veremos na Seção 3.4, o som puro mais “piano” que pode ser ouvido na frequência de 1000 Hz tem uma

intensidade de apenas 10^{-12} watts/m². De acordo com a relação (3.10), isso representa uma variação média de pressão de apenas $2,0 \times 10^{-5}$ N/m², *i.e.*, apenas $2,0 \times 10^{-10}$ vezes o valor da pressão atmosférica normal! Isso dá uma idéia de quão sensível é o ouvido.

Uma determinada fonte sonora (um instrumento musical ou um alto-falante) emite ondas sonoras em todas as direções. Em geral, a quantidade de energia emitida por segundo depende da direção em particular considerada. Seja I_1 a intensidade da onda no ponto A_1 com a onda se propagando ao longo da direção vista na Figura 3.10. Isso significa que uma quantidade de energia $I_1 a_1$ flui através da superfície a_1 por segundo. Se assumirmos que nenhuma energia é perdida nesse percurso, a mesma quantidade de energia fluirá por segundo através da superfície a_2 no ponto A_2 . Então

$$I_1 a_1 = I_2 a_2$$

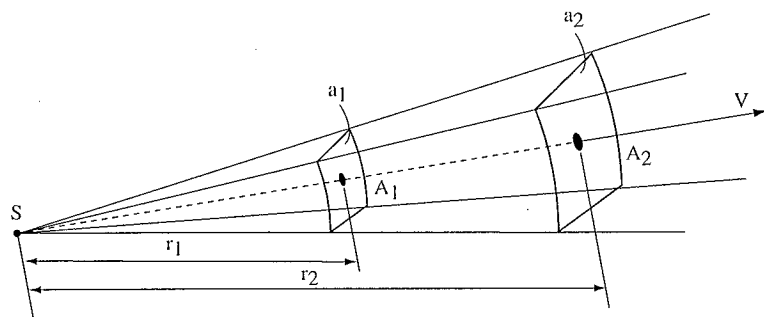


Figura 3.10

Como as áreas das superfícies a_1 e a_2 são proporcionais aos quadrados das suas respectivas distâncias r_1 e r_2 à fonte, a intensidade de uma onda sonora é inversamente proporcional ao quadrado da distância à fonte:

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

Esta lei já não será válida se levarmos em conta as reflexões e absorção sonoras.

Se imaginarmos que a fonte sonora está inserida numa esfera, a quantidade total de energia que fluirá por segundo através da superfície dessa esfera é chamada de potência efetiva da fonte. Ela representa a razão com que a fonte emite energia em todas as direções na forma de ondas sonoras. O seu valor é dado em watts (joule/s). Os instrumentos musicais irradiam entre 0,01 W (clarinete) e 6,4 W (trombone tocando fortíssimo). Um bom alto-falante *hi-fi* pode emitir até 150 W na forma de ondas sonoras.

3.3 SUPERPOSIÇÃO DE ONDAS; ONDAS ESTACIONÁRIAS

Na ausência de paredes refletoras, as ondas sonoras propagam-se em linha reta a partir da fonte. Como foi visto na seção anterior, a sua intensidade decresce rapidamente, proporcional a $1/r^2$, onde r é a distância à fonte. Se temos mais de uma fonte, as ondas emitidas por cada uma delas irão se propagar individualmente como se não existisse nenhuma outra onda, e o efeito resultante num ponto qualquer do meio (por exemplo, no canal auditivo) será uma oscilação de pressão, que é simplesmente dada pela soma algébrica das oscilações de pressão das ondas individuais⁴. Em outras palavras, as ondas sonoras se superpõem linearmente⁵. Isso ocorre, inclusive, com as ondas que se propagam na membrana basilar, para sons de baixa intensidade, dando origem a duas regiões de ressonância independentes entre si (duas envoltórias desse tipo são vistas na Figura 3.5).

4. Observe cuidadosamente que o que está sendo somado aqui são variações de pressão, e não os valores absolutos da pressão!
5. Não é verdade para ondas sonoras muito intensas (potentes), como as de uma explosão.

Vamos considerar a superposição de duas ondas sonoras puras de frequências f_1 e f_2 e, de acordo com a relação (3.7), de comprimentos de onda $\lambda_1 = v/f_1$ e $\lambda_2 = v/f_2$, propagando-se na mesma direção. Para obtermos um “instantâneo” das variações de pressão resultantes, só precisamos somar os valores das variações de pressão individuais, causadas por cada onda separadamente em cada ponto x ao longo da direção de propagação. Como a velocidade das ondas sonoras não depende da frequência (nem do padrão de vibração como um todo), todos os pontos do meio repetirão exatamente o mesmo padrão de vibração complexo — sujeitos apenas a um sincronismo diferente. O fluxo de energia — *i.e.*, a intensidade da superposição de duas (ou mais) ondas propagando-se na mesma direção com a mesma fase — é simplesmente a soma das contribuições ao fluxo de energia das componentes individuais;

$$I = I_1 + I_2 + I_3 + K \tag{3.12}$$

Um caso particularmente importante é o de duas ondas senoidais com a mesma frequência e mesma amplitude propagando-se em direções opostas. Isso, por exemplo, acontece quando uma onda senoidal é refletida num certo ponto (sem absorção) e retorna em sentido contrário, superpondo-se a si mesma. Vamos primeiro considerar as ondas transversais numa corda (Figura 3.11). Somando as contribuições de cada componente, obtemos uma outra senoidal com a mesma frequência, mas com amplitude diferente. Todavia, o fato inusitado é que esta onda resultante não se propaga! Ela permanece presa a certos pontos N_1, N_2, N_3, \dots chamados nodos, que não vibram. Os pontos entre os nodos vibram com amplitudes diferentes, dependendo das suas posições. Em particular, os pontos A_1, A_2, A_3, \dots (a meio caminho entre os nodos), chamamos antinodos*, vibram com amplitude máxima (o dobro da amplitude de cada onda componente). A Figura 3.12 mostra as sucessivas formas de uma corda quando duas ondas senoidais de mesma amplitude se propagam em direções opostas. Isso é o que se chama de onda estacionária. Os pontos oscilam, mas não há evidência

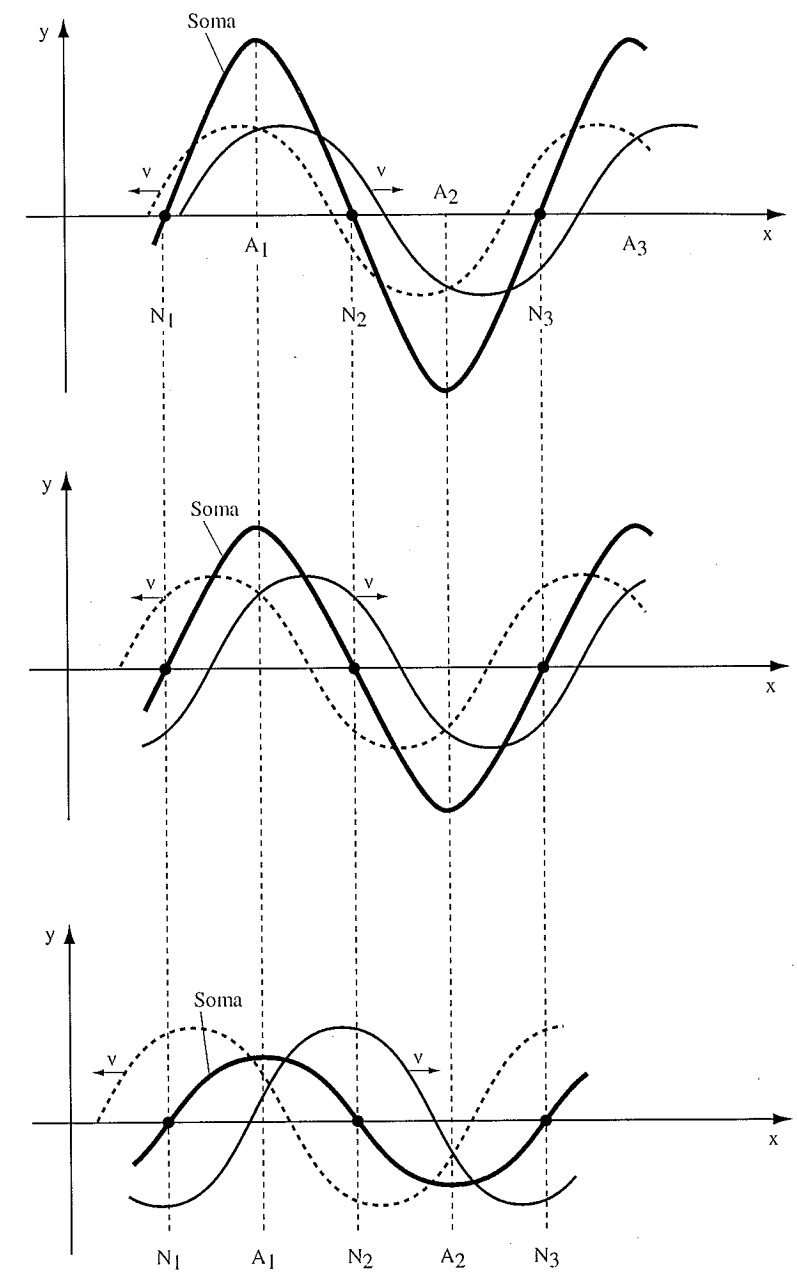


Figura 3.11

* Os antinodos são normalmente chamados “ventres” (N. da T.).

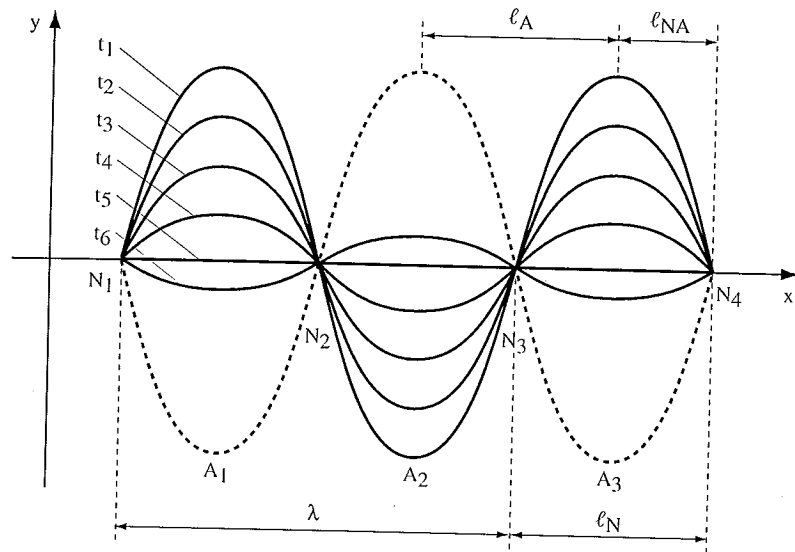


Figura 3.12 Formas sucessivas de uma corda numa oscilação de onda estacionária.

em qualquer situação intermediária (ambas as formas de energia). Além disso, numa onda em propagação, todos os pontos têm a mesma amplitude; o que varia são os tempos em que esse deslocamento máximo é atingido (os pontos estão fora de fase).

Um exame cuidadoso da Figura 3.12 mostra que a distância ℓ_N entre dois nós vizinhos, N_1, N_2 , ou a distância ℓ_A entre dois ventres, A_1, A_2 , é exatamente igual a meio comprimento de onda λ :

$$\ell_N = \ell_A \frac{\lambda}{2} \quad (3.13)$$

Por outro lado, a distância ℓ_{NA} entre um nó N_1 e um ventre A_1 é um quarto de comprimento de onda:

$$\ell_{NA} = \ell_A \frac{\lambda}{4} \quad (3.14)$$

As ondas estacionárias também podem ser longitudinais. Elas surgem quando duas ondas sonoras de mesma frequência e mesma amplitude de variação de pressão se propagam em direções opostas. Isso acontece, por exemplo, quando uma onda sonora se propaga ao longo de um tubo e é refletida na outra extremidade; as ondas estacionárias também surgem de reflexões nas paredes das salas e em outros ambientes. Elas têm as mesmas propriedades das ondas estacionárias transversais, e a discussão acima também se aplica aqui. Existe, todavia, uma importante observação a ser feita. Conforme foi ressaltado na seção anterior, as ondas sonoras podem ser convenientemente descritas como oscilações de pressão. Mostramos lá que os pontos com variação máxima de pressão têm deslocamento longitudinal zero (Figura 3.6), enquanto lugares com variação de pressão zero correspondem a pontos com deslocamento máximo. Podemos traduzir isso para o caso de uma onda sonora estacionária: nós de pressão (*i.e.*, pontos onde as variações de pressão são permanentemente zero) são ventres de vibração (pontos que oscilam com amplitude máxima), enquanto ventres de pressão (pontos onde a pressão

de propagação. O perfil da onda varia em amplitude, mas não se move nem para a direita, nem para a esquerda. Num certo instante (t_1) a corda mostra uma deformação máxima; em outro (t_5), ela não apresenta nenhuma deformação. Conforme veremos no próximo capítulo, as ondas estacionárias exercem um papel de alta importância na música, em especial nos mecanismos de geração sonora dos instrumentos musicais.

Numa onda estacionária também não há propagação resultante de energia. A corda toda age quase como uma mola vibrante: num certo momento (p. ex., t_5 na Figura 3.12), todos os pontos estão passando pela sua posição de equilíbrio, e a energia da corda toda está na forma cinética (energia de movimento). Num outro instante (p. ex., t_1 na Figura 3.12), todos os pontos estão com seu deslocamento máximo e a energia é toda potencial. Em outras palavras, numa onda estacionária todos os pontos oscilam em fase. Observe atentamente o fato de que isso não acontece com uma onda propagante: na Figura 3.3, por exemplo, num certo instante há pontos que apresentam deslocamento máximo (apenas energia potencial), assim como pontos com deslocamento nulo (apenas energia cinética), ou pontos

oscila com amplitude máxima) são nodos de vibração (pontos que estão sempre em repouso).

3.4 INTENSIDADE, NÍVEL SONORO E VOLUME

Na Seção 2.3, estabelecemos que, para um som puro, a amplitude das oscilações do tímpano leva à sensação de volume. Essa amplitude está diretamente relacionada à variação média de pressão, Δp , da onda sonora de entrada e, portanto, ao fluxo de energia acústica ou intensidade I que atinge o ouvido (relação 3.10). Vamos começar aqui investigando a gama de intensidades I de ondas sonoras puras à qual o ouvido é sensível. Existe dois limites de sensibilidade para um som de frequência definida:

1. um limite inferior ou limiar de audibilidade, representando a mínima intensidade audível;
2. um limite superior de audibilidade, além do qual surge a dor fisiológica, levando eventualmente a um comprometimento físico do mecanismo de audição. Verifica-se que esses dois limites variam de indivíduo para indivíduo e depende da frequência particular em consideração.

Em geral, para uma nota com cerca de 1 000 Hz (uma altura entre si_5 e $dó_6$), temos o maior intervalo entre os limites. A enorme gama de intensidades compreendidas entre os dois limites de audibilidade é algo de se admirar. De fato, para uma nota com 1 000 Hz, verifica-se que a intensidade-limite média está perto de 10^{-12} watt/m², enquanto o limite da dor é cerca de 1 watt/m². Isso representa uma razão de intensidades de 1 trilhão para 1, às quais o ouvido é sensível! A Tabela 3.1 mostra uma razão de intensidade sonora e sensação de volume musical, para uma nota de 1 000 Hz⁶:

6. É arbitrário considerar que a notação musical representa uma medida "absoluta" de intensidade. Os músicos, por exemplo, argumentarão que somos perfeitamente capazes de per-

Tabela 3.1

Intensidade (watt/m ²)	Volume
1	Limite da dor
10 ⁻³	fff
10 ⁻⁴	ff
10 ⁻⁵	f
10 ⁻⁶	mf
10 ⁻⁷	p
10 ⁻⁸	pp
10 ⁻⁹	ppp
10 ⁻¹²	Limite de audibilidade

Em 1 000 Hz, a gama de intensidades de interesse musical vai de 10^{-9} até 10^{-2} W/m². Isso ainda representa uma variação da ordem de 10 milhões!

Devido a esta tremenda extensão, a unidade W/m² não é prática. Existe ainda outra razão pela qual ela não é prática. A diferença no limite do observável (dlo) de certo estímulo é, em geral, um bom "calibre" físico a ser levado em conta quando tivermos que escolher uma unidade apropriada para a magnitude física correspondente. Os experimentos mostram que a dlo em intensidade de nota é aproximadamente proporcional à intensidade da nota. Assim, essa proporcionalidade sugere que a "unidade" apropriada deve crescer gradualmente, à medida que cresce a intensidade da nota que queremos descrever. Isso, obviamente, pode nos levar a uma terrível complicação, a menos que introduzamos uma magnitude diferente, que seja uma função apropriada da intensidade I . Essa nova magnitude deve satisfazer três objetivos simultâneos:

ceber fortísimos e pianísimos numa música tocada por um rádio cujo botão de volume esteja quase no mínimo. O que acontece nesse caso é que usamos outras pistas além da intensidade para fazer julgamentos subjetivos do volume "relativo". Por outro lado, experimentos sistemáticos (Patterson, 1974) revelaram que a interpretação da notação de intensidades musicais num contexto dinâmico real depende muito do instrumento e da faixa de alturas envolvidas.

1. uma “compressão” de toda a escala de intensidades audíveis numa gama de valores bem menor;
2. o uso de valores relativos (por exemplo, relativos ao limiar de audibilidade) em vez de valores absolutos;
3. a introdução de uma unidade mais conveniente, cujo valor represente bem de perto a variação mínima perceptível de intensidade sonora.

A introdução da nova quantidade é feita da seguinte maneira. Observe, na Tabela 3.1, que o que parece mais relacionado ao efeito de volume é o expoente ao qual o número 10 é elevado quando citamos o valor da intensidade sonora (coluna da esquerda): -12 para o limiar de audibilidade, -9 para um som ppp, -7 para piano, -5 para forte, -3 para forte-fortíssimo, e 0 para o limite da dor ($10^0 = 1$). Isso é uma forte sugestão que devemos usar o que em matemática se chama de função logarítmica para representar a intensidade.

O logaritmo decimal de um certo número é o expoente ao qual devemos elevar o número 10 para que possamos obter esse número. Por exemplo, o logaritmo de 100 é 2 porque $10^2 = 100$; o logaritmo de 10 000 é 4 porque $10^4 = 10\,000$; o logaritmo de 1 é zero, porque $10^0 = 1$, e o logaritmo de 0,000001 é -6 porque $10^{-6} = 0,000001$. Essas relações são escritas simbolicamente: $\log 100 = 2$; $\log 10\,000 = 4$; $\log 1 = 0$; $\log 0,000001 = -6$. Para qualquer número intermediário entre potências inteiras de 10, o logaritmo pode ser encontrado com o uso de tabelas.

Uma propriedade importante é que o logaritmo do produto de dois números é a soma dos logaritmos dos números individuais, p. ex., o logaritmo do número $10^4 \times 10^3$ é 4 + 3 (= 7), porque $10^4 \times 10^3 = 10^4 + 3 = 10^7$. Em geral, para dois números quaisquer a e b , temos a relação $\log (a \times b) = \log a + \log b$. Para o logaritmo da divisão a/b , temos $\log (a/b) = \log a - \log b$.

Os logaritmos decimais podem ser usados para definir uma magnitude mais apropriada para descrever a intensidade sonora. Em primeiro lugar, adotamos o limiar de audibilidade como nossa intensidade de referência I_0 (10^{-12} W/m² em 1 000 Hz). Introduzimos então a quantidade

$$NS = 10 \times \log \frac{I}{I_0} \tag{3.15}$$

Esta é chamada de nível de intensidade sonora. A unidade de NS é chamada de decibel, “db”. Para o limiar de audibilidade, $I/I_0 = 1$, e, portanto, $NS = 0$ db. Para o limite superior de audibilidade, $I/I_0 = 10^{12}$ e $NS = 10 \times \log 10^{12} = 120$ db. Um som “forte” (Tabela 3.1) tem um nível de intensidade sonora de 70 db; ppp corresponde a 30 db.

É importante observar que, quando uma quantidade é expressa em decibéis, o que se apresenta é uma medida relativa, em relação a algum valor de referência (p. ex., o limiar de audibilidade na definição de NS). Sempre que a intensidade I for multiplicada por um fator 10, apenas acrescentaremos 10 db ao valor de NS; quando a intensidade for multiplicada por 100, acrescentaremos 20 db, e assim por diante. Da mesma forma, quando a intensidade for dividida por 100, subtraímos 20 db de NS. A Tabela 3.2 apresenta algumas relações úteis.

Tabela 3.2

Variação em NS			O que acontece com a intensidade		
some	(subtraia)	1 db	multiplique	(divida) por	1,26
+	(-)	3 db	×	÷	2
+	(-)	10 db	×	÷	10
+	(-)	20 db	×	÷	100
+	(-)	60 db	×	÷	1 000 000

Podemos usar a relação (3.10) para exprimir a intensidade em termos da variação média de pressão, Δp , que é mais facilmente mensurável. Verificamos que o limite mínimo I_0 em 1 000 Hz corresponde mais ou menos à variação média de pressão $\Delta p_0 = 2 \times 10^{-5}$ N/m² (= 20 micropascals). Como, de acordo com a relação (3.10), I é proporcional ao quadrado de Δp , temos

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

Assim, podemos introduzir a quantidade

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

chamada nível de pressão sonora. Para uma onda em propagação, os valores numéricos de (3.15) e (3.16) são idênticos, e NS e NPS representam exatamente a mesma coisa. Para ondas estacionárias, todavia, não há fluxo de energia (Seção 3.3) e a intensidade I usada em (3.16) não pode ser definida; assim, NS perde o seu significado. Porém, o conceito de variação média de pressão Δp num certo ponto do espaço (p. ex., na entrada do canal auditivo) ainda conserva o seu significado, bem como o nível de pressão sonora. Eis por que a relação (3.16) é usada mais freqüentemente do que a relação (3.15). Observe atentamente que as definições de NS e NPS absolutamente não envolvem a freqüência da onda sonora. Embora tenhamos feito referência a uma nota de 1 000 Hz, nada nos impede de definirmos NS e NPS pelas relações (3.15) e (3.16), respectivamente, para qualquer freqüência. O que depende da freqüência, e muito, são os limites subjetivos da audibilidade (p. ex., I_0 e p_0) e, em geral, a sensação subjetiva de volume, conforme veremos mais adiante.

Coisas muito engraçadas parecem acontecer com o nível sonoro, ou nível de pressão sonora, quando superpomos dois sons de mesma freqüência (e fase). Considere a Tabela 3.2: somar duas notas de mesma intensidade, o que significa dobrar a intensidade, segundo a relação (3.12), implica somar meros 3 db ao nível sonoro do som original, qualquer que possa ser esse valor de NS. Superpor dez sons iguais (em fase) aumenta apenas em 10 db o NS resultante. Para aumentar o NS de um som em 1 db, precisamos multiplicar a sua intensidade por 1,26, o que significa que precisamos acrescentar um som cuja intensidade seja 0,26 (cerca de 1/4) da do som original.

A variação mínima necessária no NPS para resultar numa mudança detectável na sensação de volume (dlo em nível sonoro) é aproximadamente constante, e da ordem de 0,2–0,4 db nas gamas musicalmente relevantes de altura e volume. A unidade de NS ou NPS, o decibel, é uma quantidade razoável — próxima à dlo.

Existe uma forma alternativa de se encarar a dlo de intensidade ou nível sonoro. Em vez de perguntarmos o quanto devemos mudar a intensidade de certa nota para obtermos um mínimo efeito perceptível, podemos colocar a seguinte questão equivalente: qual é a intensidade mínima I_2 que uma segunda nota de mesma freqüência e fase deve ter para que possa ser notada na presença da primeira (cuja intensidade I_1 é mantida constante)? Essa intensidade mínima I_2 é chamada de limiar de mascaramento. A nota inicial, de intensidade constante I_1 , é chamada de “nota mascaradora”, e a outra é chamada de “nota mascarada”. O mascaramento tem um papel decisivo em música. Neste parágrafo, mencionamos o mascaramento de notas com freqüências (e fases) iguais à da nota mascaradora; mais adiante vamos discutir o mascaramento em freqüências diferentes. A relação entre o nível de mascaramento NM (NS da nota mascarada no limite) e a dlo de nível sonoro pode ser encontrada usando-se a relação (3.15) e levando-se em conta as propriedades dos logaritmos. Por exemplo, para uma dlo de 0,2 db, obtemos um nível de mascaramento NM que é 13 db menor do que o NS da nota mascaradora; uma dlo de 0,4 db corresponde a um NM, que é 10 db mais baixo.

Até agora, temos lidado com as quantidades físicas NS e NPS. Vamos agora examinar a magnitude psicofísica “volume”, associada com um certo NPS. Nas Seções 1.4 e 2.3, mencionamos a habilidade das pessoas de estabelecerem uma ordem para a intensidade de duas sensações do mesmo tipo, ressaltando que surgem complicações quando se tem que fazer comparações quantitativas absolutas. No caso do volume, os julgamentos de se dois sons puros soam igualmente intensos e apresentam uma dispersão razoavelmente baixa entre pessoas diferentes. Mas os julgamentos de “quanto” um som tem que ser mais intenso do que outro exigem um treinamento ou condicionamento anterior e fornecem resultados que fluam bastante de indivíduo para indivíduo.

Em geral, os sons com mesmo NPS mas com freqüências diferentes são julgados como tendo volumes diferentes. O NPS, portanto, não é uma boa medida de volume, se intercomparamos notas de freqüências diferentes. Têm-se realizado experimentos para estabelecer curvas de volume igual, tomando o NPS em 1 000 Hz como uma quantidade de

referência⁷. Estas curvas são vistas na Figura 3.13. (Fletcher e Munson, 1933). Partindo do eixo vertical centrado em 1 000 Hz e seguindo em ambas as direções de freqüências mais baixas e mais altas, respectivamente, são traçadas curvas que correspondem aos NPSs das notas que são julgadas como tendo “o mesmo volume” que a nota de referência de 1 000 Hz. Observe, por exemplo, que, enquanto um NPS de 50 db (intensidade de 10^{-7} W/m²) em 1 000 Hz é considerado “piano”, o mesmo NPS é quase inaudível em 60 Hz. Em outras palavras, para produzir certa sensação de volume em baixas freqüências, digamos um som “forte”, será necessário uma intensidade (fluxo de energia) muito maior do que em 1 000 Hz. Eis por que as notas do baixo parecem sumir muito antes do que os agudos quando gradualmente nos afastamos de uma fonte sonora fixa. Ou por que temos que pagar muito mais caro por um equipamento *hi-fi*, principalmente pelos alto-falantes, se quisermos baixos bem balanceados.

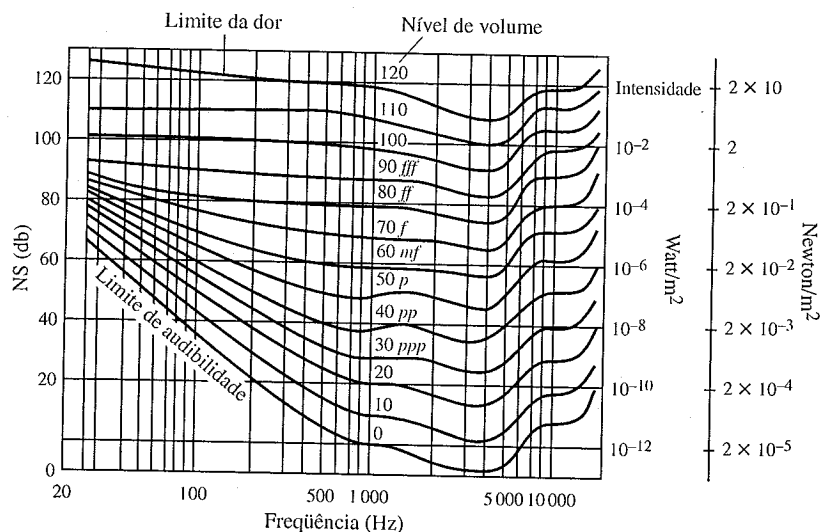


Figura 3.13 Curvas de volume igual (Fletcher e Munson, 1933) num diagrama de nível sonoro (NS) e freqüência. As escalas correspondentes de intensidade sonora e variação média de pressão também são apresentadas.

7. Obtido através de experimentos de “casamento de volumes”, conduzidos de forma análoga aos experimentos de casamento de alturas.

A curva mais baixa na Figura 3.13 representa o limiar de audibilidade para diferentes freqüências. Mais uma vez, ela mostra como a sensibilidade do ouvido decresce consideravelmente em direção às baixas freqüências (e também em direção às muito altas). A sensibilidade máxima é conseguida por volta de 3 000 Hz. A forma desta curva-limite é influenciada pelas propriedades acústicas do canal auditivo (meato) e pelas propriedades mecânicas da cadeia de ossos no ouvido médio. Em particular, a alta sensibilidade em torno dos 3 000 Hz deve-se ao fato de que o canal auditivo tem uma ressonância e forma ondas estacionárias nessa freqüência. Finalmente, devemos enfatizar que as curvas na Figura 3.13 só são válidas para sons puros individuais e que soem continuamente. Vamos discutir mais adiante o que acontece à sensação de volume se o som tiver curta duração (menos de 1 segundo). Estudos recentes (Molino, 1973) mostram que os contornos de volume igual aparentemente dependem da freqüência do som de referência (que foi 1 000 Hz na Figura 3.13).

Agora chegamos a um item às vezes confuso. Uma nova quantidade é introduzida, chamada nível de volume, *NV*. Ela é definida da seguinte maneira. O *NV* de um som de freqüência *f* é dado pelo NPS de um som de 1 000 Hz, que é julgado como igualmente forte. Isso significa que as curvas da Figura 3.13 são curvas de nível de volume constante. A unidade de *NV* é chamada *fon*. A Figura 3.13 pode ser usada para se encontrar o *NV* de uma nota de NPS determinado, em qualquer freqüência *f*. Por exemplo, considere uma nota com 70 db de NPS ($I = 10^{-5}$ W/m²) a 80 Hz. Vemos que a curva que passa por esse ponto intercepta a linha de 1 000 Hz em 50 db. Portanto, o *NV* dessa nota é igual a 50 fons. Em geral, os números vistos ao longo da linha de 1 000 Hz representam o *NV* em fons das curvas de volume constante correspondentes.

Observe atentamente que o *NV* ainda é uma magnitude física, e não psicofísica (a despeito do nome). Ele representa aquelas intensidades ou NPSs que soam igualmente intensas, mas não pretende representar o volume percebido de uma maneira absoluta: uma nota cujo *NV* é duas vezes maior que outra não soa com o dobro do volume! Muitos estudos têm sido feitos para se determinar uma escala subjetiva de volume. A Figura 3.14 (linha cheia) é o resultado (Stevens, 1955), relacionando o “volume subjetivo” *V* com o

nível de volume NV , no intervalo de interesse musical. A quantidade V que descreve o volume subjetivo é expressa em unidades chamadas sons. Observe que a relação não é linear (a escala de volume na Figura 3.14 é o que se chama de escala logarítmica). Ela é tal que, aumentando-se o NV em 10 fons, o volume V simplesmente dobra. Isso significa, por exemplo, que dez instrumentos tocando uma certa nota no mesmo NV são ouvidos soando apenas duas vezes mais forte do que um dos instrumentos tocando sozinho!

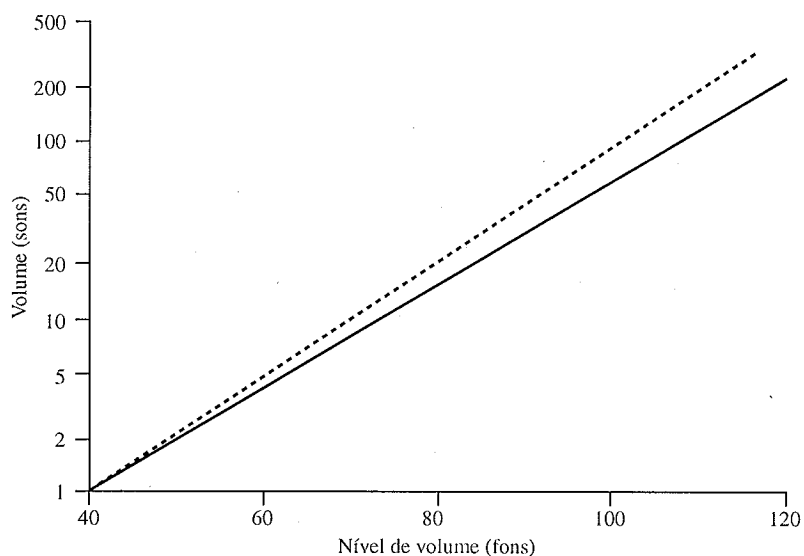


Figura 3.14 Linha cheia: relação experimental entre a magnitude psicológica volume e a magnitude física nível de volume (segundo Stevens, 1955). Linha pontilhada: relação de potência 3.17 (Stevens, 1970).

Já foi demonstrado que a relação entre V e a intensidade de onda I , ou a variação média de pressão Δp , pode ser descrita aproximadamente pela função simples (Stevens, 1970):

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

onde C_1 e C_2 são parâmetros que dependem da frequência. Isso resulta na linha pontilhada vista na Figura 3.14, que se encontra bem dentro da flutuação estatística das medições reais (não mostrada). Note que a relação logarítmica quase desapareceu. Todavia, ainda permanece uma “compressão” apreciável da escala de volume subjetivo: para variar V entre 1 a 200, a intensidade I deve variar num fator de 8 milhões.

Quando superpomos duas ou mais notas de mesma frequência (e mesma fase), a nota resultante tem uma intensidade (fluxo de energia) que é a soma das intensidades das notas componentes: $I = I_1 + I_2 + I_3 + \dots$ (3.12). Como, neste caso, as notas individuais não podem ser discernidas uma da outra, essa intensidade total determina o volume resultante pela relação (3.17). Obviamente, V não será igual à soma dos volumes das notas individuais. Situações diferentes surgem quando as ondas componentes têm frequências diferentes. Podemos identificar três regimes:

1. Se as frequências das notas componentes estiverem todas dentro da banda crítica da frequência central (Seção 2.4), o volume resultante ainda estará diretamente relacionado à intensidade total (fluxo de energia), soma das intensidades individuais:

$$V = C_1 \sqrt[3]{(I_1 + I_2 + I_3 + K)} \quad (3.18)$$

Esta propriedade, na verdade, leva a uma determinação mais precisa da banda crítica (Zwicker, Flottorp e Stevens, 1957) do que a dada na Seção 2.4.

2. Quando o espalhamento de frequência do estímulo de várias notas excede a banda crítica, o volume subjetivo resultante é maior do que o obtido por simples soma de intensidades (3.18), tendendo para um valor que é dado pela soma das contribuições individuais das bandas críticas adjacentes para o volume (Zwicker e Scharf, 1965):

$$V = V_1 + V_2 + V_3 + K \quad (3.19)$$

Os efeitos de mascaramento devem ser levados em conta se os volumes individuais v_1 , v_2 etc. diferirem consideravelmente entre si. O limite de integração do volume (3.19) nunca é atingido na prática.

3. Quando a diferença de frequência entre as notas individuais for grande, a situação será mais complicada. Em primeiro lugar, surgem dificuldades com o conceito de “volume total”. As pessoas tendem a se concentrar em apenas uma das notas componentes (p. ex., a mais forte, ou a mais aguda) e atribuem a sensação de volume total apenas àquela componente⁸:

$$V = \left. \begin{array}{l} \text{máximo} \\ \text{ou mais alto} \end{array} \right\} \text{ de } (V_1, V_2, \dots) \quad (3.20)$$

Tudo isso tem importância para a música. Por exemplo, dois tubos de órgão do mesmo tipo e de mesma frequência soam apenas 1,3 vez mais forte do que um tubo sozinho. Quando as suas frequências diferirem em um tom inteiro ou um semitom, o seu volume ainda será aproximadamente 1,3 vez o de um tubo sozinho (um tom inteiro ou um semitom estão dentro da banda crítica, Figura 2.13). Mas duas notas que formam o intervalo de terça maior soarão mais forte do que a combinação anterior. Esses fatos, há séculos, têm sido bem conhecidos pelos construtores de órgãos e compositores. Como não há possibilidade de um controle manual de volume para as notas individuais dos órgãos, assim como ocorre nos instrumentos de cordas ou de sopro, o volume no órgão só pode ser alterado mudando-se o número de registros que soam simultaneamente. Mas como, de acordo com o parágrafo anterior, a soma dos volumes é mais efetiva

8. Uma situação bem conhecida, semelhante a esse fato, surge com a sensação de dor. Se você for picado em dois lugares muito próximos um do outro, a dor pode ser “duas vezes” mais forte que a de uma única picada (equivalente ao caso 1 acima). Mas, quando os lugares são separados, você tem dificuldade em perceber o que se chama de “dor total” (caso 2). Na verdade, você tende a se concentrar apenas no estímulo que estiver dando a sensação mais intensa de dor. Muito provavelmente, a técnica chinesa de anestesia por acupuntura é baseada nesse fato.

quando as notas componentes diferem bastante em frequência (relação 3.19), os registros que soam uma (4'), duas (2'), ou mais oitavas acima da nota escrita (e abaixo (16')) são usados principalmente para esse fim⁹. Por outro lado, sem acrescentar ou retirar registros, o volume também pode ser controlado pelo número de notas tocadas simultaneamente. Cada nova voz que entra numa fuga aumenta o volume subjetivo da peça, e cada nota adicional num acorde cumpre o mesmo propósito. Alguns organistas tocam os acordes finais de uma fuga de Bach acrescentando alguns registros. Isso é um absurdo — o próprio Bach programou o aumento desejado de volume simplesmente escrevendo mais notas do que o número de vozes usado ao longo da fuga!¹⁰

Examinamos a soma de volumes de duas ou mais notas superpostas, mas ainda não discutimos o que acontece com o limiar de audibilidade de uma nota quando ela é ouvida em presença de outra. Se as frequências das duas notas coincidem, este limiar é dado pelo nível de mascaramento discutido anteriormente (p. 133). Se as suas frequências forem diferentes, ainda podemos determinar um nível de mascaramento, definido como o nível sonoro mínimo que deverá ser excedido pela nota mascarada, para que esta possa ser “selecionada” e ouvida individualmente na presença da nota mascaradora. O limiar de intensidade de sons puros isolados (curva inferior na Figura 3.13) varia apreciavelmente, *i.e.*, aumenta, se outros sons estiverem simultaneamente presentes. A experiência mais familiar de mascaramento é aquela de alguém não ser capaz de acompanhar uma conversa na presença de uma grande quantidade de ruído de fundo. O nível de mascaramento NM de um som puro com frequência f na presença de outro som puro com características fixas (frequência 415 Hz e nível sonoro NS) é mostrado pela Figura 3.15 (Egan e Hake, 1850). O nível NM atingido pela nota mascarada acima do limiar de audibilidade normal (segundo a curva inferior da Figura 3.13) está indicado para diferentes valores de

9. A sua soma, obviamente, também contribui para uma mudança de timbre (Capítulo 4).
10. Muitos órgãos de seu tempo possuíam um registro chamado *Zimbelstern*, que fazia soar pequenos pratos ou sinos bem agudos posicionados numa estrela rotativa no alto da caixa do órgão; esse registro era usado para reforçar a intensidade de um acorde final (sem interferir de nenhum modo na sua harmonia).

NS da nota mascaradora. As regiões próximas a f_0 devem ser extrapoladas para o valor deduzido da dlo de volume (p. 133); os fenômenos de batimentos têm um importante papel aí, o que não tem nada a ver com o mascaramento em si. Num NS maior surgem outras complicações devidas ao aparecimento de harmônicos aurais nas frequências $2f_0$, $3f_0$ etc. (Seção 2.5). Observe a assimetria das curvas em NS superiores (causada por esses harmônicos aurais): uma nota de frequência f_0 mascara com mais eficiência as frequências mais altas do que as mais baixas.

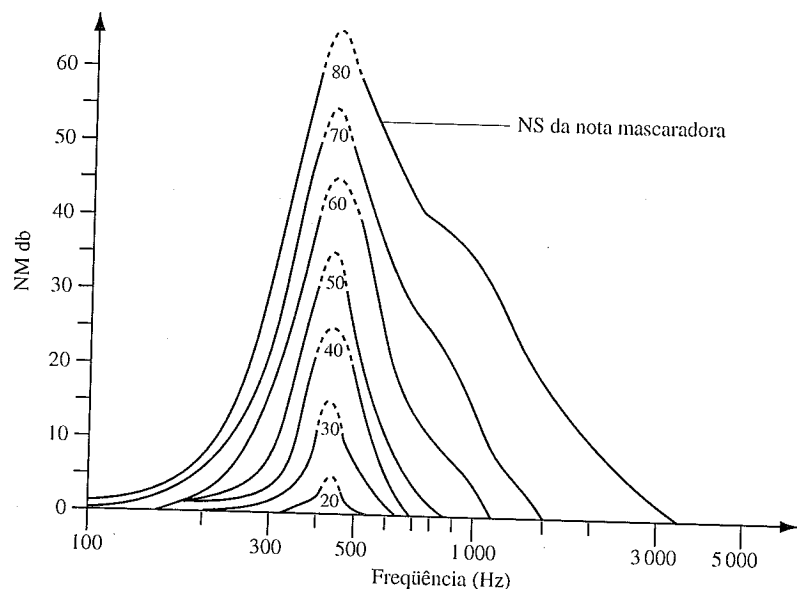


Figura 3.15 Nível de mascaramento correspondente a um som puro de 415 Hz, para vários valores de nível sonoro (Egan e Hake, 1950) do som mascarador.

O mascaramento tem um papel importante na música polifônica, particularmente na orquestração. Em muitas ocasiões, nas partituras, a participação de um determinado instrumento, como um oboé ou um fagote, pode ser totalmente irrelevante, se ele tocar ao mesmo tempo que os metais estão tocando um fortíssimo. Da mesma forma, acrescentar registros de flauta ou

registros suaves do tipo cordas a um *tutti* de diapasões, misturas e madeiras num órgão é completamente irrelevante, do ponto de vista do volume.

Finalmente, devemos mencionar o efeito da duração de um som sobre a sensação de volume (chamada de integração temporal). Em primeiro lugar, existe um limite de tempo, uma duração mínima que um determinado som puro deve ter para transmitir uma sensação de som. Essa duração mínima é de aproximadamente 10-15 milissegundos, ou pelo menos 2-3 períodos de oscilação, se a frequência for menor que 50 Hz; sons que duram menos que isso são percebidos como “estalidos”, e não como “sons”. Sons que duram mais de 15 milissegundos (ou 2-3 períodos, o que for mais longo) podem ser individualizados como sons de altura e volume definidos¹¹. Todavia, o volume subjetivo depende da duração do som (enquanto a altura não) (Plomp e Bouman, 1959). Quanto mais curto for o pulso sonoro, mais baixo será o seu volume, se a intensidade de som física (fluxo de energia) se mantiver constante (p. ex., Richards, 1977).

A Figura 3.16 mostra um gráfico do decréscimo relativo de volume, ou atenuação de volume, em função da duração do som, para diferentes frequências. Observe que o valor “final” da resposta é atingido primeiro para as frequências mais altas; após aproximadamente meio segundo, o volume atinge um valor constante que depende apenas da intensidade (relação dada na Figura 3.14). Também o mascaramento (Figura 3.15) tem características de dependência temporal significativas e bem complicadas para sons mascaradores de curta duração; para maiores detalhes, sugerimos Zwicker e Fastl (1990). Em geral, em primeira aproximação, verifica-se que, para sons curtos, a sensação de volume não está relacionada ao fluxo instantâneo de potência (intensidade da onda sonora), mas à energia acústica total liberada pelo som no ouvido (intensidade \times duração). Na verdade, existem indícios de que, neste caso, a sensação de volume está relacionada ao número total de impulsos neurais que são transmitidos em associação com o som de curta duração (Zwislocki, 1969) (veja a próxima seção).

11. De modo geral, existe um princípio físico que estabelece que a frequência de uma vibração não pode ser definida com mais precisão do que o inverso da duração total da vibração.

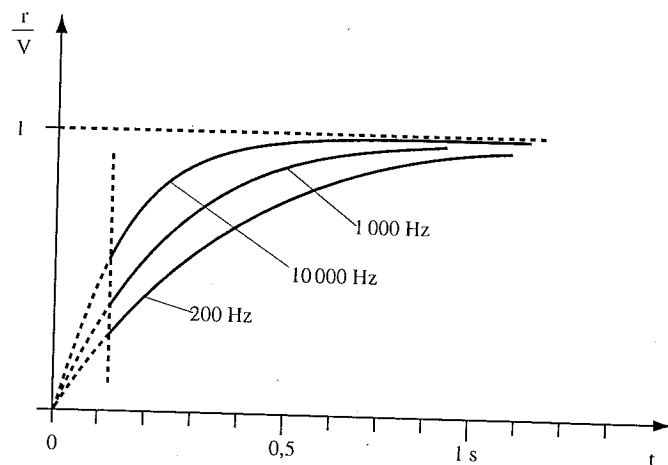


Figura 3.16 Volume relativo de sons puros de curta duração. r/v : razão entre o volume real (r) e o volume (V) de um som contínuo de mesma frequência e amplitude.

Para longos tempos de exposição estabelece-se um efeito chamado *adaptação*, que consiste numa diminuição do volume subjetivo quando um som de intensidade constante for ouvido por vários minutos¹². Embora exista grande diferença entre os indivíduos, os experimentos de casamento de volume revelam alguns aspectos comuns (Scharf, 1983). A adaptação aumenta com a frequência dos sons puros de mesmo NPS (sons de frequência mais alta são “amortecidos” mais depressa); para sons de mesma frequência, a adaptação diminui quando o NPS aumenta (não existe efeito de adaptação quando os NPSs forem acima de uns 40 db). Em geral, para frequências e NPSs musicalmente relevantes o volume subjetivo de um som se nivela depois de uns 100 segundos, permanecendo então constante.

A atenuação de volume para sons curtos tem implicações muito importantes na prática musical. Se quisermos tocar, ao piano, uma passagem *staccato* num certo volume, deveremos pressionar as teclas com mais força do que se fôssemos tocar as mesmas notas *legato* com o mesmo

12. Isso não deve ser confundido com a *fadiga* acústica, um processo psicológico pelo qual nosso cérebro consegue ignorar um som contínuo mais irrelevante.

volume¹³. Esse efeito é muito mais pronunciado com sons que não decaem, como os sons do órgão. Na verdade, o organista tem um considerável controle do volume subjetivo de uma certa nota dando-lhe a duração correta; o fraseado no órgão é a arte do controle dinâmico pela execução de cada nota com a duração apropriada (obviamente, isso só funciona para sons curtos — veja a Figura 3.16). Por outro lado, a adaptação de volume também desempenha um papel em música: os trinados foram inventados para combatê-la! Em contraste, o uso da “nota pedal” (nota do baixo que soa continuamente) é uma prova do fato afirmado acima de que a adaptação não é importante para baixas frequências. Dada a importância desses efeitos temporais para a música, justifica-se uma maior quantidade de pesquisas sobre atenuação e adaptação.

3.5 O MECANISMO DE PERCEÇÃO DO VOLUME E PROCESSOS CORRELATOS

Qual é o processo físico ou neural responsável pela diferença entre a limitada escala subjetiva de volumes e a enorme gama de intensidades detectáveis (Tabela 3.1; relação 3.17) do som original? No caso da percepção primária de alturas (Seção 2.3), já havíamos encontrado essa espécie de “compressão” (sem mencioná-la explicitamente): embora a escala de frequências audíveis varie de uns 20 Hz até cerca de 16 000 Hz, isso corresponde apenas a nove oitavas de extensão. Nesse caso, a compressão é causada principalmente pelas propriedades de ressonância mecânica da partição coclear: a curva na Figura 2.8 realmente representa uma relação aproximadamente logarítmica entre a posição x da região de ressonância na membrana basilar e a frequência.

No caso do mecanismo de detecção de volumes, a “compressão” é em parte neural e em parte mecânica. Na presença de um som puro, os neu-

13. Para notas de instrumentos musicais, existe uma complicação: durante a formação da nota, que pode durar vários décimos de segundo, ocorre na fonte uma variação “natural” de intensidade e espectro. Também os efeitos de reverberação são importantes (Seção 4.7).

neurônios primários conectados às células sensoriais localizadas no centro da região de máxima amplitude de ressonância (*i.e.*, fibras com a mesma frequência característica) aumentam a sua velocidade de disparo acima do nível espontâneo. Esse aumento é uma função monotônica da amplitude do estímulo, embora não seja linear (p. ex., Sachs e Abbas, 1974). De fato, quando esta última aumenta de um fator de, digamos, 100 (um aumento de 40 db no NPS, relação 3.16), verifica-se que a velocidade de disparo aumenta apenas de um fator de 3-4. Outro elemento que contribui para a “compressão” de volume está relacionado com o seguinte: em altos NPSs, a velocidade de disparo de um neurônio primário satura a um nível próximo daquele do disparo espontâneo (sendo que o nível real de saturação varia bastante de neurônio para neurônio). Qualquer aumento extra na intensidade não irá alterar muito essa velocidade de disparo; os neurônios simplesmente não conseguem transmitir pulsos com uma velocidade maior que o valor de saturação (determinado pelo tempo refratário após cada pulso). Fibras individuais do nervo auditivo com frequências características semelhantes possuem limites de disparo muito diferentes (o NPS necessário para aumentar a frequência de disparos acima do espontâneo). O aumento necessário no nível de estímulo para se alcançar a saturação também pode variar de 20 db a 40 db, ou mais, para diferentes fibras. Em consequência, o conjunto dos neurônios auditivos que servem uma determinada área da membrana basilar pode efetivamente cobrir a ampla extensão dinâmica do sistema acústico, através de uma divisão adequada de trabalho. Num estudo sistemático das características individuais dos neurônios auditivos, Liberman (1978) identificou três grupos:

- Fibras com altas frequências de disparo espontâneo (mais de 20 impulsos por segundo) e com os limites mais baixos.
- Fibras com frequências espontâneas muito baixas (menos de 0,5/s) e altos valores-limite ao longo de uma ampla região (até uns 50-60 db).
- Um grupo com limites e frequências espontâneas intermediárias.

Cada grupo de fibras tem características celulares distintas (diâmetro e arquitetura do contato sináptico com as células capilares internas). O fato

interessante é que cada célula capilar interna recebe fibras de todos os três grupos. Isso dá chance a *uma* célula sensorial individual de liberar sua resposta sobre uma ampla gama dinâmica! (E essa pode ser mais uma razão pela qual tantas fibras aferentes fazem contato com cada célula capilar interna). De acordo com esses resultados, a intensidade sonora é codificada tanto pela frequência de disparos como pelo tipo de fibra nervosa acústica que porta a informação.

Contudo, a afirmação anterior deve ser complementada. Quando a intensidade de um som puro aumenta, a amplitude da onda que se propaga na membrana basilar aumenta em todos os lugares, e não apenas na região do pico de ressonância. Isso dá chance aos neurônios cuja frequência característica é diferente daquela da onda sonora que chega (*i.e.*, os neurônios ligados às células capilares que estão próximas, mas não exatamente no local da ressonância) de aumentar as suas frequências de disparo quando os seus limites (fora da ressonância) forem superados (ver Figura 3.17). Em resumo, um aumento na intensidade leva a um aumento do número total de impulsos transmitidos — ou porque a velocidade de disparo de cada neurônio aumentou ou porque aumentou o número total de neurônios ativados. Este último efeito depende principalmente da forma da distribuição de amplitudes de oscilação da membrana — uma propriedade puramente mecânica.

A relação entre o volume subjetivo e a razão total de disparos explica qualitativamente as propriedades principais da soma de volumes (Seção 3.4). Para sons simultâneos com frequências que diferem mais do que a banda crítica, o total dos impulsos neurais transmitidos é mais ou menos igual à soma das velocidades de pulsos evocadas por cada componente em separado; assim, o volume total tenderá a ser a soma dos volumes de cada som (relação 3.19). Por outro lado, para sons cujas frequências se encontram dentro de uma banda crítica, com regiões de ressonância na membrana basilar superpondo-se substancialmente, o número total de pulsos será controlado pela soma das intensidades do estímulo original (relação 3.18).

A dependência da sensação de volume (e dos limites de mascaramento) em relação à duração do som (Figura 3.15) indica que a operação de processamento do sinal acústico apresenta uma estrutura dependente do

tempo. Apenas após vários décimos de segundo é que o mecanismo neural atinge um “estado estacionário” de processamento de som. É importante enfatizar mais uma vez que (felizmente para a música) durante o estabelecimento de uma situação estacionária a sensação de altura é estável e única desde o início (exceto no estágio mais inicial).

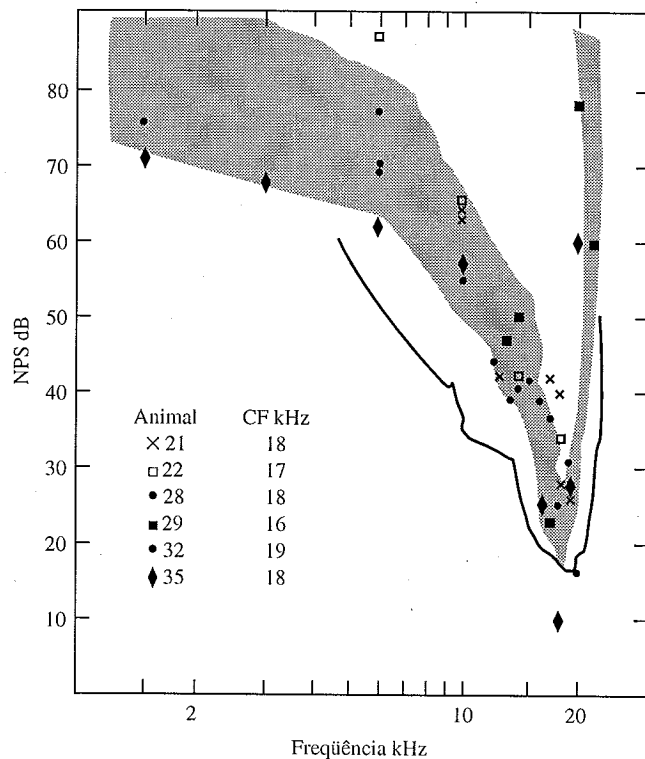


Figura 3.17 Comparação entre membrana basilar (curva cheia) e sintonia neural. A área cinza corresponde a curvas de sintonia de dez fibras nervosas acústicas conectadas à mesma região da membrana basilar (Johnstone e outros, 1983).

Existe, entretanto, uma pequena dependência da altura em relação ao volume, para um som com frequência constante. Para sons acima de uns

2 000 Hz, a altura aumenta quando o volume aumenta e vice-versa; abaixo de 1 000 Hz acontece o oposto (p. ex., Walliser, 1969). Esse efeito é pequeno e varia consideravelmente de pessoa para pessoa; são necessárias grandes variações de intensidade para se chegar a variações perceptíveis na altura¹⁴. Esse efeito, provavelmente, é causado pela assimetria na distribuição da excitação ao longo da membrana basilar (ver Figura 3.5) e pela resposta neural não-linear devida à saturação — uma variação de intensidade provoca uma mudança do ponto central de excitação (mesmo que a frequência permaneça constante), conduzindo a uma alteração na sensação de altura primária. Outro resultado dessa assimetria pode ser o pequeno, mas relevante efeito segundo o qual a altura de um som puro de frequência fixa varia ligeiramente quando outro som de frequência diferente lhe é superposto (p. ex., Walliser, 1969; Terhardt, e Fastl, 1971). Esse efeito pode ter algumas conseqüências relevantes para a entoação musical (Seção 5.4)¹⁵.

Aqui chegamos novamente à questão formulada na página 62 relativa ao mecanismo de detecção de altura primária ou espectral: se um som puro com determinada intensidade e frequência impõe à membrana basilar uma oscilação ressonante cobrindo uma gama espacial finita Δx , e a altura primária está codificada na forma da posição espacial x das fibras ativadas, como se explica que apenas uma única sensação de altura é produzida? Um mecanismo equivalente opera no sistema visual, sendo responsável por importantes efeitos de realce de contrastes (Ratliff, 1972). Nesse sistema, o processo de “depuração” é realizado por uma rede neural (parcialmente na retina) cuja função é concentrar ou “afunilar” a atividade para um número limitado de neurônios circundados por uma região de “quietude” ou inibição neural, ressaltando assim o contraste. Até os

14. O casamento de alturas mostra que, para um aumento de NPS de 40 para 80 dB, a altura de um som de 6 000 Hz aumenta 7%, e a de um som de 150 Hz diminui 3% (Walliser, 1969).
15. Devemos ressaltar que há também uma variação de altura quando a pressão do fluido coclear se modifica (p. ex., variações de altura percebidas durante um bocejo), ou quando a sua composição química se altera (p. ex., injeções de drogas no fluido cérebro-espinhal). Repare que nenhuma dessas variações de altura pode ser explicada por uma teoria que envolva sugestões temporais!

anos 70, acreditava-se que existia no sistema acústico um mecanismo neural equivalente para a depuração de alturas. Contudo, estudos recentes sobre a cóclea revelaram impressionantes propriedades eletromecânicas das células capilares externas, responsáveis por um processo de realimentação não-linear que amplifica e aguça a sintonia auditiva antes da conversão em sinais neurais. Esse processo será visto na próxima seção.

3.6 MÚSICA DOS OUVIDOS: EMISSÕES OTOACÚSTICAS E MECÂNICA COCLEAR

Música dos ouvidos? Bem... não exatamente, mas quase!

A compreensão do mecanismo da audição se desenvolveu basicamente em três épocas principais. A primeira foi dominada pela idéia de Helmholtz, (1863), de que a membrana basilar atuava como um analisador espectral, sustentando mecanicamente ondas estacionárias forçadas externas (a idéia do “pequeno teclado ressonante no ouvido”). A segunda época, de 1940 aos anos 70, teve a predominância dos resultados experimentais de Békési (1960), mostrando que o som que chega provoca uma onda progressiva hidromecânica espacialmente confinada, com a localização da amplitude máxima determinada pela frequência do sinal de entrada (a idéia da “bandeira tremulante”). A época atual começou nos anos 70 com um vasto conjunto de experimentos e estudos teóricos demonstrando que as ondas progressivas de Békési são localmente amplificadas por um processo eletromecânico no qual as células capilares externas, devido à sua motilidade (até então inesperada), funcionam como sensores e como elementos de realimentação mecânica (para uma análise mais abrangente, veja Dallos, 1992). Essa amplificação ciclo-a-ciclo funciona melhor para baixos níveis de sinal, o que explica a alta sensibilidade e gama dinâmica do ouvido. O processo de realimentação pode entrar em auto-oscilação, ou entrar em ressonância após cessar o estímulo externo, promovendo vibrações cocleares no domínio acústico que podem ser captadas como sons fracos por um microfone bem sensível posicionado no canal auditivo externo (Kemp, 1978).

Embora não tenham conseqüências diretas para a música, essas emissões otoacústicas são manifestações de um processo que pode explicar algumas das incríveis, e em parte mutuamente conflitantes, habilidades do sistema acústico: detectar um som que desloca os elementos sensoriais em apenas frações de nanômetros (10^{-9} m); ser sensível a uma gama dinâmica de intensidades de pelo menos nove potências de dez; responder a detalhes temporalmente variantes da ordem de alguns microssegundos; e resolver a frequência de um som com uma precisão bem maior do que a membrana basilar passiva pode permitir. As novas evidências obtidas no nível celular da audição revelam a cóclea como “um triunfo evolutivo de miniaturização [...] o mais complexo equipamento mecânico do corpo humano, com mais de um milhão de partes móveis [...] um amplificador acústico e analisador de frequências compactados no volume de uma bola de gude” (Hudspeth, 1985, 1989).

As medições feitas por Békési do movimento da membrana basilar foram realizadas em cócleas bem preservadas de animais mortos, revelando amplas regiões de ressonância para sons de frequência única. Quando se tornou possível registrar os impulsos neurais das fibras nervosas auditivas de pessoas vivas (ver Seção 2.8), descobriu-se que essas fibras estavam finamente sintonizadas¹⁶ a uma frequência característica, mais do que se poderia esperar de uma “análise pobre de frequências” da membrana basilar, revelando a ação de um processo de depuração que, naquela época, se supôs que fosse mediado pelo sistema neural, como na visão. Medições mais recentes, usando o efeito Mössbauer e técnicas de *laser* em animais vivos, revelaram uma sintonia bem mais refinada da membrana basilar. A Figura 3.17 mostra uma comparação entre membrana basilar e sintonia neural (Johnstone e outros, 1983). A curva sólida representa o NPS da onda sonora incidente, com uma frequência definida, à qual a membrana basilar responde com uma determinada velocidade de pico (0,04 mm/s na figura) num certo ponto onde a fonte radioativa está localizada (isso é chamado de

16. “Sintonia fina” ou “depurada” significa que a faixa de frequências a que o neurônio responde é muito estreita; “sintonia rasa” significa que as respostas ocorrem numa faixa ampla de frequências.

curva de sintonia isovelocidade). O mínimo acentuado em 18 kHz indica que, no ponto em questão, a membrana basilar responde mais aos sinais dessa frequência (são necessários NPSs de apenas 20 db para produzir essa velocidade). Uma oitava abaixo, *i.e.*, 9 kHz, é preciso um sinal 15 db mais forte para provocar a mesma resposta de velocidade. Uma “curva” de sintonia neural, por outro lado (a área cinza na Figura 3.17 define os limites de dez curvas de sintonia normais), representa o NPS mínimo de um sinal de determinada frequência ao qual a fibra responde (p. ex., aumenta a sua frequência de disparos acima do limite). Existe muito mais semelhança entre a membrana basilar e a sintonia neural do que se poderia antecipar pelas primeiras medições em cócleas de animais mortos. Os registros das células capilares internas, feitos com microeletrodos, mostram que as suas curvas de sintonia são muito semelhantes às das fibras nervosas aferentes; isso indica claramente que o processo de depuração nas cócleas vivas deve ocorrer em algum lugar da partição coclear, mais provavelmente no espaço subtectorial (Figura 2.7a).

Técnicas de *laser* e Mössbauer bastante aperfeiçoadas permitem agora medições muito mais precisas do movimento da membrana basilar em animais vivos (p. ex., Johnstone e outros, 1986; Ruggero e Rich, 1991). Os resultados mostram que, sob condições reais, as curvas de sintonia da membrana basilar são realmente muito semelhantes às das células capilares internas ou das fibras neurais associadas ao mesmo local da membrana. Além disso, as medições mostram que a nitidez da curva de sintonia da membrana basilar aumenta dramaticamente quando o NPS do estímulo diminui até o limite; isso revela a atuação de um mecanismo de amplificação não-linear capaz de reforçar a vibração da membrana basilar na região de ressonância em mais de dez mil vezes para os sinais fracos (o mecanismo é “não-linear” porque sua sensibilidade não aumenta na mesma proporção para todas as frequências e amplitudes).

Estudos teóricos mostram que o suprimento de energia para o mecanismo de amplificação não pode ser garantido pela entrada sonora original (p. ex., De Boer, 1983). As propriedades de sensibilidade, seletividade de frequências bem definidas e de resposta não-linear da membrana basilar devem provir da ação de realimentação mecânica ativa das células capilares externas.

Um fato convincente é que essas células são capazes de mudar de forma — principalmente no comprimento — em frequências na faixa de radiofrequências sob estímulo elétrico (Kachar e outros, 1986). É concebível que essas alterações de forma afetem a deformação local propagante induzida pelas ondas na partição coclear, realimentando energia de modo que a onda agregada resultante será amplificada (Dallos, 1992). A motilidade das células capilares externas é produzida pela ação combinada de moléculas protéicas contráteis encontradas nas membranas das células; a sua velocidade de ação é tão incrivelmente alta que supera os mecanismos eletromecânicos como os que se encontram nas fibras musculares normais. Embora pareça certo que a ativação do mecanismo de realimentação coclear seja iniciada localmente pela ativação mecânica das próprias células capilares externas, existe uma grande possibilidade de que esse processo possa ser modificado por comandos neurais do sistema nervoso central via fibras eferentes do feixe olivococlear. Finalmente, dada a saturação da resposta das células capilares externas para altos níveis de estímulo, a ação de realimentação estará limitada a baixos níveis de entrada, dentro de 40 db acima do limiar.

Em resumo, a concepção atual (Dallos, 1992; Allen e Neely, 1992) acerca do processo de realimentação não-linear que reforça a resposta mecânica da partição coclear é a seguinte, para sons de frequência única:

1. A oscilação senoidal do “pistão” na janela redonda (Figura 2.6b) induz uma onda progressiva na membrana basilar que tem amplitude máxima numa posição ao longo da membrana que depende da frequência.
2. A oscilação local é captada pelos *estereocílios* das células capilares externas, um processo que dispara um sinal elétrico em cada célula, provocando uma contração das células motoras em sua membrana. Os sinais neurais do feixe aferente olivococlear podem influenciar nesse processo, de acordo com os comandos do cérebro.
3. O efeito coletivo da resposta das células motoras é uma alteração na forma da célula (provavelmente no comprimento), em sincronia com a oscilação aplicada.

4. Esta reação mecânica em-fase das células capilares externas reforça (hidrodinamicamente, ou por um movimento de “chacoalhar” da membrana tectorial) a oscilação local da membrana basilar.
5. Se algum limite superior for atingido, a reação celular se nivela, ou satura; caso contrário, volta-se ao passo 2!

Questão: por que esse processo atua apenas na região de ressonância máxima da membrana basilar para essa frequência de entrada? Afinal de contas, longe do local de amplitude máxima da onda progressiva ainda há oscilação, embora com amplitude cada vez menor (Figura 3.5)! Em outras palavras, por que o padrão de vibração ao longo da membrana basilar não é amplificado *por inteiro*? Em primeiro lugar, um mecanismo de realimentação não-linear positivo é altamente seletivo: a amplificação será mais eficiente quanto maior for a amplitude da oscilação! Isso o faz seletivo quanto às frequências, porque o fator de amplificação será mais elevado na região de ressonância pertinente. Além deste, entretanto, existe um segundo mecanismo de sintonia atuando. De fato, existem evidências experimentais de que os próprios estereocílios estão sintonizados na frequência de ressonância local e, portanto, responderão eficientemente apenas se a frequência da onda que passa coincidir com a frequência de ressonância local; isso só acontece perto do máximo de ressonância correspondente à frequência de entrada¹⁷.

Outra questão complexa se refere aos efeitos do ruído térmico. Um estereocílio se apóia em sua base e dispara uma resposta elétrica excitadora na célula hospedeira quando ele é desviado numa certa direção (na direção oposta, a resposta será inibidora, enquanto as respostas ortogonais são

17. As pesquisas eletrofisiológicas mostram que as frequências a que as células capilares são mais sensíveis são inversamente proporcionais aos comprimentos dos seus feixes capilares — e isso acontece de modo que esses comprimentos aumentam várias vezes ao longo da membrana basilar, desde a base (extremidade de alta frequência) até o ápice (extremidade de baixa frequência). Há também alguns indícios de que a sintonia pode ocorrer pela ação de uma ressonância elétrica na membrana da célula capilar (Hudspeth, 1985). Tudo isso foi observado em cócleas de vertebrados inferiores; ainda não é claro em que extensão isso também se aplica aos mamíferos superiores e aos seres humanos.

ineficazes — isso explica o disparo relativo à fase realizado pela fibra nervosa auditiva vista na Figura 2.22!). O que é surpreendente é que o limiar de audibilidade ocorre para uma deflexão do cílio de apenas 0,003 graus! Pode-se demonstrar que deslocamentos assim tão pequenos ocorrem mesmo na ausência de qualquer estímulo acústico, devido ao chamado movimento browniano do líquido endolinfático (pequenas frações do líquido oscilam agitadoamente por causa das flutuações do equilíbrio térmico). Mas o ruído aleatório resultante é amenizado pelo fato de que os *estereocílios* estão unidos por filamentos extremamente finos que os faz moverem-se como um feixe de ação combinada! Na verdade, o verdadeiro limite de audibilidade é determinado pelo ponto onde os sons mais tênues são abafados pelo ruído térmico dos componentes do ouvido (Hudspeth, 1989).

Poderão todos os fascinantes processos acima ocorrer numa escala de tempo de frações de microssegundos, como exige o poder de resolução temporal do sistema acústico? Em outras palavras, o equipamento eletromecânico da cóclea poderá realmente ser melhor (temporalmente) do que o equipamento neural? A geração de impulsos elétricos numa célula capilar está baseada na troca de íons (cálcio e potássio) com o meio circundante. Experimentos mostram a existência de poros, que atravessam a membrana, através dos quais os íons podem passar quando o feixe capilar for convenientemente defletido (Hudspeth, 1985). O pulso de voltagem gerado pela corrente elétrica conduzida pelos íons é responsável pela resposta geral da célula; para uma célula capilar externa, isso inclui a contração das moléculas motoras, o que acontece num microssegundo após a aplicação do estímulo. Fica claro que, para satisfazer às incríveis demandas de velocidade no sistema auditivo, a amplificação e detecção de sinais devem ocorrer segundo processos realizados dentro de uma célula: o aparato acústico não pode aceitar o “andamento vagaroso” do sistema nervoso, que opera numa escala de tempo de milissegundos ou mais!

Finalmente, vamos voltar à “música dos ouvidos”. Como foi mencionado acima, as emissões otoacústicas são sons muito tênues que podem ser detectados no canal auditivo bem fechado; o seu nível, em geral, está bem abaixo do limiar da audição. Elas são de dois tipos básicos: emissões

espontâneas e emissões evocadas por estímulos sonoros externos. A Figura 3.18 mostra um exemplo de um espectro de emissão otoacústica espontânea (Zwicker e Fastl, 1990) (NPS zero corresponde ao limiar da audição). As emissões espontâneas variam enormemente de indivíduo para indivíduo. Cerca de 50% das pessoas com audição normal apresentam uma ou mais dessas emissões, e a maior parte dos animais as têm; a não-existência *não* indica uma anormalidade na audição (*tinnitus*, ou “zumbido nos ouvidos”, parece não estar relacionado a essas emissões, na maioria dos casos). A gama de frequências vai de 400 a 4 000 Hz, embora a maior parte ocorra entre 1 000 e 2 000 Hz; para uma determinada pessoa, a sua intensidade pode variar apreciavelmente no intervalo de um dia, ou mesmo ser intermitente em menor escala de tempo, mas as frequências em que elas ocorrem são bastante estáveis em cada cóclea. Emissões otoacústicas provocadas são mais difíceis de se detectar; é

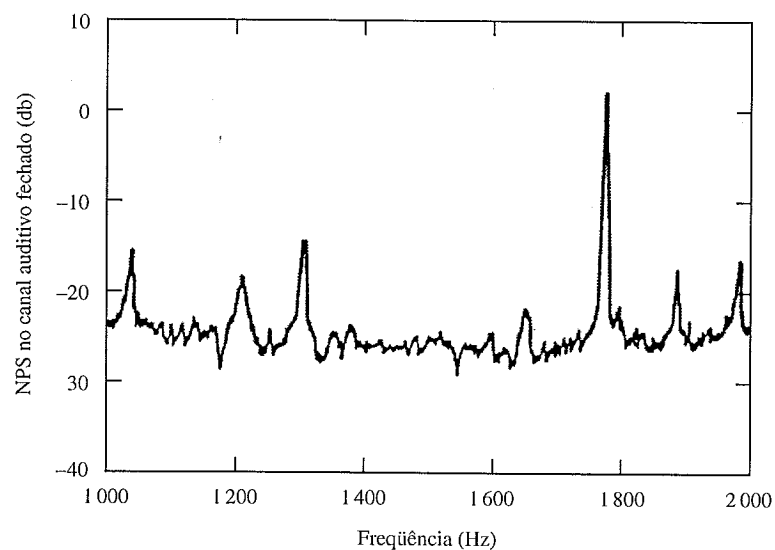


Figura 3.18 Exemplo de um espectro de frequências de emissões otoacústicas (Zwicker e Fastl, 1990) registradas por um microfone sensível posicionado no canal auditivo externo fechado. NPS zero corresponde ao limiar de audibilidade.

necessário um transmissor sonoro na sonda e condições mais rigorosas para a análise das informações do microfone. Elas podem ser divididas em duas classes: as emissões que ocorrem simultaneamente a um som de teste contínuo, e as respostas atrasadas a impulsos sonoros curtos. Não há dúvida de que as emissões otoacústicas são um subproduto do processo de realimentação não-linear descrito acima, mediado pelas células capilares externas; por que elas ocorrem e por que aparecem em frequências selecionadas não se sabe; apenas podemos observar que os circuitos de realimentação ativa em geral não têm capacidade de auto-excitação ou ressonância pós-estímulo. Como as emissões otoacústicas parecem não ter nenhum papel em música (a menos que sejam reproduzidas num volume bem amplificado!), indicamos ao leitor interessado em maiores informações a literatura técnica (p. ex., Zwicker e Fastl, 1990; Gelfand, 1990 e as referências aí contidas).

GERAÇÃO DE SONS MUSICAIS, SONS COMPLEXOS E A PERCEÇÃO DO TIMBRE

Nos dois capítulos anteriores foram analisados os dois atributos principais do som, a altura e o volume, com base principalmente em sons puros de frequência única. Entretanto, esses não são os sons que desempenham papéis ativos em música. A música é constituída por sons complexos, cada um dos quais consistindo em uma superposição de sons puros misturados em uma certa proporção, de tal modo que pareça ao nosso cérebro como todos não-analisáveis. Surge, assim, um terceiro atributo fundamental do som: a qualidade do som, ou timbre, relacionada ao tipo de mistura de sons puros, ou componentes harmônicos, existente num som complexo (Seção 1.2).

A maior parte dos instrumentos musicais gera ondas sonoras por meio de cordas vibrantes ou colunas de ar. No Capítulo 1, nós chamamos esses elementos de elementos vibrantes primários. A energia necessária para sustentar a sua vibração é fornecida por um mecanismo de excitação, e a saída final de energia acústica em muitos instrumentos é controlada por um ressonador. O local onde o instrumento musical está sendo tocado pode ser considerado uma extensão natural do próprio instrumento, exercendo um papel substancial na modelagem do som real que chega até o ouvinte.

Neste capítulo, vamos discutir como os sons musicais reais são realmente produzidos nos instrumentos, como são constituídos por superposições de sons puros, como interagem com o ambiente nas salas e auditórios, e como tudo isso leva à percepção do timbre e do reconhecimento dos instrumentos. O capítulo termina com uma breve revisão das funções cognitivas do cérebro relevantes à percepção de sons complexos.

4.1 ONDAS ESTACIONÁRIAS NUMA CORDA

Vamos considerar o caso de uma corda tensa, presa aos pontos fixos P e Q (Figura 4.1), de comprimento L e massa por unidade de comprimento d , esticada com uma certa força T que pode ser variada à vontade, digamos, alterando-se a massa do corpo suspenso na corda, como mostra a figura. Em seguida, dedilhamos ou percutimos a corda num certo ponto. Dois pulsos ondulatórios elásticos transversais vão se propagar para a esquerda e para a direita, afastando-se da região de perturbação inicial, na forma discutida na Seção 3.2, com uma velocidade dada pela relação 3.3. Esses pulsos ondulatórios serão refletidos ao atingirem os pontos fixos P e Q ; um pulso positivo, ou “para cima”, retornará negativo, ou “para baixo”, e vice-versa. Após certo tempo (extremamente curto, devido à velocidade das ondas numa corda tensa), haverá ondas se propagando simultaneamente para a frente e para trás, em ambas as direções da corda. Em outras palavras, teremos energia ondulatória elástica “presa” na corda entre P e Q , sendo que esses dois pontos estarão sempre em repouso. Se não houvesse perdas, essa situação iria permanecer assim para sempre, e a corda continuaria a vibrar indefinidamente. Todavia, a fricção e as perdas através de P e Q eventualmente dissiparão a energia armazenada e as ondas decairão¹. Por enquanto, vamos ignorar essa dissipação.

1. É esta “razão de vazamento” de energia pelos pontos “fixos” (especialmente na ponte) que é transformada em potência sonora na caixa de ressonância de um instrumento de cordas.

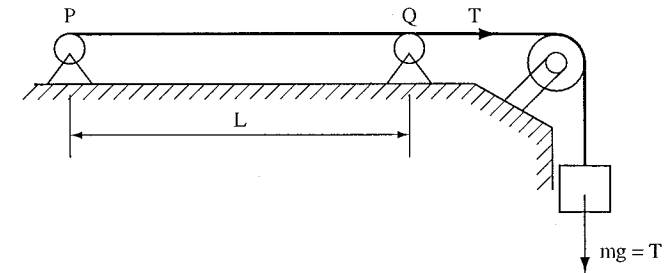


Figura 4.1

Em vista da discussão na Seção 3.3, imaginamos que o quadro acima, de ondas propagando-se para a frente e para trás ao longo de uma corda, assemelha-se bastante à situação que surge com uma onda estacionária. Na verdade, pode-se mostrar matematicamente que as ondas estacionárias são a única forma estável possível de vibração em uma corda com as extremidades fixas, com esses pontos P e Q exercendo o papel de nodos.

Isso tem uma conseqüência muito importante. Entre todas as formas imagináveis de ondas estacionárias, só são possíveis aquelas cujos nodos estejam em P e Q . Em outras palavras, são permitidas apenas aquelas ondas estacionárias senoidais que “caibam” um número inteiro de vezes entre P e Q (Figura 4.2), *i.e.*, para as quais o comprimento da corda L é um múltiplo inteiro da distância entre nodos, ℓ_N , dada pela relação (3.13). Levando em conta essa relação, obtemos a condição $L = n \ell_N = n \lambda/2$, onde n é um número inteiro qualquer 1, 2, 3, ... Isso nos diz que são permitidos apenas os seguintes comprimentos de ondas (Figura 4.2):

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

Usando a relação (3.8), verificamos que uma corda só pode vibrar com as seguintes freqüências:

$$f_n = \frac{1}{\lambda_n} \sqrt{\frac{T}{d}} = \frac{n}{2L} \sqrt{\frac{T}{d}} = n f_1 \quad (4.2)$$

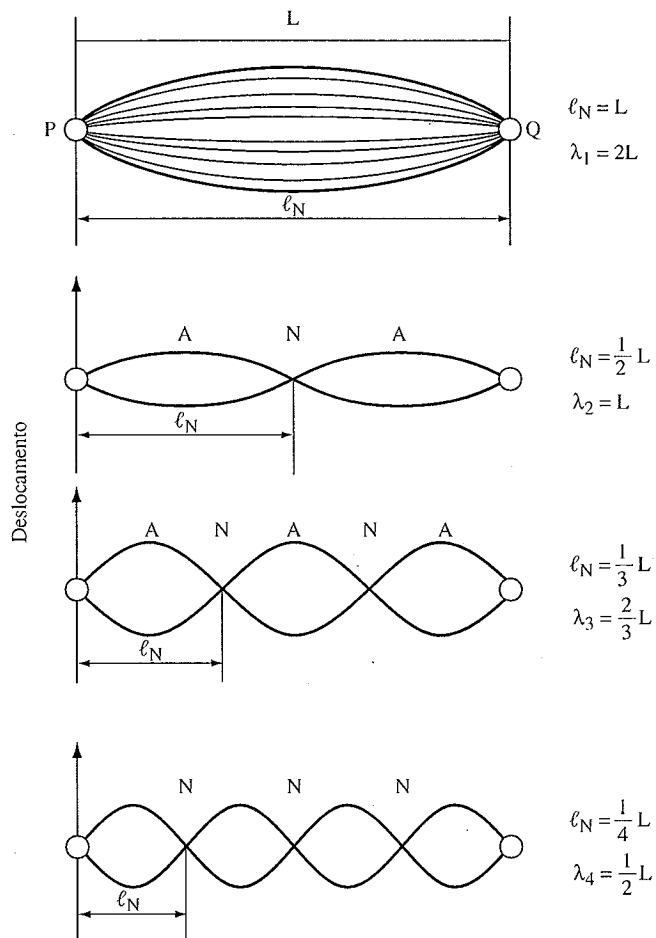


Figura 4.2 Modos da onda estacionária numa corda vibrante.

A frequência mais baixa possível é obtida para $n = 1$:

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

Esta é chamada de frequência fundamental da corda. Observe em (4.2) que todas as outras frequências possíveis são múltiplos inteiros da fre-

quência fundamental. Elas são chamadas de harmônicos superiores de f_1 (Seção 2.7). Observe, em particular, que o primeiro harmônico ($n = 1$) é idêntico à frequência fundamental; o segundo harmônico, f_2 , é a oitava acima de f_1 ; o terceiro harmônico é a décima segunda (uma quinta acima da oitava); o quarto harmônico é a décima quinta (duas oitavas) etc. (Figura 4.3). Os harmônicos superiores também são chamados de sobretons².

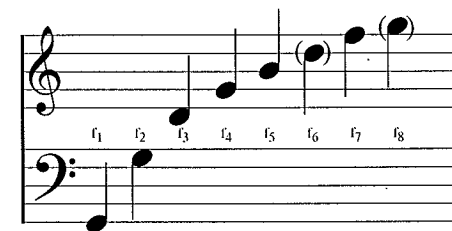


Figura 4.3 Os primeiros oito harmônicos de $f_1 = 98$ Hz (sol₂). (As notas entre parênteses não correspondem exatamente a f_6 e f_8).

A relação (4.3) nos diz que a frequência fundamental de oscilação de uma corda é proporcional à raiz quadrada da tensão e inversamente proporcional ao seu comprimento e à raiz quadrada da sua massa por unidade de comprimento. Isso explica muitos aspectos característicos das cordas dos pianos: na parte superior do teclado, as cordas ficam cada vez mais curtas (frequência fundamental mais alta, f_1); se tivermos que afinar uma certa corda um pouco mais alto, devemos aumentar a tensão (subir f_1) e vice-versa; na região grave, para economizar espaço e maximizar a potência de saída, em vez de aumentar o comprimento da corda, aumenta-se a sua massa por unidade de comprimento, d , envolvendo-a com um enrolamento extra de fio metálico (abaixando f_1). No violino, onde temos apenas quatro cordas de comprimentos aproximadamente iguais, cada uma deve ter uma tensão diferente (ou massa), a fim de possuir uma altu-

2. Mais precisamente, os sobretons são os componentes de frequências mais altas de uma vibração complexa, a despeito de suas frequências serem ou não múltiplos inteiros da frequência fundamental.

ra básica diferente. Para variar a frequência fundamental f_1 de uma certa corda, deve-se mudar o seu comprimento vibrante L pressionando-se a corda contra o braço do instrumento, introduzindo assim um nodo no ponto de contato.

O aparecimento, de maneira natural, de frequências discretas relacionadas a uma frequência fundamental fixada pelas condições do sistema, estando todas as outras frequências “proibidas”, recebe o nome de “quantização” e exerce um papel fundamental em todos os campos da física. As diferentes formas discretas de vibração possíveis em um sistema físico são chamadas de modos de vibração. A fundamental, a oitava, a décima segunda etc. são o primeiro, o segundo, o terceiro etc. modos de vibração de uma corda tensa. Todas as frequências dos possíveis modos de vibração de uma corda são dadas pela relação (4.2). Nessa relação aparecem apenas quantidades que dependem da corda; assim, os modos de vibração são uma característica permanente do sistema físico particular. Em qual dos possíveis modos uma corda de fato vibrará? Isso é determinado pela maneira como as vibrações são iniciadas, *i.e.*, pelo mecanismo primário de excitação. Devido à capacidade de superposição linear das ondas, muitos modos diferentes podem coexistir simultaneamente sem que perturbem uns aos outros. Nesta seção, vamos concentrar nossa atenção em como uma corda pode vibrar. Na próxima seção, vamos examinar a questão de como uma corda realmente vibra.

Vamos considerar o caso de uma corda vibrante em que apenas um modo é excitado. Isso pode ser conseguido facilmente no laboratório, fazendo-se com que uma corrente elétrica alternada de frequência definida percorra uma corda metálica tensa, estirada no vão de um ímã permanente forte. As forças magnéticas sobre a corrente na corda vão induzir uma vibração transversal com a frequência da corrente. Sempre que essa frequência estiver próxima de um dos harmônicos da corda, produz-se uma grande onda estacionária; pode-se observar visualmente os nodos e ventres (como se vê na Figura 4.2), e ouvir claramente o som produzido (desde que a corda esteja montada sobre uma caixa ressonadora). O uso de um estroboscópio (uma fonte de luz que pisca a uma frequência definida e controlável) permite “congelar” a forma da corda, ou observá-la em “câmera lenta”.

Uma experiência mais acessível e amplamente conhecida pode ser realizada com um piano. Abaixo lentamente a tecla de uma nota grave, digamos, sol_2 (Figura 4.3), mantendo-a abaixada, mas de modo que não se produza nenhum som (o martelo não percute); o abafador, todavia, permanece afastado da corda. Agora toque de maneira forte e *staccato* a tecla da nota oitava acima (sol_3). Depois que esse som desaparece, você ouve claramente a corda sol_2 vibrando uma oitava acima: ela foi excitada (por ressonância) em seu segundo modo harmônico, sol_3 ! Agora repita o mesmo experimento percutando a décima segunda ($ré_4$), enquanto mantém abaixada a tecla sol_2 : você ouvirá a corda sol_2 vibrando em $ré_4$. Continue com sol_4 , si_4 etc. Para testar, toque $lá_3$ ou $fá_3$ enquanto segura sol_2 — nada acontecerá: a corda sol_2 permanece em repouso. O motivo é que $lá_3$ e $fá_3$ não são harmônicos superiores de sol_2 , e a corda sol_2 simplesmente não consegue sustentar vibrações estáveis nessas frequências.

A relação (4.2), na verdade, é apenas aproximada, em especial para os modos mais superiores. A razão é que a velocidade de uma onda transversal numa corda depende ligeiramente da frequência (ou comprimento de onda) da onda (isso é o que se chama dispersão) e as expressões (3.3) e (3.8) não são totalmente corretas. Na verdade, a velocidade da onda V_T é ligeiramente maior do que o valor dado pela relação (3.3). Esse desvio aumenta à medida que aumenta a distorção da corda, *i.e.*, torna-se mais importante para comprimentos de onda menores e amplitudes maiores. O resultado é que as frequências dos modos de vibração superiores de uma corda de piano são ligeiramente superiores aos valores dados (4.2)³. Em geral, quando as frequências dos modos superiores de vibração de um sistema não são múltiplos inteiros da frequência fundamental, nós chamamos esses modos de não-harmônicos. Outros corpos sólidos vibrantes que não sejam cordas — p. ex., as barras de um xilofone, sinos, campainhas — possuem muitos modos de vibração que não são harmônicos, cujas frequências absolutamente não são múltiplos inteiros da frequência fundamental. Na sequência desta obra, vamos assumir, por simplicidade, que os

3. Isso não afeta muito a altura do som complexo resultante (Apêndice II), mas afeta a afinação do piano nos registros agudo e grave (quando a afinação é feita por oitavas).

sobretons de uma corda vibrante coincidem com os harmônicos superiores e que a relação (4.2) é exata. Assim, usaremos indistintamente os termos “harmônicos superiores”, “modos” ou “sobretons”, embora fisicamente se trate de conceitos diferentes no caso de cordas reais.

4.2 GERAÇÃO DE VIBRAÇÕES ESTACIONÁRIAS COMPLEXAS EM INSTRUMENTOS DE CORDAS

Existem duas maneiras fundamentais de se excitar a vibração de uma corda tensa:

1. Um fornecimento momentâneo de energia, pela ação de percutir ou puxar a corda (piano, no primeiro caso; violão, cravo etc., no segundo caso).
2. Um fornecimento contínuo de energia pela ação do arco (família dos violinos).

Em ambos os casos, o efeito resultante será uma superposição de muitos modos de vibração ativados simultaneamente. Em outras palavras, os sons musicais individuais gerados naturalmente pelas cordas contêm muitas frequências diferentes ao mesmo tempo — aquelas correspondentes aos harmônicos do sistema vibrante. A Figura 4.4 mostra como isso pode acontecer na prática: acrescentando o primeiro harmônico e, digamos, o terceiro, obtemos uma superposição resultante que, num certo instante, pode parecer com o que vemos na Figura 4.4. Cada modo se comporta independentemente, e a forma instantânea da corda é dada pela superposição (soma) dos deslocamentos individuais, devido a cada modo em separado. É possível utilizar o aparato experimental já citado, da corda vibrante percorrida por uma corrente elétrica e colocada num campo magnético, usando dessa vez a saída combinada de dois geradores de voltagem senoidais, cujas frequências serão iguais a dois dos harmônicos da corda, respectivamente. Com o uso de um estroboscópio, é possível visualizar cla-

ramente a forma instantânea da corda quando ela está vibrando em dois modos ao mesmo tempo. A proporção com que cada sobreton intervém na vibração resultante determina, em grande parte (Seções 1.2, 4.8), o caráter, a qualidade ou o timbre particular do som gerado. A altura do som complexo de uma corda é determinada pela frequência fundamental (4.3), conforme já foi antecipado na Seção 2.7

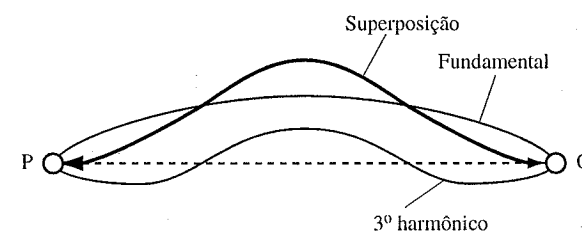


Figura 4.4

Um experimento muito simples com um piano pode mostrar de maneira convincente que uma corda pode realmente vibrar em mais de um modo ao mesmo tempo. Pressione e mantenha abaixada uma certa tecla (digamos sol_2) (Figura 4.3), de modo a retirar o abafador da corda. A seguir toque, forte e *staccato*, simultaneamente ré_4 , sol_4 e si_4 . Depois que o seu som desaparece, é possível ouvir claramente a corda sol_2 vibrar em todos os três modos simultaneamente — temos apenas uma corda sozinha soando como uma tríade completa de sol maior! O que acontece é que os três modos (terceiro, quarto e quinto harmônicos) foram excitados em amplitudes razoavelmente semelhantes (por ressonância). Um experimento mais drástico é o seguinte: mantenha a tecla sol_2 abaixada — e toque com o seu antebraço direito todas as teclas brancas e pretas de duas ou mais oitavas de sol_3 — depois que a explosão inicial de ruído decair, a corda sol_2 vibra maravilhosamente no acorde de dominante com sétima sol_3 , ré_4 , sol_4 , si_4 , ré_5 , fá_5 , sol_5 , ... (Figura 4.3). Qualquer uma das outras notas tocadas não puderam excitar ou manter uma vibração estável na corda sol_2 .

Enquanto o experimento acima mostra que uma certa corda de piano pode vibrar simultaneamente em modos diferentes, o experimento seguinte mostra que uma corda de piano, tocada normalmente, vibra realmente em diversos modos harmônicos. Escolha novamente uma nota grave, digamos sol_2 . Mas, dessa vez, mantenha abaixada a tecla sol_3 , sem fazê-la soar. Em seguida toque, forte e *staccato*, sol_2 . A corda sol_3 começa a vibrar em seu próprio modo fundamental (*i.e.*, sol_3). O motivo é que esse modo foi excitado (por ressonância) pelo segundo harmônico da corda sol_2 . Se, em vez de sol_2 , você tivesse tocado $lá_2$, a corda sol_3 teria permanecido em silêncio. Repita agora o mesmo experimento várias vezes, mantendo abaixadas sucessivamente as teclas $ré_4$, sol_4 , si_4 , $ré_5$, ... etc. Cada uma delas será excitada pelo modo harmônico correspondente na corda sol_2 ⁴.

Muitos modos de vibração aparecem juntos quando se faz vibrar uma corda. O que determinará quais deles e quantos deles? Inicialmente, isso é controlado pela maneira particular com que a corda é posta em vibração, *i.e.*, pelo mecanismo primário de excitação. Dependendo de como e de onde puxamos, percutimos ou friccionamos a corda, obteremos misturas diferentes de sobretons e, portanto, qualidades diferentes do som resultante. Podemos explicar isso com base nos exemplos seguintes. Vamos atribuir a uma corda a forma inicial vista na Figura 4.5a (embora isso seja um tanto difícil de se conseguir na prática). Como a forma está mais ou menos de acordo com o modo fundamental (Figura 4.2), a corda naturalmente começará a vibrar nesse modo quando for solta. Agora, se a forma inicial for a da Figura 4.5b, a corda vibrará no terceiro modo quando for solta (Figura 4.2). Mas o que acontecerá se a forma inicial tiver o aspecto mais realístico da Figura 4.6, que é obtida quando puxamos a corda no ponto central *A* entre *P* e *Q*? Para descobriremos, vamos superpor, *i.e.*, somar linearmente, os casos das Figuras 4.5a e b. Obteremos a forma vista na Figura 4.7a, que se assemelha bem à configuração inicial da corda puxa-

4. O objetivo principal do mecanismo de pedal do piano está baseado nesse fenômeno: o ato de abaixar o pedal libera todos os abafadores, e as cordas podem vibrar livremente por ressonância. Se uma nota é tocada, todas as outras notas (cordas) que pertencem à série de harmônicos da primeira serão induzidas a vibrar.

da (Figura 4.6). Assim, antecipamos que o modo fundamental e pelo menos o terceiro harmônico devem estar simultaneamente presentes na vibração de uma corda dedilhada no ponto central. Podemos melhorar bastante a aproximação à forma da Figura 4.6, acrescentando mais harmônicos superiores em proporções apropriadas (Figura 4.7b). Podemos eliminar as pontas restantes nesta figura acrescentando mais e mais harmônicos superiores em proporções adequadas até que consigamos reproduzir quase com exatidão a forma desejada.

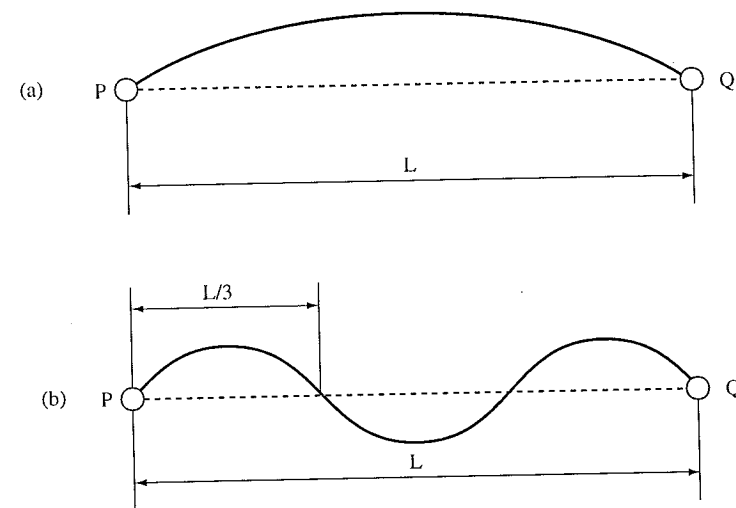


Figura 4.5

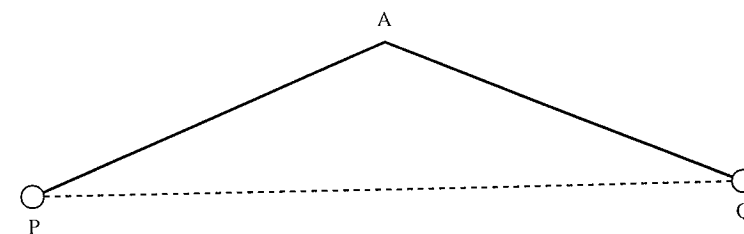


Figura 4.6

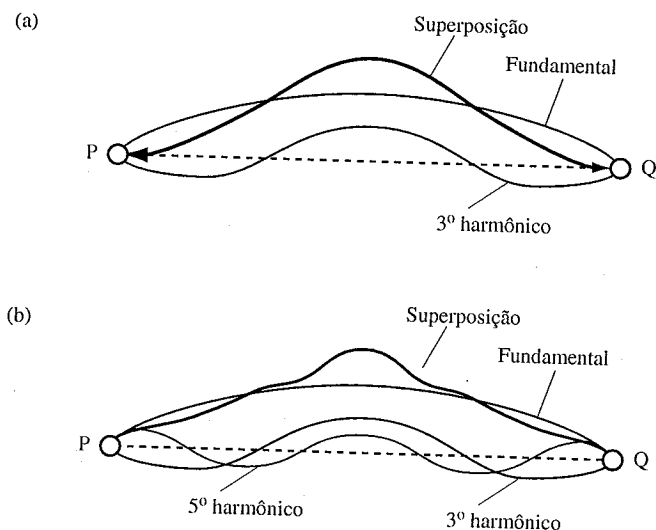


Figura 4.7

Um fato a se destacar é que não existe nenhuma suposição envolvida nisso: tudo pode ser conseguido de uma forma rigorosamente matemática! De fato, pode-se mostrar que qualquer forma inicial arbitrária de uma corda pode ser reproduzida a um grau arbitrário de precisão por uma certa superposição de formas geométricas correspondentes aos modos de vibração harmônicos da corda (ondas estacionárias). É essa superposição “matemática” de formas, e em particular a proporção entre as suas amplitudes e fases, que define a superposição física de harmônicos com que a corda irá realmente vibrar quando for liberada de sua configuração inicial. Em outras palavras, cada uma das ondas estacionárias componentes que, quando somadas, constituem-se na forma inicial da corda (p. ex., Figura 4.7b), apresenta o seu próprio modo de vibração, com sua frequência e amplitude características, quando a corda é liberada. Com o passar do tempo, a forma instantânea da corda muda periodicamente de um jeito muito complicado; mas, a cada vez que um período fundamental $\tau_1 = 1/f_1$ se passa, todos os modos componentes se encontrarão novamente na mesma relação do início, e a corda terá a mesma forma que tinha a princípio. É muito

importante destacar aqui que a configuração inicial da corda determina não só as amplitudes dos modos harmônicos de vibração, mas também as suas fases (defasagens relativas). O ponto onde a corda é puxada determinará a proporção particular de harmônicos superiores, *i.e.*, o timbre inicial do som emitido (Seção 1.2). Se puxarmos no centro, teremos a situação vista na Figura 4.7b, e surgirão apenas harmônicos ímpares. Por outro lado, quanto mais próximo aos pontos fixos puxarmos, mais rica será a proporção de harmônicos superiores. Em geral, todos os harmônicos que têm um nodo no ponto dedilhado serão eliminados (p. ex., todos os harmônicos pares no exemplo da Figura 4.6), enquanto os que têm um ventre nesse ponto serão realçados. Esse efeito é explorado com eficiência pelo harpista para controlar o timbre do seu instrumento.

Numa corda que é dedilhada para ser posta em vibração, temos uma situação em que o mecanismo primário de excitação fornece uma certa energia potencial ao sistema, pela deformação da corda. Após a liberação, essa energia inicial é convertida periodicamente, de forma oscilante, em energia cinética de vibração (Seção 3.1). Por outro lado, quando a corda é vibrada por percussão, uma certa quantidade de energia cinética é fornecida inicialmente ao mecanismo de percussão (p. ex., o martelo do piano), colocando em movimento os pontos da corda inicialmente não-deformada. Essa energia inicial é convertida periodicamente em energia potencial de deformação. Pode-se mostrar matematicamente que, a partir do conhecimento das velocidades iniciais dos pontos da corda percutida, é possível deduzir-se a superposição resultante de harmônicos. Assim, uma corda percutida no ponto central oscilará principalmente com a frequência fundamental, mais uma mistura de intensidades decrescentes de harmônicos ímpares. Quanto mais próximo às extremidades *P* e *Q* a corda é percutida, mais rico em harmônicos superiores será o som. Como acontecia com a corda dedilhada, os harmônicos cujos nodos estão no ponto de percussão ou próximos a ele serão excluídos, e os que tiverem ventres nessas posições serão realçados. Na situação mais realística de um martelo de piano percutindo a corda, a teoria aliada a medições cuidadosas (Hall e Askenfelt, 1988) mostra que a duração do contato martelo-corda influencia a mistura de modos harmônicos superiores de forma significativa: quanto mais longo for o contato, mais

pobre em harmônicos superiores será a vibração da corda (modos com períodos mais curtos que a duração do contato serão excluídos).

Quando fazemos uma corda vibrar puxando-a ou percutindo-a, observamos que a vibração desaparece com certa rapidez. Isso é devido à ação das forças dissipativas: o atrito elástico dentro da corda e, com maior importância, as forças que colocam num pequeno movimento ondulatório o que estiver segurando a corda em seus pontos fixos. Apenas parte dessa perda de energia é realmente convertida em energia ondulatória sonora. Uma corda que vibra livremente, instalada numa estrutura pesada e rígida, produz apenas um som fraco: a maior parte da energia de vibração desaparece na forma de energia de atrito (calor). A conversão em energia ondulatória sonora pode ser melhorada instalando-se a corda numa caixa de propriedades elásticas especiais, chamada ressonador (a caixa de ressonância do piano, ou o corpo do violino). Nesse caso, permite-se que as extremidades da corda vibrem um pouco (tão pouco que, em comparação com o resto das vibrações da corda, essas extremidades ainda funcionam tecnicamente como nodos), e a energia da corda pode ser gradualmente convertida em energia de vibração da caixa. Devido à superfície geralmente muito grande dessa caixa, essa energia será então convertida com maior eficiência em energia ondulatória sonora. O som resultante é muito mais forte do que no caso de uma corda instalada rigidamente — mas ele decai mais depressa, por causa da razão consideravelmente maior em que a quantidade disponível de energia da corda é gasta (potência, Seção 3.1).

Vamos examinar o processo de declínio da vibração com mais detalhes. Simplificando, vamos considerar uma corda que esteja vibrando livremente apenas em seu modo fundamental. Vamos concentrar nossa atenção na amplitude gradualmente decrescente da oscilação da corda num ponto, digamos antinodal (ventre). As medições mostram que, para uma certa corda, as oscilações amortecidas e com amplitude maior decaem mais rapidamente do que as de menor amplitude. O movimento resultante é visto na Figura 4.8. Observe a declividade da curva “envoltória”, que acompanha o decréscimo da amplitude. Isso é chamado de decaimento exponencial da amplitude. O mais importante (e felizmente para a música!) é que a frequência da oscilação amortecida permanece constante.

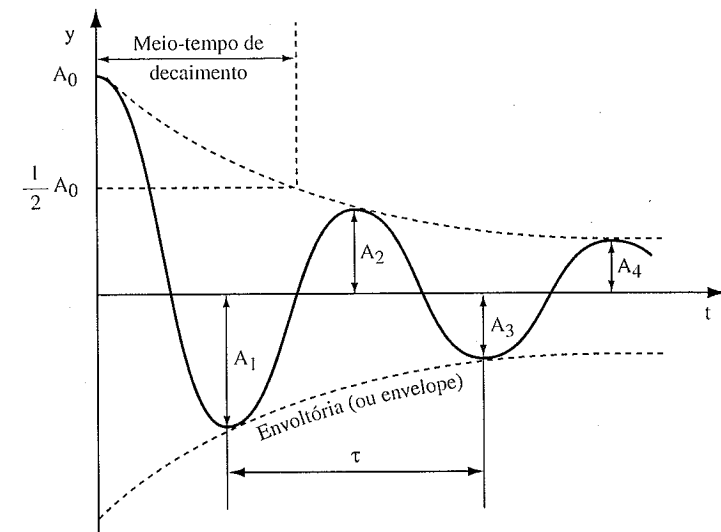


Figura 4.8

Isso é mais ou menos a maneira como uma corda se comporta quando vibra livremente num certo modo, após ter sido puxada ou percutida. Se ela estiver instalada numa base rígida, a perda de energia será relativamente pequena, o mesmo acontecendo com o amortecimento da amplitude (Figura 4.9a). Se, por outro lado, estiver instalada numa caixa de ressonância, ela perderá energia numa razão maior, por colocar a caixa e o ar circunvizinho em oscilação. Assim, as oscilações decairão mais rapidamente (Figura 4.9b).

Uma quantidade característica é o que se chama de meio-tempo de decaimento. Trata-se do intervalo de tempo após o qual a amplitude das oscilações fica reduzida à metade do valor inicial (Figura 4.8). O fato importante de um decaimento exponencial é que o seu meio-tempo é sempre o mesmo durante todo o decaimento: leva sempre o mesmo tempo para a amplitude se reduzir à metade, não importando o seu valor. O meio-tempo de decaimento é, portanto, uma constante característica de uma oscilação amortecida. O meio-tempo típico de decaimento típico de uma corda de piano é de cerca de 0,4 s.

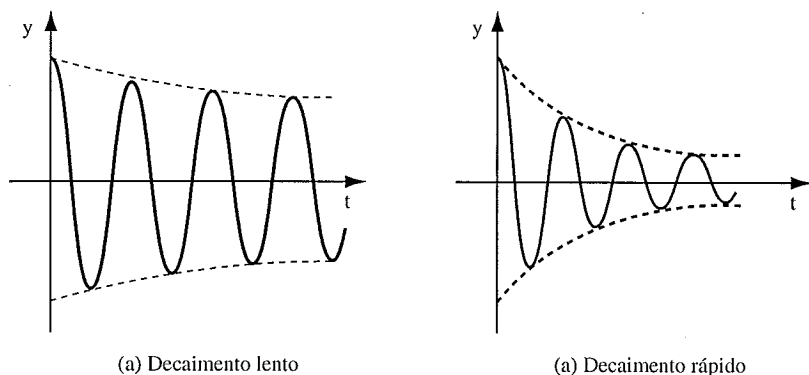


Figura 4.9

Quando uma corda vibra em vários modos ao mesmo tempo, a situação é mais complexa. Todavia, ainda constatamos que cada modo decai exponencialmente, e os meios-tempos de decaimento serão diferentes para modos diferentes. O som complexo resultante, portanto, não apenas diminui em volume, mas também o seu timbre muda gradualmente. Nas cordas do piano, os modos de maior frequência decaem bem mais rápido do que os harmônicos mais baixos; num sino em vibração, os harmônicos mais baixos continuam a soar bem depois que os mais agudos já decaíram. Por outro lado, o comportamento global de uma corda livre vibrante é exclusivamente determinado pela maneira com que a vibração foi inicialmente excitada (corda pinçada ou percutida).

Tem havido uma duradoura disputa entre pianistas e físicos sobre o que se chama de *touché* na execução pianística. Os pianistas prestam muita atenção à maneira como a tecla do piano é pressionada e afirmam que isso influencia o som resultante muito mais do que apenas determinando o seu volume. O físico responde que, uma vez que o martelo está num arremesso livre totalmente desvinculado do instrumentista durante a última porção do seu movimento, o som resultante deve depender de apenas um parâmetro: a velocidade com que o martelo percute a corda. Portanto, no

caso de um único som, o *touché* do piano nada mais é que uma intensidade com um timbre que está irrevogavelmente acoplado a essa intensidade e o decaimento decorrente⁵. A única coisa que o músico pode realmente fazer é controlar a velocidade final do martelo; o timbre não pode ser mudado independentemente do volume e o toque “bonito” ou “feio” não pode existir para sons isolados, diz o físico. O *touché* que sem dúvida existe quando uma peça musical é interpretada está relacionado a outros efeitos psicoacústicos tais como o controle sutil de duração do som, pequenas variações de intensidade de som para som, a melodia que soa acima do acompanhamento, diferenças de intensidade e de tempo entre as notas de um acorde, e mesmo o componente percussivo dado pelo “baque” das teclas quando elas batem na barra de anteparo (Askenfelt e Jansson, 1990). Todavia, há uma esperança para os pianistas que participam dessa “disputa do *touché*”: medições recentes (Askenfelt e Jansson, 1990) revelaram que o movimento detalhado do martelo em arremesso livre como um corpo elástico oscilante e em rotação pode ser ligeiramente diferente para diferentes tipos de toque (mais precisamente, para diferentes acelerações do martelo, controladas pelo instrumentista, antes do seu disparo)! Isso pode levar a um movimento de fricção, relacionado ao toque, realizado contra a corda durante o contato — mas ainda não se demonstrou que esse efeito realmente influencia a excitação da corda de uma forma mensurável.

5. Existe uma razão física um tanto complicada para o acoplamento timbre-volume. Como já foi mencionado na página 172, a duração do contato entre martelo e corda influencia a proporção relativa de modos harmônicos superiores, e, quanto mais longo for o contato, menor o número de modos superiores. A duração do contato, por sua vez, depende da rigidez do feltro sobre a cabeça do martelo: se for mais macio, terá contato mais longo com a corda do que no caso de um feltro mais duro (considerando a mesma intensidade de golpe). Mas existe um fato notável (Hall e Askenfelt, 1988): a rigidez efetiva de um certo martelo depende da velocidade de impacto com que o martelo atinge a corda. A rigidez efetiva será maior para maiores velocidades de impacto e vice-versa (isso também é chamado de comportamento não-linear da rigidez). Em consequência disso tudo, percutir a tecla de um piano com mais força não só aumentará a amplitude de oscilação da corda (maior volume de som), mas também diminuirá o tempo de contato e assim, *automaticamente*, aumentará a proporção de harmônicos superiores (timbre mais brilhante).

O que podemos fazer para evitar o amortecimento das vibrações de uma corda? Obviamente, precisamos compensar a perda de energia de alguma forma fornecendo energia extra ao nosso sistema vibrante numa razão igual à potência dissipada. Se a potência fornecida excede de uma certa quantidade a razão de perda de energia, a amplitude aumentará gradualmente. Mas esse aumento não prosseguirá indefinidamente: enquanto a potência fornecida permanece constante, a dissipação de potência aumentará com a amplitude, e obter-se-á um regime em que a potência dissipada se torna igual à potência fornecida (Figura 4.10). Isso acontece durante a formação do som em qualquer instrumento com capacidade de soar continuamente (a corda friccionada do violino, a flauta, o tubo de órgão etc.). Nesse caso, verifica-se que cada harmônico se forma independentemente, como se houvesse um mecanismo individual de fornecimento de potência para cada modo. Quanto maior for esse fornecimento de potência, maior será o nível de intensidade final.

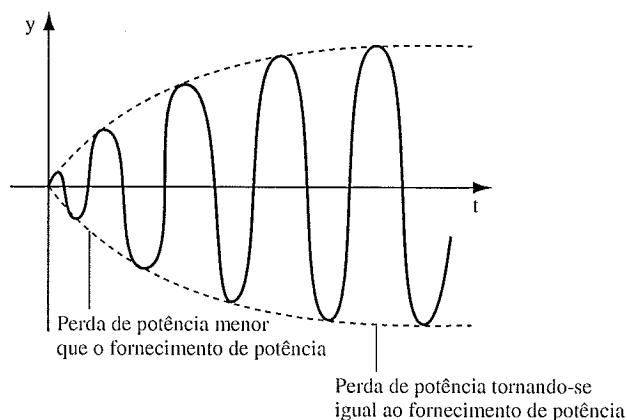


Figura 4.10

A atuação do arco é um bom exemplo de como as oscilações das cordas podem ser mantidas num regime constante. O problema físico é matematicamente complicado e só pode ser tratado após serem feitas várias suposi-

ções simplificadoras (Friedlander, 1953; Keller, 1953). Aqui só podemos apresentar uma descrição qualitativa da teoria. A interação entre o arco e a corda é produzida por forças de atrito. De modo geral, distinguimos dois tipos de interação de atrito. O primeiro é chamado de atrito estático, que surge quando não há deslocamento entre os corpos que interagem. Isso acontece quando a corda “gruda” no arco, movendo-se então com a mesma velocidade dele (ou, em termos mais familiares, quando você tenta empurrar uma mesa pesada e ela “gruda” no chão). O segundo tipo é o atrito dinâmico, que surge quando dois corpos interagentes (as suas superfícies de contato) deslizam um contra o outro. Isso acontece quando a corda “responde bruscamente” e se move na direção oposta à do arco (e acontece quando você continua a empurrar a mesa depois que ela já se moveu). O atrito dinâmico é mais fraco do que o atrito estático; ambos os mecanismos são controlados pela força perpendicular à superfície de contato, que um corpo exerce sobre o outro. No caso de uma corda friccionada, essa força perpendicular é chamada de pressão do arco — um nome terrível para os ouvidos de um físico, já que não se trata de uma pressão, mas de uma força. No Apêndice I discutiremos com mais detalhes uma situação idealizada. As principais conclusões físicas são as seguintes:

1. A amplitude de vibração de uma corda friccionada (volume do som) é controlada unicamente pela velocidade do arco, mas, a fim de manter constante a natureza ou tipo de movimento da corda (timbre do som), devemos manter a pressão do arco proporcional à sua velocidade. Disso sabem muito bem os instrumentistas de cordas, que aumentam a velocidade do arco b e a pressão P simultaneamente para produzir um aumento de volume sem mudança de timbre, ou aumentam b e diminuem P para produzir um aumento de volume com mudança de timbre.
2. Uma corda friccionada sempre tem uma forma instantânea formada por segmentos de linhas retas; há muito tempo esse resultado já foi verificado experimentalmente. Um estudo do balanço de energia no sistema do arco revela que a maior parte da energia fornecida à corda

pelo arco durante as porções “aderentes” do movimento é gasta na forma de calor de atrito (trabalho da força de atrito dinâmico) durante as fases deslizantes. Apenas uma pequena fração é realmente convertida em energia sonora!⁶

Como no caso da corda pinçada ou percutida, a mistura particular de modos harmônicos de vibração dependerá da posição de toque do arco. O arco junto à ponte (*sul ponticello*) realçará os harmônicos superiores e tornará o som mais brilhante; o arco perto da região central das cordas (*sul tasto*) reduzirá a intensidade dos harmônicos superiores de maneira considerável, e o som será “mais macio”.

Na discussão anterior assumimos tacitamente que o arco é deslocado exatamente perpendicular à corda. Se ele tiver uma pequena componente de movimento paralelo, alguns modos longitudinais de vibração poderão ser excitados na corda. A sua frequência é muito maior do que a frequência fundamental dos modos transversais; essas oscilações longitudinais são responsáveis pelos sons rangentes que se ouvem quando um aluno iniciante está tocando.

4.3 ESPECTROS DE VIBRAÇÃO SONORA E RESSONÂNCIA

Quando uma onda vibra em vários modos diferentes ao mesmo tempo, as ondas sonoras geradas também são complexas. Cada componente harmônico da vibração original da corda contribui com a sua própria cota para a onda sonora resultante, com frequência igual à do modo correspondente e com intensidade e fase que estão relacionadas à intensidade e fase desse modo através de processos intervenientes de transformação. O resultado é uma superposição de ondas sonoras combinadas numa única onda complexa com frequência fundamental f_1 (igual à frequência fundamental do elemento vibrante (4.3)), e com uma série de harmônicos supe-

6. Para uma discussão detalhada de cordas friccionadas com arco, veja Schelleng (1973).

riores de frequências $2f_1, 3f_1, 4f_1, \dots$. A vibração resultante é periódica, repetindo-se com um período igual a $\tau_1 = 1/f_1$. Em outras palavras, a frequência fundamental f_1 também representa a frequência de repetição da vibração complexa resultante (Seção 2.7). A forma da curva resultante depende de quais harmônicos estão presentes, da intensidade de cada um (*i.e.*, de suas amplitudes relativas) e de suas defasagens relativas (*i.e.*, de suas fases relativas).

E aqui chegamos a um teorema matemático que teve um impacto absolutamente avassalador sobre praticamente todos os campos da física — em particular na física da música. Resumidamente, o teorema estabelece que qualquer vibração periódica, por mais complicada que seja, pode ser representada como a superposição de vibrações harmônicas puras, cuja frequência fundamental é dada pela frequência de repetição da vibração periódica. Mas isso não é tudo: esse teorema também fornece todas as “receitas” matemáticas para a determinação numérica das amplitudes e fases dos componentes harmônicos superiores! Ele é conhecido como teorema de Fourier, em homenagem a um famoso matemático francês do século XIX. A determinação dos componentes harmônicos de um dado movimento periódico complexo é chamada de análise de Fourier; a determinação do movimento periódico complexo resultante de certa combinação de componentes harmônicos é chamada de síntese de Fourier. Da mesma forma, dado um som complexo, o processo de determinação dos seus componentes harmônicos é chamado de análise de frequência. No caso oposto, dado um grupo de componentes harmônicos, a operação de misturá-los para formar um som complexo é chamada de síntese sonora.

Vamos discutir outro exemplo de análise de Fourier. Obviamente, não podemos apresentar aqui toda a operação matemática necessária para se obter os resultados numéricos; teremos que aceitá-los, concentrando-nos em sua interpretação física. Vamos escolher o movimento periódico visto na Figura 4.11, que corresponde a uma onda “dente-de-serra”. O período é τ e a velocidade de repetição ou frequência fundamental é $f_1 = 1/\tau$. Esse tipo de vibração pode ser gerado eletronicamente. Até certo grau, ele representa, numa forma idealizada, o movimento de uma corda friccionada. A Figura 4.12 mostra como esse movimento periódico pode ser constituído

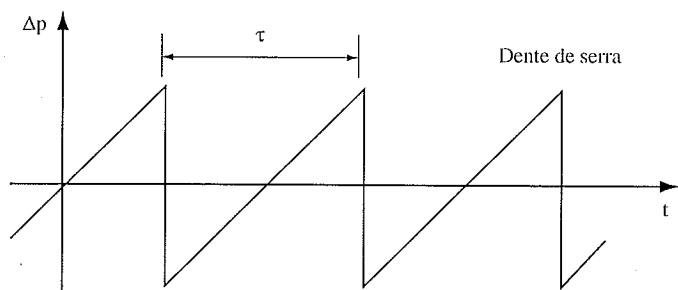


Figura 4.11

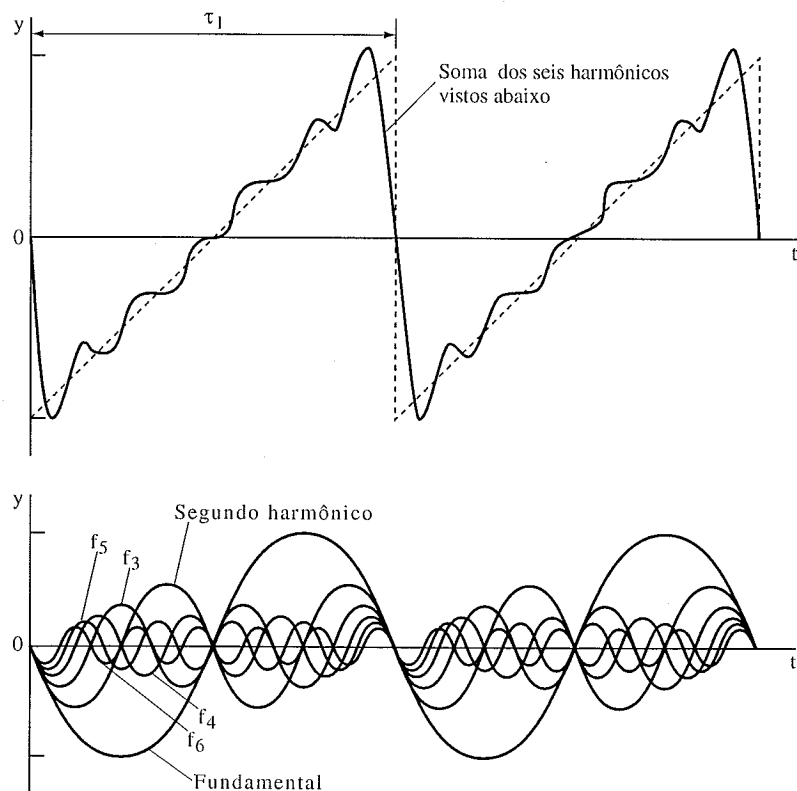


Figura 4.12 Análise de Fourier (até o sexto harmônico) de uma onda dente-de-serra.

de vibrações harmônicas puras. Naturalmente, ainda restam muitas pontas na curva correspondente à soma. Mas isso se deve ao fato de que, por questões de clareza, nós somamos apenas até o sexto harmônico. A soma de mais e mais harmônicos superiores (com amplitude e fase obtidas pelo método de análise de Fourier) eliminará essas pontas, e a forma de dente-de-serra será reproduzida cada vez com mais precisão. Observe atentamente as amplitudes e defasagens relativas dos componentes harmônicos, e como as porções negativas e positivas se somam para formar a curva resultante. Com intuição e um olho treinado é possível prever qualitativamente os principais componentes harmônicos de um movimento periódico de praticamente qualquer forma.

Precisamos agora encontrar um modo de caracterizar fisicamente um som complexo. Em princípio, devemos especificar três séries distintas de quantidades: as frequências dos componentes harmônicos, as amplitudes de variação de pressão ou intensidade dos componentes, e as suas fases ou defasagens relativas (p. ex., Figura 2.5). Na prática, todavia, é de hábito especificar apenas a frequência fundamental f_1 , e as intensidades dos componentes harmônicos porque, em primeiro lugar, compreende-se que todas as frequências superiores são apenas múltiplos inteiros da frequência fundamental, f_1 , e, em segundo lugar, as fases dos componentes exercem apenas um papel secundário na percepção do timbre, em particular para a primeira (e mais importante) meia dúzia de harmônicos (Seção 4.8).

A seqüência de valores de intensidade I_1, I_2, I_3, \dots dos componentes harmônicos de um som complexo representa o que se chama de espectro do som. Dois sons complexos de mesma altura e mesmo volume, mas com espectros diferentes soarão diferentes, *i.e.*, terão timbres diferentes. A diferença no espectro nos dá uma pista importante para distinguir entre sons de instrumentos diferentes — mas outras pistas, em particular o ataque e decaimento do som, também são necessárias para a identificação do instrumento (Seção 4.9). O fato de uma grande quantidade de parâmetros físicos (I_1, I_2, I_3, \dots) estar relacionada ao timbre indica que este último é uma magnitude psicofísica multidimensional.

Os espectros sonoros podem ser representados graficamente registrando-se para cada frequência harmônica (eixo horizontal) a intensida-

de com que esse componente harmônico intervém (eixo vertical) (p. ex., Figura 4.16). Frequentemente, para representar um espectro, são usados valores de NS (3.15) ou NPS (3.16), em vez de intensidades. Além disso, são usados valores de intensidade ou NS relativos à fundamental, ou relativos à intensidade total $I = I_1 + I_2 + I_3 + \dots$. Existem muitos livros onde se encontram reproduzidos espectros sonoros de instrumentos musicais reais (p. ex., Culver, 1956). É necessário, todavia, uma palavra de precaução. De um ponto de vista psicofísico, uma representação harmônica convencional (Fourier) de um espectro sonoro não tem sentido real além do sexto ou sétimo harmônico porque nesse ponto os componentes vizinhos começam a entrar numa banda crítica. Como esta é a unidade de coleta e integração de informação acústica elementar do ouvido (Seção 2.4), o sistema auditivo não conseguiria resolver as intensidades individuais desses harmônicos superiores (caso 1 na página 137). Uma representação psicofisicamente mais significativa dos espectros sonoros é obtida listando-se os valores integrados de intensidade por banda crítica (intervalos de frequência de mais ou menos 1/3 de oitava de extensão).

Apenas os sons constantes podem ser resolvidos numa superposição de harmônicos de frequências discretas que são múltiplos inteiros da fundamental. Quando um padrão de vibração é variável com o tempo, isso não é mais possível. Todavia, pode-se aplicar uma espécie de análise de Fourier numa “versão expandida”. Pode-se mostrar matematicamente que um som dependente do tempo leva a um espectro contínuo no qual todas as frequências estão representadas, com uma certa intensidade para cada intervalo infinitesimal de frequência. Se o som for lentamente dependente do tempo, as frequências discretas (as dos harmônicos) ainda estarão representadas com maior intensidade (“picos” espectrais), mas, se a mudança temporal for apreciável de ciclo para ciclo, o caráter discreto desaparecerá, e o espectro tenderá a se tornar uma curva contínua cobrindo toda a gama de frequências (mesmo que o som fosse originalmente puro). Esse fato leva a outra observação importante no que concerne aos equipamentos de “alta-fidelidade” (veja também página 66). Falamos na Seção 1.2, e voltaremos a isso na Seção 4.8, que os transientes, *i.e.*, as

variações temporais rápidas do padrão de vibração de um som, desempenham um papel determinante na percepção da qualidade ou timbre. Assim, para reproduzir corretamente os transientes de um certo som, os sistemas de gravação e reprodução devem manter o espectro sonoro não-distorcido sobre toda a gama de frequências. Nosso ouvido não precisa dos componentes de frequência situados muito acima dos 5 000 Hz num som constante, mas o sistema de reprodução precisa deles para nos dar uma versão correta das porções rapidamente variáveis do som!

O espectro sonoro de um instrumento de corda não é nada igual ao das vibrações das cordas. A razão se encontra na eficiência, dependente da frequência, do ressonador (caixa de ressonância do piano ou do violino), cuja função principal é extrair energia da corda vibrante numa proporção privilegiada e convertê-la mais eficientemente em potência de onda sonora. Como já foi mencionado antes, as vibrações das cordas são convertidas em vibrações do ressonador num processo em que é permitido às extremidades “fixas” da corda (em particular aquela situada na ponte) que vibrem um pouco. Essa vibração é tão pequena que não invalida o fato de que, do ponto de vista da corda, esses pontos ainda sejam nodos de vibração. Apesar de serem pequenas, essas vibrações envolvem na verdade uma apreciável transferência de energia⁷. A explicação é encontrada na própria definição de trabalho (Seção 3.1): embora o deslocamento das extremidades da corda seja extremamente pequeno, as forças aplicadas a elas são grandes (da ordem da tensão da corda), de modo que o produto força *versus* deslocamento (trabalho) pode ser bastante apreciável. Devido à ampla superfície de um ressonador típico, a conversão da sua energia de vibração em energia de onda sonora é muito eficiente — milhares ou até milhões de vezes mais eficiente do que a conversão direta da energia da corda vibrante em som.

Assim como uma corda, a estrutura elástica complexa da caixa de ressonância de um piano ou violino tem modos favoritos de oscilação.

7. A função da surdina, quando aplicada à ponte de um instrumento de corda, é diminuir essa transferência de energia para os componentes de frequências mais altas, alterando assim o timbre do som resultante.

Nesse caso, entretanto, não há uma relação de múltiplos inteiros simples entre as frequências associadas como na relação (4.2). Além disso, existem tantos modos com frequências quase superpostas que obtemos um todo contínuo de frequências de vibração preferidas, em vez de valores discretos⁸. Vamos discutir brevemente como esses modos de vibração surgem. Para isso, em vez de considerarmos o corpo do violino como um todo, vamos examinar a vibração de apenas uma das placas do corpo. Para encontrar os possíveis modos de vibração é necessário excitar a placa com um oscilador mecânico senoidal de frequência única (p. ex., num ponto com o qual a ponte esteja normalmente em contato). As ondas elásticas afastam-se do ponto de excitação em duas dimensões e são refletidas nas bordas da placa. Os únicos modos de vibração estáveis são ondas estacionárias compatíveis com as condições de contorno particulares da placa. Esse processo é muito difícil de ser tratado matematicamente. No laboratório, todavia, é possível tornar visíveis as vibrações da placa por meio de uma moderna técnica de *laser* chamada holografia (Reinicke e Cremer, 1970). O modo mais simples de oscilação (chamado “modo anel”) é aquele em que a região central da placa se move senoidalmente para cima e para baixo, com o contorno agindo como linha nodal. Os modos de anel das placas do violino determinam a “nota de percussão”, o som que surge quando se percute ligeiramente o corpo do instrumento. As Figuras 4.13 e 4.14 (Jansson e outros, 1970) mostram hologramas de quatro modos de vibração sucessivos no tampo superior (com as aberturas em *f* e a “alma”, mas sem o espelho) e no tampo inferior de um violino, respectivamente. Cada uma das curvas escuras representa um contorno de mesma amplitude de deformação. A diferença de amplitude entre franjas vizinhas é de aproximadamente 2×10^{-5} cm. No instrumento montado, os modos de vibração do tampo superior (Figura 4.13) permanecem quase inalterados, mas surgem novos modos de vibração (na região de baixas frequências).

A resposta de vibração de um ressonador a um certo sinal de amplitude fixa (proveniente ou de um vibrador mecânico ou de uma corda

8. Nem uma corda real, de espessura finita, tem modos de oscilação discretos e “nítidos”.

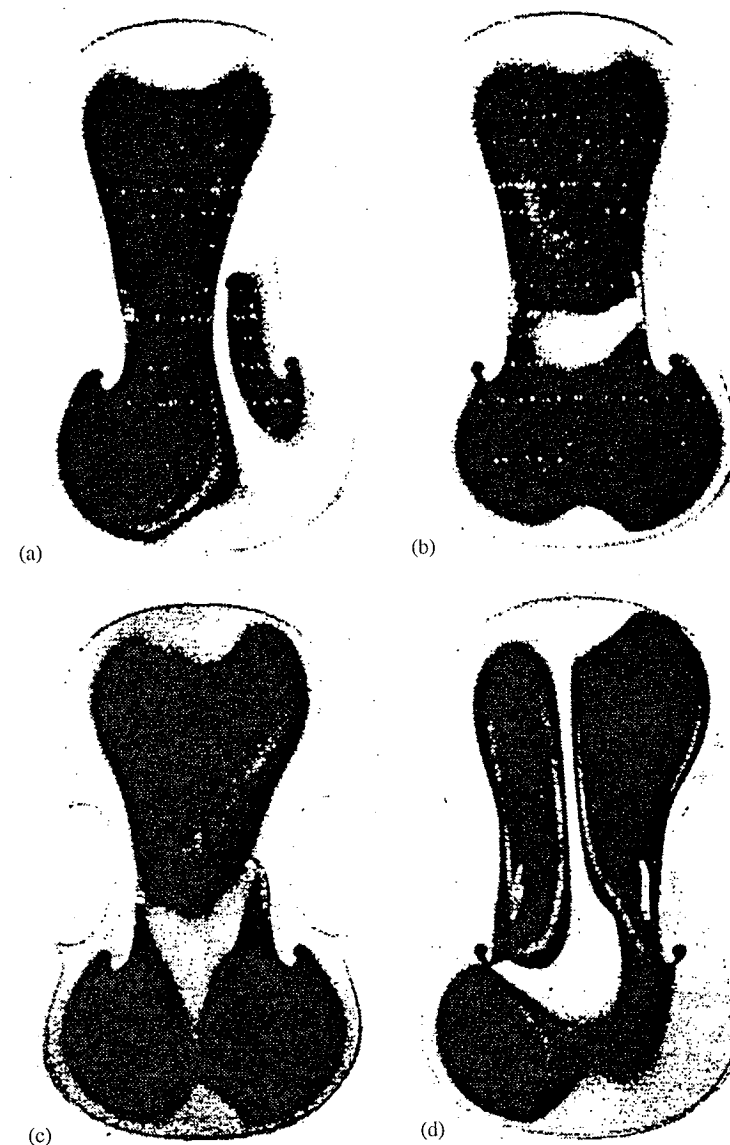


Figura 4.13 Hologramas mostrando os primeiros quatro modos de vibração no tampo superior de um violino (com as aberturas em *f* e a “alma” sem espelho). Cada uma das curvas escuras representa um contorno de mesma amplitude de deformação. (a) 540 Hz, (b) 775 Hz, (c) 800 Hz, (d) 980 Hz.

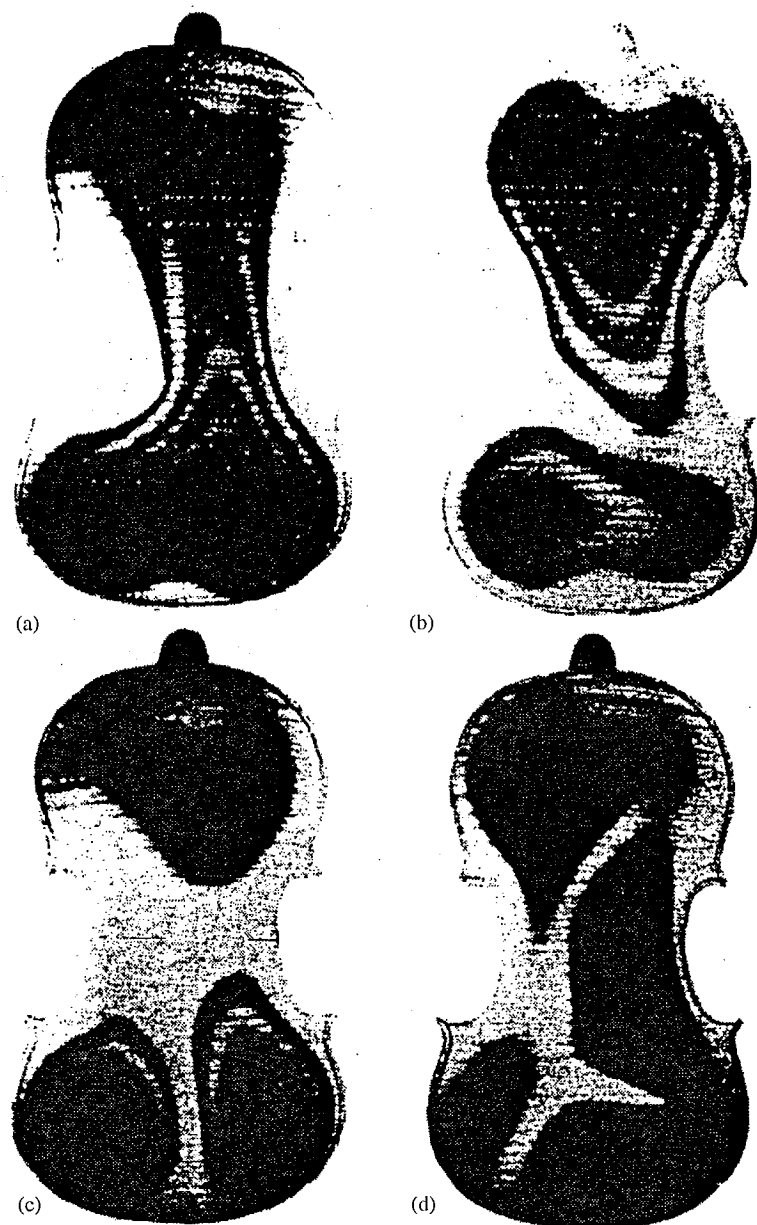


Figura 4.14 O mesmo que a Figura 4.13, para o tampo inferior do violino. (a) 740 Hz, (b) 820 Hz, (c) 960 Hz, (d) 1110 Hz.

vibrante montada nesse ressonador) depende fortemente da frequência das oscilações primárias. Por esse motivo, uma caixa de ressonância reage diferentemente a vibrações de frequências diferentes. Algumas frequências serão preferencialmente realçadas, enquanto outras podem não receber nenhuma amplificação. Uma frequência para a qual a conversão de energia é especialmente eficiente é chamada de frequência de ressonância do ressonador. Um ressonador pode ter muitas frequências de ressonância diferentes; elas podem ser bem definidas (ressonância “nítida”) ou podem estar espalhadas por uma ampla faixa de frequências. O gráfico obtido traçando-se o sinal de saída (por exemplo, aquele medido pela intensidade da onda sonora emergente) como função da frequência de uma vibração senoidal de entrada com amplitude constante é chamado de curva de ressonância ou curva de resposta. Normalmente, a intensidade do sinal de saída I é representada em relação a algum sinal de referência dado I_{ref} e expressa em decibéis (Seção 3.4):

$$R = 10 \log \frac{I}{I_{ref}} \quad (\text{em db}) \quad (4.4)$$

onde R é o valor da função de resposta. A dependência de R com a frequência nos fornece a curva de ressonância mencionada acima. A Figura 4.15 é um exemplo, relativo ao corpo de um violino (Hutchins e Fielding, 1968). Alguns dos picos de baixa frequência correspondem aos modos de vibração do ar dentro do corpo. A primeira subida pronunciada vista na Figura 4.15 corresponde à nota de percussão do corpo⁹. Todas as outras ressonâncias correspondem a modos de vibração mais complicados. A curva de resposta da caixa de ressonância de um piano é ainda mais complicada; essa complicação, entretanto, assegura uma amplificação relativamente uniforme sobre uma grande faixa de frequências.

9. A posição (na frequência) e a forma desse pico de ressonância em particular são da maior importância para o timbre do instrumento de corda (Hutchins e Fielding, 1968). Veja também a página 222.

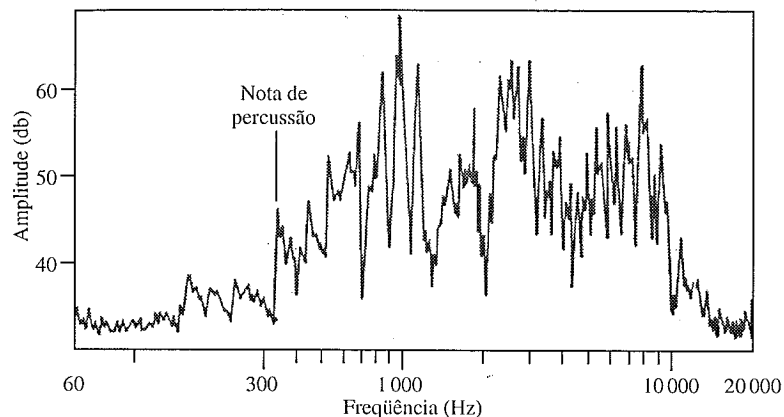


Figura 4.15 Curva de ressonância de um violino (Hutchins e Fielding, 1968).

A Figura 4.15 representa a curva de resposta de um ressonador a uma única vibração harmônica de frequência fixa f . O que acontecerá se ele for excitado por uma corda vibrando com um espectro complexo de harmônicos, de frequências $f_1, f_2 = 2f_1, f_3 = 3f_1, \dots$ e intensidades I_1, I_2, I_3, \dots ? Cada componente harmônico será convertido independentemente, segundo o valor da curva de resposta R correspondente a essa frequência. Assim, a qualidade do som resultante será governada tanto pelo espectro de vibração original da corda quanto pela curva de resposta do ressonador. Fica então bem claro que o ressonador influencia fortemente o timbre de uma corda vibrante.

Como exemplo, considere o espectro hipotético de uma corda vibrante visto na Figura 4.16. Essa corda está montada sobre uma caixa de ressonância hipotética cuja curva de resposta também é apresentada. O espectro sonoro de saída é visto no gráfico da direita (valores relativos). A fundamental está consideravelmente reduzida; por outro lado, o quinto harmônico aparece realçado acima dos outros. Nesse exemplo, de acordo com a Figura 4.16, será obtida maior potência do quinto harmônico do que de qualquer outro. Se a corda tivesse sido inicialmente puxada ou percutida, esse harmônico decairia mais depressa do que os outros porque o seu reservatório de energia se esgotaria mais depressa do que os dos outros. Isso leva a uma variação temporal do espectro ou do timbre, enquanto o

som decai. Se, por outro lado, a corda fosse friccionada, a perda de energia em cada modo seria automaticamente compensada pelo mecanismo do arco, com o timbre resultante permanecendo constante no tempo.

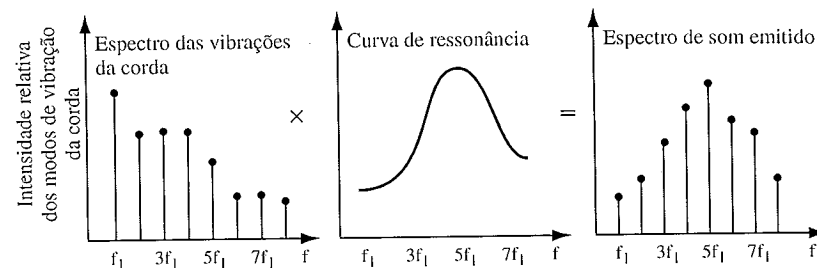


Figura 4.16

Finalmente, chegamos a um ponto muito importante para a música. A curva de resposta de um ressonador é uma característica imutável de um instrumento musical. Se, por exemplo, ele tiver uma região de ressonância em torno de, digamos 1 000 Hz, ele intensificará todos os harmônicos superiores cujas frequências estão perto de 1 000 Hz, não importando qual nota se ouça (desde que, evidentemente a sua frequência fundamental esteja abaixo de 1 000 Hz), nem qual era o espectro de vibração inicial da corda. Uma ampla região de ressonância que realce os harmônicos que se encontrem numa faixa fixa de frequências é chamada de formante. Um instrumento musical (seu ressonador) pode ter muitos formantes. Acredita-se que os formantes, *i.e.*, as intensificações de harmônicos em certos intervalos de frequência fixos e característicos, sejam usados pelo sistema auditivo como uma “marca” muito importante de um som complexo no processo de identificação de um instrumento musical (Seção 4.9). Uma das razões a favor dessa hipótese é o fato de que os formantes são a única característica invariável comum à maioria — senão de todas — das notas de um certo instrumento, enquanto o espectro das notas individuais pode variar consideravelmente de uma nota para outra.

4.4 ONDAS ESTACIONÁRIAS LONGITUDINAIS NUMA COLUNA DE AR IDEALIZADA

Vamos considerar um cilindro muito fino e longo, aberto em ambas as extremidades (Figura 4.17). O ar dentro dele pode ser considerado como um meio elástico unidimensional (Seção 3.2) através do qual as ondas longitudinais podem se propagar. Em qualquer ponto dentro do cilindro é permitido à pressão, momentaneamente, crescer, decrescer ou oscilar consideravelmente em relação à pressão atmosférica normal externa — as paredes rígidas e a inércia da coluna de ar restante mantêm o balanço necessário das forças (3.1), as quais surgem devido à diferença de pressão. Mas, nos pontos extremos, P e Q , não são possíveis grandes variações de pressão, mesmo nos menores intervalos de tempo, porque não há nada que possa balancear as diferenças de pressão que surgem. Assim, esses pontos devem exercer o papel de nodos de pressão, e qualquer onda sonora causada por uma perturbação dentro do tubo e que se propague por ele será refletida em qualquer uma das extremidades abertas. Dessa forma, temos uma situação formalmente análoga à da corda vibrante, discutida na Seção 4.1: as ondas sonoras geradas no tubo permanecem presas lá dentro, e os únicos modos de vibração estáveis possíveis são ondas sonoras longitudinais estacionárias com nodos de pressão nas extremidades abertas P e Q (Figura 4.17). Observe que, em vista da nossa discussão na Seção 3.3, na página 125, as extremidades abertas são ventres de deslocamento, *i.e.*, pontos com máxima amplitude de vibração.

A coluna de ar aberta não precisa necessariamente estar fisicamente definida da maneira vista na Figura 4.17. Por exemplo, existe uma coluna aberta de ar limitada entre os pontos P e Q do tubo, visto na Figura 4.18. Na verdade, como há buracos em P e Q , a pressão do ar nesses pontos deve permanecer constante e igual à pressão externa. Assim, P e Q devem desempenhar o papel de extremidades abertas da coluna de ar. A Figura 4.18 corresponde ao caso de uma flauta idealizada, onde P é a embocadura e Q , o primeiro orifício aberto.

Num tubo aberto real de diâmetro finito, os nodos de pressão não ocorrem exatamente nas extremidades, mas um pouco mais para fora (correção de extremidade, p. 201). As relações dadas abaixo, portanto, são apenas primeiras aproximações.

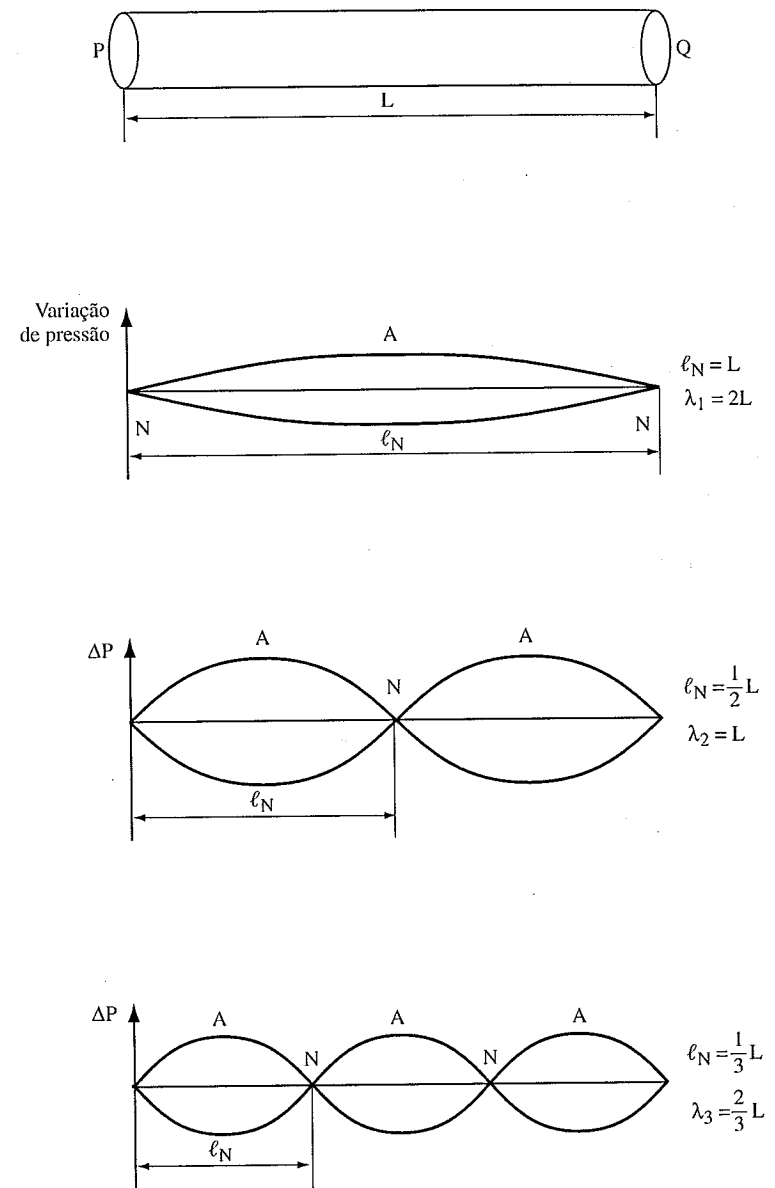


Figura 4.17 Modos de ondas estacionárias (variações de pressão) num tubo cilíndrico idealizado, aberto em ambas as extremidades.

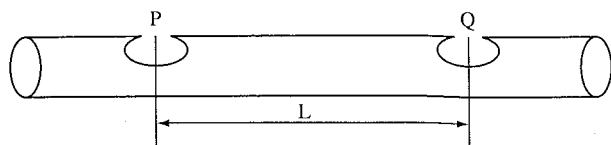


Figura 4.18

Da Figura 4.17 e da relação (3.6), obtemos as frequências dos modos de vibração de um tubo cilíndrico aberto:

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

f_1 é a frequência fundamental:

$$f_1 = \frac{10,5}{L} \sqrt{t_A} \quad (4.6)$$

Lembre-se que t_A é a temperatura absoluta do ar no tubo, dada por (3.5). Em (4.5) e (4.6), L deve ser expresso em metros. Levando-se em conta que o comprimento de onda λ_1 da nota fundamental está relacionado com o comprimento L do tubo por $\lambda_1 = 2L$ (Figura 4.17) e examinando a Figura 3.7, pode-se ter uma idéia dos comprimentos típicos de flautas, tubos de órgão tipo flauta, e flautas doces, em função da frequência. Um aumento na frequência requer uma diminuição no comprimento. A relação (4.6) também mostra o efeito da temperatura do ar sobre a altura fundamental de uma coluna de ar cilíndrica vibrante. Um aumento na temperatura causa um aumento na frequência (nota mais aguda). Assim, as flautas e os tubos de órgão tipo flauta devem ser afinados na temperatura em que se espera que eles sejam tocadas. Felizmente, a frequência fundamental (4.6) é controlada pela temperatura absoluta t_A , que aparece numa raiz quadrada. Ambos os fatos tornam a influência das variações de temperatura sobre

a altura relativamente fraca, mas suficiente para que mereça atenção — como bem sabem os flautistas e organistas.

Vejamos agora o caso de um cilindro estreito fechado em uma das extremidades (Figura 4.19). Supomos que, enquanto a pressão na extremidade aberta P deve permanecer constante e igual à do ar exterior (nodo de pressão), a pressão interna na extremidade fechada Q pode aumentar ou decrescer sem restrições. Na verdade, forma-se um ventre de pressão em A . Isso pode ser compreendido mais facilmente considerando-se o movimento vibratório real dos pontos do meio. Obviamente, deve haver um nodo de vibração para todas as moléculas de ar perto de Q : elas não podem oscilar longitudinalmente para a frente e para trás devido à parede do tubo. De acordo com a discussão na Seção 3.3, esse nodo de vibração corresponde a um ventre de pressão.

A Figura 4.19 mostra como os modos de onda estacionária se adaptam a um tubo fechado, de modo que haja sempre um nodo de pressão na extremidade aberta e um ventre de pressão na extremidade fechada. Para a frequência fundamental, obtemos a relação

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

(L em metros, t_A temperatura absoluta (3.51)). Isso é exatamente metade da frequência fundamental (4.6) de um tubo aberto de mesmo comprimento. Em outras palavras, um tubo cilíndrico fechado idealizado soa uma oitava abaixo do que soa um tubo semelhante aberto nas duas extremidades.

Com relação aos modos superiores de um tubo cilíndrico fechado, a observação da Figura 4.19 (e a conversão do comprimento de onda em frequência) revela que apenas os múltiplos ímpares da frequência fundamental f_1 (4.7) são permitidos:

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

As frequências $2f_1, 4f_1, 6f_1, \dots$ são proibidas — os seus modos não podem ser sustentados de maneira estável num tubo cilíndrico fechado e ideal-

mente fino. Em outras palavras, os sobretons de um tubo fechado são os harmônicos ímpares da sua fundamental.

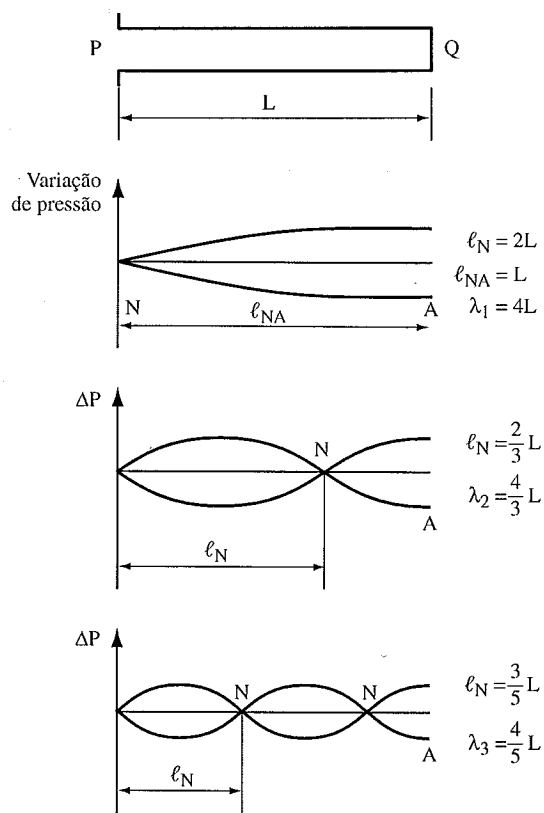


Figura 4.19 Modo de onda estacionária num tubo cilíndrico idealizado e fechado numa extremidade.

A clarineta talvez seja o exemplo mais familiar de um instrumento que se comporta quase como um tubo cilíndrico fechado. A embocadura com a palheta comporta-se como a extremidade fechada, e a campana ou o primeiro orifício aberto definem a extremidade aberta. A altura fundamental de uma nota tocada numa clarineta está realmente uma oitava abaixo da nota correspondente ao mesmo comprimento de coluna de ar tocada numa flauta.

Os órgãos possuem várias filas de tubos fechados. Uma das razões é uma economia de dinheiro e espaço: os tubos graves abertos são muito longos (de acordo com a relação (4.6); um tubo $dó_1$ aberto tem 5.3 m de comprimento). O mesmo tubo, se estiver fechado, terá apenas 2.65 m de comprimento. Naturalmente, há algo mais do que dinheiro em jogo: um tubo fechado tem um timbre diferente de um aberto de mesma frequência fundamental.

Nosso último caso em discussão aqui é o de um tubo cônico (bem estreito), fechado no vértice P (Figura 4.20). A determinação dos modos de vibração exige uma análise matemática muito complexa. Todavia, os resultados podem ser resumidos de uma forma bem simples: um tubo cônico estreito ideal fechado no vértice tem os mesmos modos de vibração que um tubo aberto de mesmo comprimento. Em outras palavras, as relações (4.5, 4.6) podem ser aplicadas. Um cone truncado (estreito) (Figura 4.21), fechado na extremidade P , tem uma série de modos de vibração que não se encaixam na relação de números inteiros: na região de frequência mais baixas (perto da fundamental), eles correspondem aproximadamente aos modos de um tubo aberto com o mesmo comprimento L , mas, nas frequências superiores, eles se aproximam dos modos de um tubo cilíndrico fechado de comprimento L . Em outras palavras, os modos de vibração não são harmônicos.

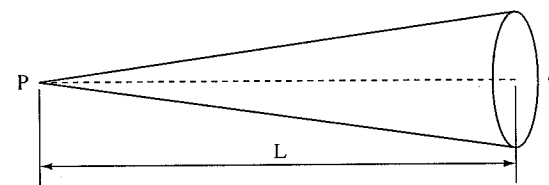


Figura 4.20

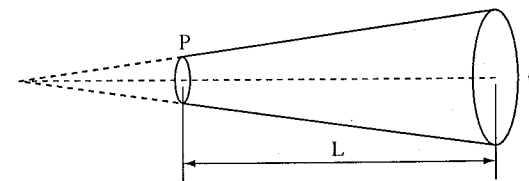


Figura 4.21

4.5 GERAÇÃO DE VIBRAÇÕES ESTACIONÁRIAS COMPLEXAS EM INSTRUMENTOS DE SOPRO

Os resultados da seção anterior são idealizações que só seriam verdade para colunas de ar hipotéticas (cilindros, cones) e diâmetros que são muito pequenos em comparação com o comprimento L . Este, entretanto, não é o caso dos instrumentos musicais e tubos de órgão reais. Além disso, as cavidades de ar nesses instrumentos só são cilíndricas ou cônicas numa certa porção do seu comprimento, apresentando formas mais complexas perto da embocadura e da extremidade aberta (orifícios abertos, campanas etc.).

Para analisarmos o comportamento físico dos instrumentos de sopro reais, precisamos examinar com mais detalhes todos os fenômenos envolvidos. Vamos inicialmente observar o mecanismo de excitação. No caso de uma coluna de ar, não há equivalente ao ato de “puxar” ou “percutir” uma corda. A razão é que as vibrações de uma coluna de ar livremente oscilante decaem quase instantaneamente. Podemos verificar isso facilmente batendo com a mão em uma das extremidades de um tubo aberto (quanto mais longo, melhor), ou percutindo incisivamente a sua parede, enquanto aproximamos o ouvido da outra extremidade. Um pulso sonoro de altura igual à frequência fundamental do tubo será ouvido, mas o decaimento ocorre numa fração de segundo. Assim, é necessário que haja um mecanismo de excitação primária equivalente ao arco numa corda, fornecendo continuamente energia a uma certa razão à coluna de ar vibrante.

Existem dois tipos distintos desses mecanismos. O primeiro consiste em uma corrente de ar de alta velocidade soprada contra uma cunha aguda e rígida E (Figura 4.22), localizada a uma certa distância d exatamente acima da fenda S . Esse sistema é aerodinamicamente instável: a corrente de ar se alterna para a frente e para trás entre ambos os lados da cunha, partindo-se em redemoinhos de ar chamados “vórtices” ou “turbilhões”, que se propagam para cima, por ambos os lados da cunha. À medida que a velocidade da corrente aumenta, aumenta a proporção em que os vórtices são criados. Como eles representam uma perturbação periódica do ar ambiente, ondas sonoras serão geradas quando a proporção de geração de

vórtices se encontrar na região de áudio. O som resultante é chamado de som de borda¹⁰. O mecanismo do som de borda é o processo de excitação primária de todos os instrumentos de sopro da família das flautas e dos tubos de órgão do tipo flauta. As oscilações da corrente de ar em geral são complexas; para intensidades de fluxo muito pequenas, elas se tornam quase senoidais. A frequência fundamental de um som livre de borda depende da velocidade da corrente de ar v e da distância até a cunha d (Figura 4.22). Na região de baixas frequências, ela é proporcional à razão v/d , *i.e.*, ela aumenta com o aumento de v e com a diminuição de d .

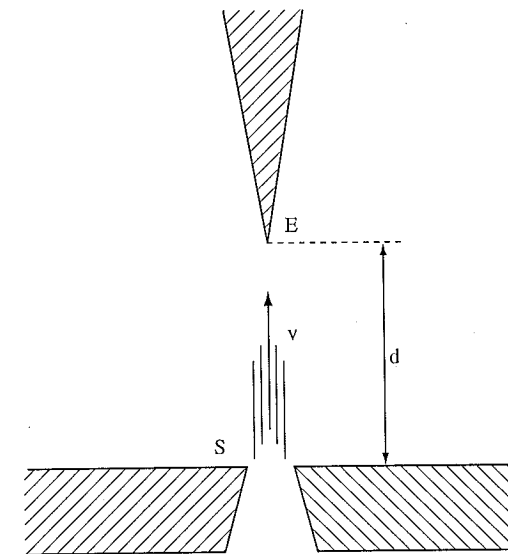


Figura 4.22

O outro mecanismo de excitação que tem importância em música é a palheta, uma placa fina feita de cana, plástico ou metal, colocada em fren-

10. Os vórtices são formados, inclusive, na ausência de bordas ou cunhas, desde que a fenda S seja suficientemente pequena e a velocidade v bastante grande. Isso representa a física básica do assobio humano, em que o tamanho da fenda (abertura dos lábios) e a velocidade da corrente de ar (pressão de sopro) determinam a frequência fundamental.

te a uma fenda que tem quase a mesma forma e com dimensões ligeiramente menores que as da palheta (Figura 4.23). Quando o ar é soprado por baixo na cavidade (*i.e.*, a pressão interna aumenta), o ar em excesso flui pelo pequeno espaço existente entre a palheta e a fenda, para dentro da chalota. Durante esse fluxo, a palheta é puxada para a fenda¹¹. Isso eventualmente interrompe o fluxo; a elasticidade própria da palheta abre a fenda de novo e o jogo começa outra vez. Em outras palavras, a palheta começa a oscilar para a frente e para trás, alternadamente fechando (parcial ou totalmente) e abrindo a fenda. O ar penetra na chalota em golfadas periódicas, dando origem a um som chamado de som de palheta. A frequência fundamental de um som livre de palheta depende tanto das propriedades elásticas da palheta quanto da pressão em excesso na cavidade (pressão de sopro). Em geral, o movimento vibratório de uma palheta livre é complexo, exceto em amplitudes muito pequenas, onde ele é quase senoidal. Alguns instrumentos (oboé, fagote) possuem palhetas duplas, batendo uma contra a outra. Também os lábios de um instrumentista de sopro de embocadura podem ser considerados como um sistema de palhetas duplas (bem maciças).

As notas de borda e de palheta discutidas acima raramente são usadas sozinhas (notas “livres” de borda e de palheta). Nos instrumentos de sopro de madeira, elas servem meramente como mecanismo de excitação primária: como fornecedor de energia para a coluna de ar dentro do tubo. Nesses casos não apenas o espectro, mas também a frequência das vibrações da corrente de ar ou da palheta, é controlado pela coluna de ar via um mecanismo de realimentação (não-linear). Isso é realizado pelas ondas sonoras na coluna de ar: o primeiro pulso ondulatório de compressão produzido propaga-se pelo tubo, é refletido na outra extremidade (aberta ou fechada) e retorna à embocadura (como um pulso de rarefação nos tubos abertos, ou um pulso de compressão nos fechados). Aí ele ocasiona uma variação de pressão que, no caso das madeiras, supera todas as outras forças (aerodinâmicas ou elásticas) e controla, assim, o movimento da corrente de ar

11. Pela diferença de pressão dinâmica (não estática) em ambos os lados da palheta — o mesmo efeito que mantém um avião no ar!

ou da palheta. A altura resultante é bem diferente (em geral, bem menor) que aquela originada quando o mecanismo correspondente de cunha ou de palheta é ativado livremente, na ausência do tubo. Isso é muito diferente do caso de uma corda montada numa caixa de ressonância, cuja altura permanece inalterada pelo ressonador (realimentação zero). No caso dos instrumentos de sopro de metal, a massa dos lábios do instrumentista é tão grande que a realimentação só pode influir na sua vibração, mas nunca superá-la; esta deve ser controlada pelo próprio instrumentista, ajustando a tensão dos seus lábios. Existem alguns instrumentos musicais com palhetas abertas (acordeão, gaita, harmônio).

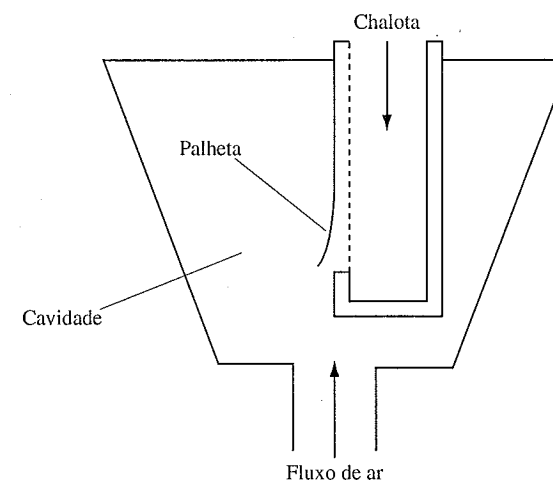


Figura 4.23

O processo de formação do som num instrumento de sopro é muito complicado e ainda não totalmente pesquisado. Mas é de importância fundamental para a música. Em muitos instrumentos, os harmônicos superiores se formam mais depressa do que a fundamental; às vezes, isso pode ser realçado artificialmente, dando uma cor muito característica ao som resultante.

Para compreender a geração de som em estado estacionário nas madeiras, metais e tubos de órgão, é necessário analisar as propriedades de ressonância das suas colunas de ar e o acoplamento destas com o mecanismo de excitação primária (corrente de ar, palheta ou lábios). Para isso, vamos estabelecer os seguintes fatos verificados experimentalmente:

1. O mecanismo de excitação primária sustenta uma oscilação que é complexa, tem uma certa frequência fundamental e uma série de harmônicos de espectro definido.
2. A frequência fundamental e o espectro da oscilação primária são controlados (influenciados, no caso dos metais) pelas propriedades de ressonância da coluna de ar; a amplitude total das oscilações é determinada pelo fornecimento primário de energia (fluxo total da corrente de ar, pressão de sopro).
3. O espectro das oscilações de pressão fora do instrumento (onda sonora gerada) está relacionado ao espectro interno por uma transformação que é governada pela forma e distribuição detalhadas dos orifícios e/ou pela forma da campana (para uma discussão detalhada, veja por exemplo, Benade, 1976).

Para explorar as propriedades de ressonância, *i.e.*, a curva de ressonância, da coluna de ar de um instrumento de sopro, precisamos imaginar um aparato experimental análogo à corda metálica vibrada por corrente alternada (Seção 4.1) ou à placa de violino excitada por um vibrador (Seção 4.3). Isso é conseguido substituindo-se o mecanismo de excitação “natural” por um propulsor mecânico de oscilações (p. ex., uma membrana de alto-falante adequada) e medindo-se, com um pequeno microfone, as amplitudes de oscilação de pressão na peça do bocal (onde, nesse experimento, deve-se formar um ventre de pressão). A curva de ressonância será obtida, então, traçando-se as amplitudes de oscilação de pressão em função da frequência, para uma amplitude de oscilação constante do propulsor. As amplitudes medidas, em geral, são expressas em decibéis (R na expressão 4.4), com referência a um certo nível-padrão. As curvas desse tipo também são chamadas gráficos de impedância de entrada. A Figura

4.24 esboça curvas de ressonância típicas obtidas para colunas de ar do tipo do clarinete e do oboé (Benade, 1971) (sem embocaduras, campanas ou orifícios abertos). É importante observar que os picos de ressonância obtidos dessa maneira correspondem aos modos de vibração de uma coluna de ar fechada na extremidade do excitador primário, *i.e.*, ao caso real em que se usa uma palheta como mecanismo de excitação em uma extremidade da coluna (ventre de pressão e nodo de vibração na posição da palheta). Para encontrar a curva de ressonância da mesma coluna de ar correspondente ao caso em que ela é excitada por uma corrente de ar (flautas transversais ou doces, tubos de órgão tipo flauta), é suficiente traçar o

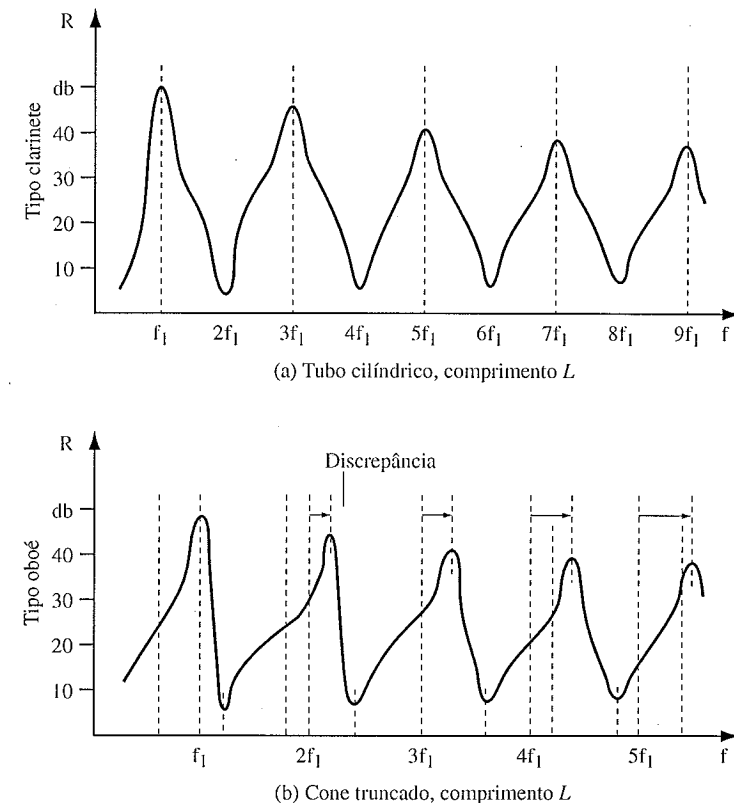


Figura 4.24 Curvas típicas de ressonância (Benade, 1971) para colunas de ar tipo clarinete (cilíndrica) e tipo oboé (cônica) — sem embocaduras, campanas ou orifícios abertos.

gráfico do negativo $-R$ dos valores obtidos na medição anterior¹²: os picos de ressonância tornam-se vales e os vales tornam-se picos (basta virar a Figura 4.24 de cabeça para baixo). A justificativa principal para esse procedimento é que, na embocadura, o ventre de pressão (que aparece quando se usa uma palheta) é substituído por um ventre de vibração, *i.e.*, um nodo de pressão, no caso de uma flauta com embocadura aberta.

Observe, na Figura 4.24, que os picos de ressonância não são muito bem definidos; assim, os possíveis modos de oscilação não correspondem a frequências únicas e discretas, como parecia ser o caso das colunas de ar infinitamente finas (Seção 4.4). Além disso, no caso do cone truncado (b), os picos de ressonância são assimétricos e não-harmônicos (veja a discrepância para os múltiplos harmônicos $2f_1, 3f_1$ etc.). Na região das altas frequências, os picos de ressonância do cone truncado assemelham-se aos do cilindro. Por outro lado, se o cone estivesse completo (até o ápice), os picos de ressonância estariam todos próximos aos vales da curva (a) (harmônicos pares do cilindro) com pouca inarmonicidade¹³.

Vamos discutir qualitativamente como a curva de ressonância controla o mecanismo de excitação primária, por exemplo, de uma palheta. Para intensidades muito baixas (pequenas amplitudes da palheta), o seu movimento é quase senoidal e, em princípio, qualquer frequência de pico de ressonância (Figura 4.24) pode ser estimulada. Na prática, entretanto, verifica-se que apenas a frequência correspondente ao pico de ressonância mais alto é excitada num nível de intensidade muito baixo (pianíssimo). Normalmente, esse é o pico com frequência de ressonância mais baixa; o som obtido pertence ao “registro grave” do instrumento.

À medida que a amplitude da oscilação da palheta aumenta (aumentando-se a pressão de sopro), o caráter não-linear da realimentação da coluna de ar destrói a vibração harmônica e senoidal da palheta, os harmônicos superiores aparecem com intensidade maior (em geral), a intensidade do n -ésimo harmônico cresce proporcionalmente à $2n$ -ésima potência da inten-

12. Apenas se R for expresso em decibéis.

13. Nos instrumentos reais, as curvas de ressonância do tipo clarinete (b) também apresentam uma discrepância em relação à harmonicidade (Backus, 1974).

sidade da fundamental), e o som resultante torna-se “mais brilhante”. Ao mesmo tempo, a frequência fundamental se reajusta se os picos de ressonância superiores forem um pouco inarmônicos. A regra que governa esse reajuste de altura é a seguinte: a frequência fundamental fixa uma posição de tal modo que maximize a altura média ponderada de todos os valores de ressonância R_1, R_2, R_3, \dots ¹⁴, correspondentes aos harmônicos $f_1, 2f_1, 3f_1, \dots$ (Benade, 1971). Se, por exemplo, os picos de ressonância superiores se desviam da harmonicidade, como se vê na Figura 4.24b, o som deve se tornar mais agudo à medida que a sua intensidade aumenta, a fim de acomodar o grupo de harmônicos superiores que vão se tornando importantes e posicionar cada um deles o mais próximo possível a cada pico de ressonância. Devido a esse efeito, um cone truncado não funcionará como um instrumento de sopro de madeira, a menos que algo seja feito para minimizar a inarmonicidade.

Uma situação interessante surge com a curva de ressonância tipo clarinete (Figura 4.24a). Nela, os picos de ressonância estão situados apenas nos múltiplos inteiros ímpares da fundamental (veja também a Seção 4.4). Assim, todos os componentes harmônicos pares da oscilação da palheta serão fortemente atenuados. Partindo de um piano-pianíssimo no registro grave (excitação de frequência única no pico fundamental) e aumentando gradualmente, a pressão de sopro irá, a princípio, tentar estimular o segundo harmônico. A sua energia, entretanto, será inibida eficientemente por causa do vale situado na frequência correspondente (Figura 4.24a). Assim, o aumento resultante no volume (e no “brilho”) será consideravelmente menor para um dado aumento da pressão de sopro do que numa coluna de ar do tipo oboé (curva b), onde o segundo harmônico pode se formar livremente. Eis por que as transições de *ppp* a *pp* são executadas mais facilmente num clarinete do que num oboé ou num saxofone.

Finalmente, outro fato digno de nota na Figura 4.24 é a localização quase idêntica dos vales nas duas curvas. Quando esses vales são convertidos em picos, o que se obtém traçando-se o gráfico de $-R$ em vez de R para se ter a curva de ressonância dessas colunas de ar quando elas são

14. Ponderados com os valores correspondentes de intensidade espectral I_1, I_2, I_3, \dots

excitadas por uma corrente de ar (embocadura aberta), encontramos uma série harmônica praticamente idêntica em ambos os exemplos. Conseqüentemente, o cilindro e o cone truncado podem ser usados quase indistintamente para fazer instrumentos do tipo flauta.

Até agora consideramos ondas estacionárias nas quais a frequência fundamental é determinada pelo primeiro pico da curva de ressonância (frequência mais baixa), fornecendo as notas do registro grave dos instrumentos de madeira. No “registro médio”, a frequência fundamental encontra-se próxima ao segundo pico de ressonância. Nos instrumentos de palheta, isso é conseguido diminuindo-se o tamanho do primeiro pico de ressonância, de modo que fique menor que o segundo, e afastando a sua posição da série harmônica. O orifício de registro cumpre essa função. Na flauta, essa transição, ou “sobre-som”, é realizada por meio de uma variação (aumento) na velocidade do ar soprado na fenda. Repare que, no primeiro sobre-som de uma palheta do tipo clarinete, a altura salta para o terceiro harmônico ou décima segunda (segundo pico, Figura 4.24a), enquanto as palhetas de tubo cilíndrico (e todas as flautas) têm o seu primeiro sobre-som na oitava (segundo harmônico, Figura 4.24b).

Mais um sobre-som leva ao registro superior das madeiras, com as frequências fundamentais baseadas no terceiro e/ou quarto picos de ressonância. Para se conseguir isso num instrumento de palheta, os dois primeiros picos de ressonância devem ser diminuídos e ter a frequência mudada para destruir a relação harmônica.

Os tubos de órgãos operam essencialmente sob os mesmos fundamentos que uma flauta (tubos tipo flauta, abertos e fechados) ou um instrumento de palheta (tubos de palheta). A diferença principal é que, como existe um tubo para cada nota de um dado registro, orifícios de notas e sobre-som são desnecessários. Os tubos de órgão sempre são operados no registro grave (com algumas poucas exceções nos órgãos “românticos”). As curvas de ressonância dos tubos de órgãos do tipo flauta abertos têm picos localizados próximo aos múltiplos inteiros da frequência fundamental, com uma ligeira inarmonicidade que depende da razão $r = \text{diâmetro/comprimento}$. As curvas de ressonância dos tubos fechados assemelham-se às do gráfico superior da Figura 4.24, com máximos nos múlti-

plos ímpares da fundamental. Quanto maior o valor de r , maior será a inarmonicidade das ressonâncias superiores. Como resultado, haverá um desvio da frequência fundamental do som resultante, além de uma crescente atenuação dos harmônicos superiores — que estarão cada vez mais afastados dos picos de ressonância (inarmônica). Conseqüentemente, o som dos grandes tubos de órgão é menos rico em harmônicos superiores (som “como flauta”). Os tubos estreitos (r pequeno) têm picos de ressonância que se encontram próximos de múltiplos inteiros da frequência fundamental, e haverá uma excitação mais forte dos harmônicos superiores (o som é mais “brilhante” ou “como cordas”). A frequência fundamental está ligeiramente deslocada com relação ao valor dado pela relação 4.6 (tubo aberto) ou 4.7 (tubo fechado). Todavia, essas relações ainda podem ser usadas se se acrescentar ao comprimento L uma correção de $0,3 \times \text{diâmetro}$ para cada extremidade fechada (“correção de extremidade”). Os tubos do órgão de palheta variam desde o tipo em que a vibração da palheta é fortemente controlada pela realimentação da coluna de ar (p. ex., os registros da família dos trompetes) até o tipo em que as vibrações da palheta são praticamente autônomas (registros da família regal).

4.6 ESPECTROS DE SONS DE INSTRUMENTOS DE SOPRO

A ressonância característica da coluna de ar e o mecanismo de excitação colaboram para determinar o espectro e a intensidade da onda estacionária no tubo do instrumento. A Figura 4.24 mostrou duas curvas de ressonância hipotéticas; os instrumentos reais, entretanto, apresentam um comportamento mais complicado devido à forma peculiar da embocadura (e da peça de embocadura), à forma e distribuição dos orifícios abertos, ao efeito da campana e, no caso das flautas, ao efeito da velocidade do ar sobre a largura e posição dos picos de ressonância (Benade, 1971). Aqui, só podemos resumir brevemente os efeitos mais importantes. Os orifícios, além de obviamente determinarem o comprimento efetivo da coluna de ar e, assim, a posição absoluta dos picos de ressonância, são parcialmente

responsáveis pelo corte dos picos de ressonância que variam entre 1 500 e 2 000 Hz. Esse corte tem um efeito importante sobre o timbre (atenuação de harmônicos superiores) e sobre o controle dinâmico dos sons fortes das madeiras, particularmente nos registros médio e agudo. No oboé, a cavidade da palheta e o estreitamento na base contribuem para diminuir a inarmonicidade das ressonâncias do cone truncado.

Os metais merecem uma atenção especial nesta seção. Como já foi dito, o mecanismo de realimentação é menos eficiente na determinação da frequência fundamental, e o instrumentista deve fazer com que a frequência de vibração dos seus lábios se aproxime da frequência desejada para conseguir a altura certa. Num instrumento de metal, os harmônicos superiores são criados pelas propriedades de ressonância oscilante da embocadura, causada pelo fechamento e abertura dos lábios (Backus e Hundley, 1971), em vez de um movimento não-senoidal de realimentação controlada dos lábios. A embocadura, a peça de embocadura cônica, o tubo principal cilíndrico e a campana combinam-se de modo a gerar uma curva de ressonância característica diferente daquela de um instrumento de madeira. A Figura 4.25 mostra um exemplo (Benade, 1971). Observe a frequência de corte assinalada (determinada principalmente pela campana) e o grande aglomerado de picos e vales na região de frequências médias (governado principalmente pela forma da embocadura). Esse aglomerado tem um papel crucial na moldagem do timbre dos instrumentos de metal. Finalmente, o primeiro pico de ressonância encontra-se abaixo da frequência fundamental (marcada pela seta) correspondente ao restante dos picos. Observe também a assimetria característica dos picos na região de baixas frequências (semelhante ao caso do cone truncado, Figura 4.24b), em comparação com a região de altas frequências (onde a forma deles é invertida).

Os metais não têm orifícios de dedilhado para alterar o comprimento efetivo das suas colunas de ar — as variações de altura são efetuadas basicamente por sobre-som, *i.e.*, fazendo-se a frequência fundamental saltar de um pico de ressonância para outro. Isso é conseguido ajustando-se adequadamente a tensão dos lábios. Com um trompete pode-se atingir até o 8º harmônico, e até o 16º harmônico com a trompa. Para obter notas entre os picos de ressonância, existe um sistema de válvulas que oferece uma esco-

lha limitada de comprimentos de tubo ligeiramente diferentes. No trombone, consegue-se, com a vara, uma variação contínua do comprimento do tubo (e, portanto, da altura). Como o pico de ressonância mais baixo está desafinado em relação ao resto de quase toda a série harmônica de picos, ele não pode ser usado. A frequência fundamental das vibrações dos lábios é ajustada para o valor da fundamental ausente (seta na Figura 4.25), que corresponde ao segundo, terceiro etc. picos. Isso leva à chamada “nota pedal” dos instrumentos de metal (usada principalmente no trombone). Ela só pode ser executada em níveis de volume consideráveis.

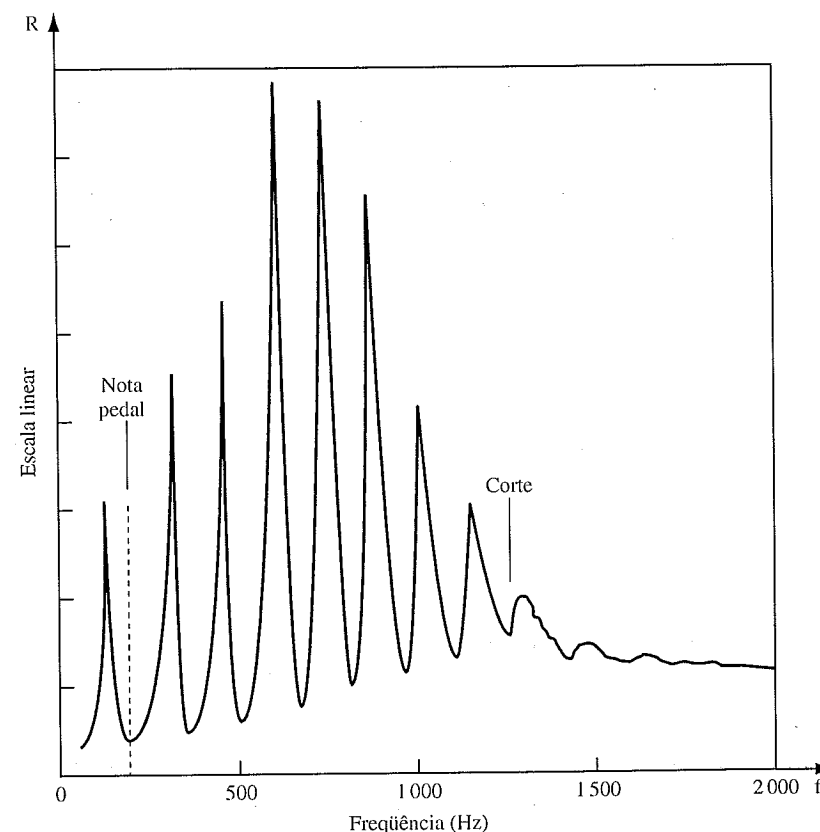


Figura 4.25 Curva de ressonância de um trompete (Benade, 1971) (dada em escala linear).

A composição espectral das ondas sonoras emitidas por um instrumento de sopro é diferente daquela das vibrações estacionárias sustentadas em sua coluna de ar. A campana e/ou os orifícios são os principais responsáveis por essa transformação espectral. Essa transformação deixa a composição espectral relativamente inalterada acima da frequência de corte, tendendo a atenuar os harmônicos mais inferiores. Em outras palavras, os espectros sonoros das madeiras e dos metais são mais ricos em harmônicos superiores do que as vibrações que são realmente produzidas dentro do instrumento (Benade, 1973, 1976).

Alguns espectros sonoros de instrumentos de sopro possuem formantes, *i.e.*, características que são independentes da frequência fundamental do som (Seção 4.3). O fagote e o corne inglês são exemplos, com uma acentuação espectral (não muito bem definida) por volta de 450 a 1 100 Hz, respectivamente. Esses formantes, porém, são causados pelas características do espectro de excitação das *palhetas duplas*; eles não são determinados pelas propriedades de ressonância do tubo do instrumento. Embora não sendo um tópico explícito deste livro, devemos mencionar que a voz humana é provavelmente o exemplo mais notável de um “instrumento de sopro” onde os formantes desempenham um papel crucial: eles são a característica determinante de todos os sons vogais. Os formantes na voz humana são determinados principalmente pelas propriedades de ressonância da cavidade nasofaríngea (Flanagan, 1972). A forma dessa cavidade determina uma entre duas faixas principais de frequência das vibrações das cordas vocais que será realçada. Essas frequências, por sua vez, determinarão se o som resultante será “a”, “i”, “o” etc.

4.7 CONFINAMENTO E ABSORÇÃO DE ONDAS SONORAS NUM AMBIENTE FECHADO

Em geral, os instrumentos musicais são tocados em salas, salões, auditórios e igrejas. O som que um ouvinte percebe nessas condições é bem diferente do som produzido pelo instrumento. Por essa razão, o local onde o instrumento é tocado pode ser considerado uma extensão natural dele, com a

diferença que as propriedades acústicas do ambiente variam muito de caso para caso e de lugar para lugar, enquanto um instrumento musical tem certas propriedades acústicas que são imutáveis. O tema da acústica de auditórios é tão importante para a música como a física dos instrumentos musicais.

Para analisar o efeito de um ambiente sobre uma fonte sonora musical colocada nele, vamos considerar um instrumento musical na posição S e um ouvinte na posição L (que poderá coincidir com S se o ouvinte for o instrumentista) numa sala com paredes perfeitamente refletoras (Figura 4.26). O instrumento começa a tocar uma certa nota no instante $t = 0$, mantendo constante a sua intensidade. Assumimos que o som é emitido por igual em todas as direções. À medida que as ondas sonoras se afastam de S , o ouvinte receberá um primeiro sinal depois do curto intervalo de tempo SL/V , que é o tempo necessário para o som direto ir de S a L [p. ex., se $SL = 10$ m, $V = 334$ m/s (relação 3.6) — o tempo de chegada direta é 0,03 s; para o instrumentista, é praticamente zero]. Conforme veremos na Seção 5.1, o som direto desempenha um papel crucial no processo de percepção (efeito de precedência). Imediatamente depois, as ondas refletidas (trajetórias 2, 3, 4, 5, ... etc.) passarão pelo ponto L em rápida seqüência (na figura, as reflexões no chão e no teto foram ignoradas). As primeiras reflexões, se forem muito pronunciadas e estiverem bem separadas uma da outra, são chamadas ecos. Esse jogo continua com as reflexões secundárias, terciárias e múltiplas (não mostradas na figura). À medida que o tempo passa e o instrumento continua soando, a energia acústica que passa pelo ponto L continua se somando. Se não houvesse nenhuma absorção, as ondas sonoras “encheriam” a sala, percorrendo-a em todas as direções, e a energia acústica emitida pelo instrumento se acumularia, permanecendo confinada na sala; assim, o volume aumentaria gradualmente em todos os pontos internos¹⁵. No caso real, naturalmente, existe absorção sempre que uma onda sonora é refletida. Logo, a intensidade da onda sonora não irá crescer indefinidamente, mas irá se nivelar

15. Para praticamente qualquer tipo de forma de sala e posição da fonte aí inserida poderá haver regiões quase inacessíveis às ondas sonoras emitidas por S (pontos “cegos”), ou regiões onde as ondas sonoras estejam focalizadas (p. ex., os pontos focais em recintos elípticos).

quando a potência dissipada nos processos de absorção se tornar igual à razão em que a energia é fornecida pela fonte (uma situação semelhante à formação das vibrações numa corda friccionada, Figura 4.10). Esse nível de intensidade de “equilíbrio”, I_m , do som difuso é muito maior do que o do som direto (exceto nas vizinhanças da fonte sonora).

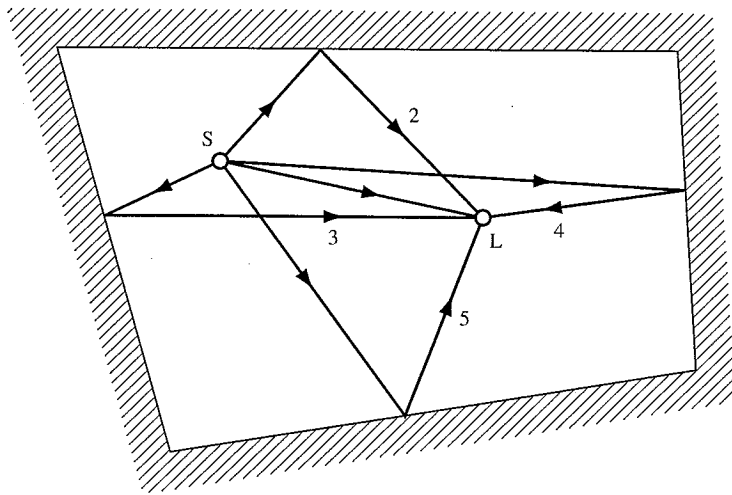


Figura 4.26

Quando a fonte sonora é desligada, desenvolve-se um processo inverso: inicialmente, o som direto desaparece, e depois a primeira, segunda etc. reflexões. A Figura 4.27 mostra esquematicamente o comportamento da intensidade sonora num certo ponto de um ambiente fechado. O decaimento sonoro, depois que a fonte foi desligada, é chamado de reverberação e representa um efeito da maior importância em acústica arquitetônica. Esse decaimento é quase exponencial (p. ex., veja Figura 4.8); de modo arbitrário, podemos definir tempo de reverberação como o intervalo que o nível sonoro leva para decair 60 db. De acordo com a Tabela 3.2, isso representa um decréscimo de intensidade da ordem de 1 milhão. Os tempos de reverberação desejáveis em boas salas de concerto de tamanho

médio são da ordem de 1,5 a 2 segundos. Tempos maiores “embolariam” bastante as seqüências sonoras; tempos menores tornariam a música “seca” e “opaca” (veja Seção 4.8).

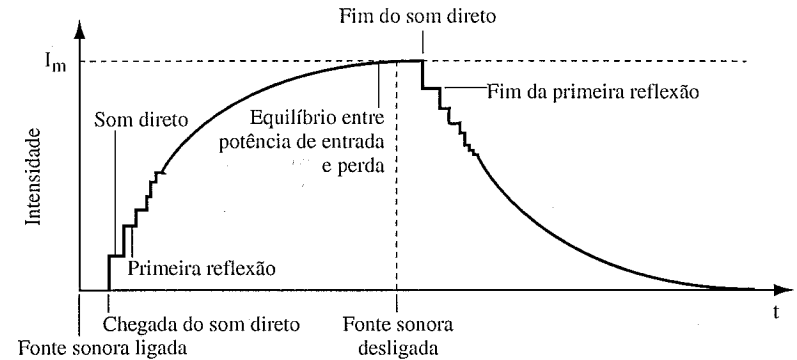


Figura 4.27 Crescimento e decaimento típicos de intensidade sonora numa sala (escala linear).

Podemos discutir algumas relações matemáticas simples que aparecem em acústica arquitetônica. Vamos imaginar uma sala com paredes perfeitamente refletoras, sem nenhuma absorção, mas com um orifício de área A . Sempre que a intensidade máxima I_m é atingida (Figura 4.27), a energia acústica escapará pelo orifício numa razão dada pelo produto $I_m A^{16}$. Como isso corresponde ao estado estacionário em que a potência P fornecida pelo instrumento se iguala à razão de perda de energia, podemos estabelecer $P = I_m A$, ou

$$I_m = \frac{P}{A} \tag{4.9}$$

Num caso real, naturalmente não temos paredes perfeitamente refletoras com buracos. Todavia, ainda podemos imaginar uma parede real absor-

16. Assumimos aqui que I_m representa o fluxo de energia sonora difuso e omnidirecional.

vedora como se fosse feita de um material perfeitamente refletor permeado de buracos, estes últimos representando uma fração a da sua superfície total. Essa fração a é chamada de coeficiente de absorção do material de que é feita a parede. Uma superfície de S metros quadrados, com coeficiente de absorção a , tem as mesmas propriedades de absorção de uma parede perfeitamente refletora de mesmo tamanho, mas com um buraco de área $A = Sa$. Os coeficientes de absorção dependem da frequência do som (em geral, aumentam para as frequências mais altas) e têm valores que variam desde 0,01 (mármore; um refletor quase perfeito) até 0,9 (telhas acústicas). Levando isso em conta, podemos reescrever a relação 4.9 em termos das superfícies reais das paredes S_1, S_2, \dots com os correspondentes coeficientes de absorção a_1, a_2, \dots :

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

Essa relação pode ser usada em estimativas de dimensões de auditórios necessárias para se obterem valores desejados de I_m , para uma determinada potência P de instrumento e uma dada distribuição de materiais absorventes nas paredes.

Verifica-se que o tempo de reverberação τ_r é proporcional ao volume V da sala e inversamente proporcional à área absorvedora das paredes $A = S_1 a_1 + S_2 a_2 \dots$. Os experimentos mostram que, aproximadamente,

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

com V em metros cúbicos, S em metros quadrados e τ_r em segundos. Como os coeficientes de absorção em geral aumentam com a frequência do som, τ_r diminuirá com o aumento da altura: notas graves reverberam mais do que notas agudas.

Um dos problemas em acústica arquitetônica é que o público influi muito (aumenta) nas propriedades absorvedoras de uma sala. Isso deve ser levado em conta no projeto de auditórios. Para minimizar os efeitos rela-

cionados ao tamanho imprevisível do público e de sua distribuição espacial, seria necessário construir assentos cujo coeficiente de absorção fique praticamente inalterado, quer ele esteja ocupado, quer não. O efeito absorvedor do público é maximizado em ambientes com grandes tempos de reverberação, como em igrejas e catedrais. Nenhum músico individual está tão exposto a um ambiente acústico variável (e é incomodado por isso) como um organista.

A distribuição sonora, a formação e decaimento e o fato dos coeficientes de absorção dependerem da frequência, tudo isso tem um profundo efeito nas características físicas dos sons musicais emitidos por um instrumento num ambiente real e, portanto, na percepção que um ouvinte tem da música. A dependência temporal dos sons é profundamente afetada: as características transientes são alteradas e, por exemplo, uma nota *staccato* se prolonga mais no tempo dependendo das propriedades de reverberação da sala. O espectro sonoro também é afetado, já que os coeficientes de absorção são todos dependentes da frequência. Finalmente, levando em conta que as fases das ondas que passam por um certo ponto num campo de som reverberante são misturadas aleatoriamente, podemos mostrar que o NPS resultante de cada componente harmônico também flutuará aleatoriamente, introduzindo um limite na habilidade do ouvinte de reconhecer timbres numa sala musical fechada (Plomp e Steeneken, 1973).

Existem outros efeitos de segunda ordem, normalmente desprezados, relacionados a um fenômeno ondulatório chamado difração. Quando uma onda sonora atinge um obstáculo (p. ex., um pilar numa igreja, ou uma pessoa que esteja sentada à frente do ouvinte), três situações poderão surgir:

1. Se o comprimento de onda da onda sonora for muito menor que o tamanho (diâmetro) do obstáculo (p. ex., uma nota bem aguda) (Figura 4.28a), poderá se formar uma “sombra acústica” atrás do obstáculo, com a reflexão “normal” ocorrendo na parte da frente.
2. Se tanto o obstáculo como o comprimento de onda tiverem mais ou menos a mesma magnitude, surge uma situação mais complicada, em que o próprio obstáculo age como um reemissor de som, irradiando em todas as direções (não mostrado na Figura 4.28).

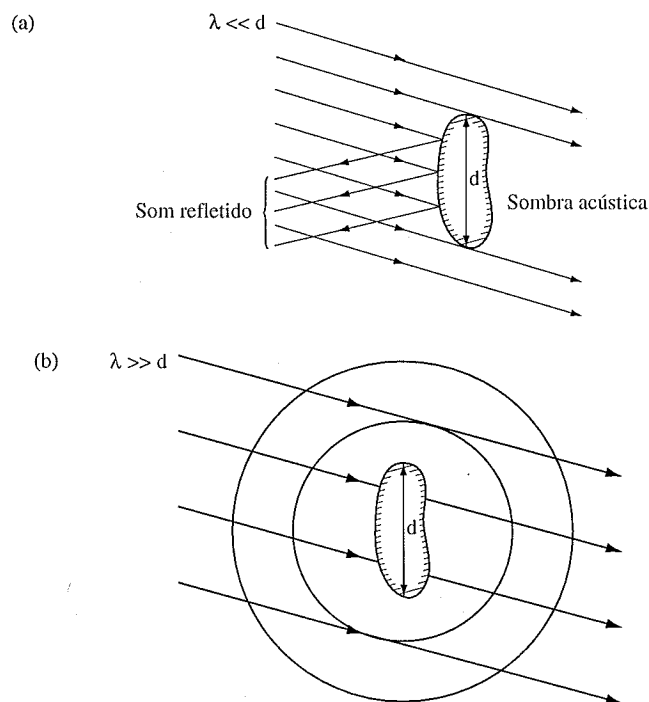


Figura 4.28

3. Se o comprimento de onda for muito maior que o obstáculo (Figura 4.28b) (p. ex. notas graves), este não perturbará a onda sonora, que se propagará praticamente inalterada. Arranjos regulares de obstáculos (como a distribuição de assentos ou de pessoas na platéia) podem levar a padrões de interferência para certos comprimentos de onda e certas direções de propagação. Finalmente, podem-se estabelecer ondas estacionárias para certas configurações do recinto, certas frequências e certas posições da fonte. Isso leva à formação de nodos e ventres perturbadores na sala (Seção 3.3). Para detalhes científicos e técnicos de acústica ambiental, veja Ando (1985).

A difração e o efeito de precedência desempenham um papel importante na reprodução eletroacústica do som. A compensação eletrônica para os

efeitos de difração da onda sonora na cabeça do ouvinte é desejável quando os sinais estereofônicos são recebidos a partir de dois alto-falantes (Damaske, 1971). Uma correta reprodução estéreo do “som direto” (*i.e.*, o efeito de precedência) é necessária para evitar que os fones de ouvido dêem a sensação de uma imagem sonora localizada “dentro da cabeça”.

4.8 A PERCEPÇÃO DE ALTURA E DE TIMBRE DE SONS MUSICAIS

Apesar de haver um número considerável de pesquisas desenvolvidas sobre a percepção da altura e intensidade de sons puros (Seções 2.3, 9; 3.4, 5), muito ainda precisa ser feito no estudo da percepção da qualidade ou timbre dos sons complexos (p. ex., veja Yost e Watson, 1987). Há muitos séculos já se sabe que o timbre de um som pode ser modificado reforçando-se certos harmônicos. Na realidade, a verdadeira síntese sonora foi feita pela primeira vez por construtores de órgãos de tubo nos séculos XIII e XIV. Os órgãos dessa época não possuíam vários registros; ao contrário, cada tecla fazia soar um número fixo de tubos chamados *Blockwerk*, que constituíam um ou vários tubos afinados na frequência fundamental da nota escrita, mais uma série de tubos afinados, respectivamente, na oitava, décima segunda, décima quinta etc., seguindo a série harmônica (excluindo-se a sétima). A combinação particular de intensidades escolhidas para cada tubo componente determinava a qualidade particular do som do instrumento. Mais tarde, surgiram os primeiros registros múltiplos: eles permitiam ao organista ligar ou desligar seletivamente várias fileiras de tubos, correspondendo aos harmônicos no *Blockwerk*, podendo assim escolher dentre várias opções a qualidade particular do som do órgão (e alterar o volume — Seção 3.4). Somente um ou dois séculos depois é que foram acrescentados novos registros independentes, na forma de fileiras de tubos com timbres individualmente diferentes.

A síntese sonora, portanto, é uma antiga prática. Todavia, a análise sonora, *i.e.*, a individualização dos harmônicos que aparecem simultaneamente num som produzido naturalmente, não foi mencionada explicitamente

mente na literatura antes de 1636, quando o notável músico-cientista-filósofo francês padre Mersenne publicou o primeiro estudo sobre a análise (qualitativa) dos harmônicos presentes num som complexo.

Duas questões principais surgem, relativas à percepção do timbre:

1. Por que um som complexo, formado pela superposição de diferentes frequências, dá origem a apenas uma sensação de altura?
2. O que é que nos permite distinguir um espectro sonoro de outro, mesmo que a altura e o volume sejam iguais?

Embora a primeira dessas questões já tenha sido respondida em parte (Seção 2.9), vamos analisar o processo de percepção auditiva passo a passo, dessa vez para uma onda sonora complexa que chega ao tímpano. O tímpano se move para dentro e para fora periodicamente com um padrão de vibração imposto pelo padrão de vibração complexo e não-senoidal da onda. Esse movimento é transmitido mecanicamente pela cadeia de ossículos à membrana da janela oval, que reproduz exatamente o mesmo padrão de vibração complexo. Nem o tímpano nem a cadeia de ossos “sabem” que a vibração que eles estão transmitindo é formada pela superposição de diferentes harmônicos. Essa análise só é feita no próximo estágio.

A vibração complexa da membrana da janela oval emite ondas que se propagam no fluido coclear. Este é o estágio em que ocorre a separação nos diferentes componentes da frequência. Conforme foi afirmado nas Seções 2.3 e 3.2, a região de ressonância para um dado componente da frequência (região da membrana basilar onde as ondas sonoras causam excitação máxima) está localizada numa posição que depende da frequência. Assim, um som complexo dará origem a toda uma multiplicidade de regiões de ressonância (Figura 2.25), uma para cada harmônico, cujas posições reais podem ser obtidas com o auxílio da Figura 2.8. Devido à relação quase logarítmica entre x e f , as regiões de ressonância estarão cada vez mais agrupadas, à medida que subimos a série harmônica (Figura 2.25). Como cada região de ressonância espelha-se por um certo comprimento (Seção 2.4), ocorrerá uma superposição das regiões

de ressonância vizinhas, particularmente para os harmônicos superiores. Na verdade, além do sexto harmônico, mais ou menos todas as regiões se superpõem, localizando-se dentro de uma banda crítica, e não faz muito sentido considerá-las individualmente¹⁷. Cada uma das regiões de ressonância da membrana basilar oscila com sua própria frequência de ressonância e com uma fase que está relacionada (mas não é igual) à fase do harmônico correspondente (componente de Fourier), presente na oscilação inicial do tímpano. A superposição, naturalmente, complica bastante esse quadro.

Assim, um único som complexo produz uma situação extremamente complicada na cóclea. Por que, então, percebemos esse som como uma entidade única, com altura, intensidade e timbre bem definidos? Parece que isso se deve ao aspecto que essas múltiplas regiões de ressonância têm em comum: como foi dito na Seção 2.9, isso pode resultar principalmente de um processo de reconhecimento de padrões espaciais (*i.e.*, uma percepção auditiva gestáltica). O aspecto característico que é reconhecido nesse processo, comum a todos os sons periódicos independentemente de sua frequência fundamental e espectro de Fourier, é a relação de distância praticamente invariante entre máximos de ressonância na membrana basilar. A sensação de altura deve ser encarada como o “sinal de saída final” desse processo de reconhecimento¹⁸. Os componentes cruciais do som, que ativam o processador central de alturas responsável por esse processo de reconhecimento, são os primeiros seis a oito harmônicos. O mecanismo de reconhecimento do processador de alturas (o “ajuste de modelos” ou “autocorrelação espacial” discutido na Seção 2.9) pode funcionar mesmo que parte da entrada esteja faltando (p. ex., uma fundamental ausente). Nesse caso, ele pode perpetrar erros de casamento ou fornecer sensações de altura ambíguas ou múltiplas (Seção 2.7). Uma discussão mais detalhada de como esse processo de reconhecimento pode realmente operar é desenvolvida na Seção 4.10 e no Apêndice II.

17. Experimentos têm mostrado que, para instrumentos musicais comuns, devido a essa superposição (discriminação de frequências, Seção 2.4), os harmônicos, além do sétimo, mais ou menos, não podem ser ouvidos (Plomp, 1964).

18. Recomendamos enfaticamente que o leitor releia a Seção 2.9.

Devemos lembrar mais uma vez que a análise da distribuição temporal de pulsos neurais (Seção 2.9) também pode participar da determinação da altura subjetiva de um som complexo. Na verdade, quase esquecemos o outro aspecto invariante notável do padrão de excitação gerado por um som complexo (Seção 2.7): *pares de componentes consecutivos de uma mesma série harmônica fornecem uma frequência de repetição comum* (Figura 2.19). Como já foi mencionado na Seção 2.8, um processo com sugestões de tempo deve funcionar melhor na faixa dos harmônicos superiores, onde as regiões de ressonância vizinhas se superpõem, dando às fibras nervosas acústicas pertinentes uma chance de experimentar o padrão de vibração superposto (p. ex., um padrão como o da Figura 2.20). Por outro lado, como vimos na página 96, os experimentos psicoacústicos realizados com sons de baixa frequência e de duração muito curta também parecem favorecer um processo de sugestão temporal na percepção da altura: apenas uns 2 ou 3 ciclos podem dar origem a uma sensação clara de altura — um fato que não pode ser explicado por um mecanismo de análise espectral.

A percepção da altura de um som complexo é apenas uma saída dentre toda uma seqüência de tarefas neurais. É o mecanismo psicofísico que transforma o padrão de atividade periférico evocado por um som musical em outro padrão, de tal modo que todos os estímulos com a mesma periodicidade são representados semelhantemente — fornecendo a mesma sensação de altura. Um análogo óptico seria o reconhecimento de uma certa letra — independentemente da sua orientação, tamanho, cor ou tipo. Sempre que nosso tímpano é posto em vibração periódica, por mais complexa que seja, o fato de percebermos uma altura nos diz que nosso cérebro recebeu a mensagem: “Ahá! um som periódico” (em contraste com o ruído ou com as alturas múltiplas e ambíguas que percebemos num som inarmônico).

Nossa reação subjetiva aos sons complexos depende bastante do contexto do qual eles fazem parte (Houtsma e Goldstein, 1972). A realização de tarefas musicalmente “significativas”, como a identificação da fonte sonora, *i.e.*, o instrumento, e o reconhecimento de melodias ou harmonias, influencia bastante a forma como os sons complexos são processados no cérebro. Isso se aplica mesmo para a percepção de altura: experimentos

com componentes espectrais gerados eletronicamente mostram convincentemente que a atribuição de uma única altura a sons complexos é grandemente facilitada (e às vezes é até necessário) pela apresentação dos sons em teste na forma de uma melodia “significativa”¹⁹ (Houtsma e Goldstein, 1972). Os sons complexos individuais sintetizados eletronicamente, quando retirados de um contexto musical, podem freqüentemente levar a sensações ambíguas ou de altura múltipla (Seção 2.7).

Ao ouvir um som complexo, nosso sistema auditivo presta mais atenção à resposta do mecanismo central de alturas (que fornece uma única sensação de altura) do que à altura primária dos componentes harmônicos individuais. Se quisermos “tomar consciência” dos primeiros seis ou sete harmônicos de um som complexo soando continuamente, devemos ordenar um “desligamento” (inibição) do mecanismo de altura subjetiva dominante e direcionar nossa atenção para a resposta, a princípio obstruída, do mecanismo de altura espectral ou primária mais primitivo, determinado pela posição espacial das regiões ativadas da membrana basilar (Seção 2.3). Essa inibição e redirecionamento leva um certo tempo — consideravelmente maior do que o necessário para acionar o mecanismo geral de processamento de som (Seção 3.5). Eis por que os harmônicos superiores não podem ser ouvidos isoladamente em sons curtos ou de rápido decaimento²⁰. É importante observar que, em vista da assimetria da distribui-

19. Um interessante experimento para testar esse efeito de rastreamento de altura fundamental dependente do contexto pode ser realizado com o órgão. Toque uma peça (p. ex., o coral “Wenn wir in höchsten Nöten sein”, do *Orgelbüchlein* de Bach) com uma única melodia de soprano numa combinação tipo Cornet $8' + 4' + 2 \frac{2}{3}' + 2' + 1 \frac{3}{5}' + 1 \frac{1}{3}' + 1'$, acompanhada o tempo todo por um $8' + 4'$ e $16' + 8'$, respectivamente. Solicite a ouvintes musicalmente treinados que observem cuidadosamente a altura da melodia, mas alerte-os que haverá mudanças de timbre. Após os primeiros 5-6 compassos, repita a peça, mas elimine o $8'$ da melodia. Repita mais uma vez, eliminando o $4'$, depois o $2'$ e finalmente o $1'$. No final, explique aos ouvintes o que restou na voz superior e mostre que a altura da nota escrita estava totalmente ausente (em qualquer das oitavas) — eles dificilmente acreditarão! Uma repetição do experimento, porém, provavelmente fracassará, porque os ouvintes redirecionarão suas estratégias de processamento de alturas!
20. O fato de o sétimo harmônico ser uma dissonância tem incomodado os músicos há muito tempo. Todavia, essa preocupação é infundada. O sétimo harmônico é extremamente difícil de ser individualizado, mesmo em sons complexos contínuos, gerados eletronicamente

ção de atividade ao longo da membrana basilar (p. ex., Figura 3.5) e do efeito discutido rapidamente na página 145, o casamento de altura primária dos sobretons de um som complexo sempre resulta em intervalos ligeiramente “distendidos”, por exemplo, entre o primeiro e o segundo harmônicos (oitava distendida), e assim por diante (Terhardt, 1971). Esse desvio é causado pela influência perturbadora do conjunto de todos os outros harmônicos sobre o harmônico cuja frequência primária está sendo casada. O efeito é pequeno (alguns pontos percentuais), mas pode ser musicalmente relevante (Seções 5.4 e 5.5).

Obviamente, existem outros aspectos importantes do estímulo auditivo primário (ignorado pelo processador de alturas) que fornecem uma resposta, em nível de percepção, de outros estágios do processo de reconhecimento de padrões sonoros. De um som complexo percebemos o *volume* (ligado à rapidez total dos impulsos neurais, Seção 3.5) e a *qualidade* ou *timbre*. Precisamos agora fazer uma clara diferenciação entre a situação estática que surge quando ouvimos um som complexo contínuo com frequência fundamental, intensidade e espectro constantes, e a situação dinâmica mais realística na qual um som complexo com características transientes é percebido num contexto musicalmente relevante. Vamos analisar o caso estático. Experimentos psicoacústicos com sons complexos contínuos gerados eletronicamente, com mesma altura e volume, mas diferentes espectros e fases entre os harmônicos, mostram que a sensação de timbre é controlada primariamente pelo espectro de potência (Plomp, 1970, 1976). As mudanças de fase, embora claramente perceptíveis, particularmente quando operadas entre componentes de baixa frequência, exercem apenas um papel secundário. Assim, a sensação estática de qualidade ou timbre emerge como o correlato perceptivo da distribuição de atividade evocada ao longo da membrana basilar — desde que a relação correta de distância entre os picos de ressonância esteja presente para unir tudo numa sensação de “som único”. Dividindo a gama de frequências audíveis em faixas de

(Plomp, 1964). Na verdade, alguns órgãos modernos têm um registro $1/7'$ soando o sétimo harmônico da nota escrita, o que resulta num timbre muito particular quando usado sabiamente com outros registros.

cerca de um terço de oitava cada uma (correspondendo aproximadamente a uma banda crítica, Seção 2.4) e medindo a intensidade ou fluxo de energia sonora que, para determinado som complexo, está contida em cada faixa, podemos definir “índices de dissimilaridade” quantitativos para os sons (contínuos) dos diversos instrumentos musicais. Esses índices têm um correlato muito grande com os julgamentos de similaridade e dissimilaridade de timbre determinados psicofisicamente (Plomp e Steeneken, 1971). É importante ressaltar que a sensação de timbre é controlada pela distribuição absoluta de energia sonora nas bandas críticas, e não pelos valores de intensidade relativos ao da fundamental. Uma única medida dessa distribuição é dada pela “frequência centróide”, a frequência média do espectro de um som (Grey e Gordon, 1978). Isso pode ser verificado facilmente ouvindo-se um disco ou fita tocado em rotação alterada. Esse procedimento deixa os espectros de potência relativos inalterados, simplesmente deslocando todas as frequências para cima e para baixo; ainda assim, uma clara mudança no timbre de todos os instrumentos é percebida.

A sensação estática de timbre é uma magnitude psicológica “multidimensional” relacionada não com um mas com todo o conjunto de parâmetros físicos do estímulo acústico original (o conjunto das intensidades em todas as bandas críticas)²¹. Essa é a principal razão pela qual é mais difícil fazer descrições semânticas do timbre do que da altura ou volume, que são “unidimensionais”. Com exceção das amplas denominações que vão de “opaco” ou “abafado” (poucos harmônicos superiores) a “nasal” (principalmente harmônicos ímpares) e a “brilhante” ou “metálico” (muitos harmônicos superiores realçados), a maior parte das qualificações dadas pelos músicos invoca uma comparação com sonoridades instrumentais (“como flauta”, “como cordas”, “como madeira”, “como metal”, “som de órgão” etc.).

21. Embora haja cerca de 15 bandas críticas na extensão das frequências musicalmente relevantes, cujas intensidades devem ser especificadas a fim de se determinar o espectro, um estudo de identificação de vogais (Klein, Plomp e Pols, 1970) indica que apenas quatro parâmetros independentes de intensidade (cada um sendo uma combinação linear específica das intensidades em todas as bandas críticas) são suficientes para especificar um som complexo dentro da “capacidade de resolução de timbre” do sistema auditivo.

Os aspectos dinâmicos da percepção do timbre estão intimamente relacionados ao processo de identificação da fonte sonora, *i.e.*, à identificação do instrumento (Seção 4.9). Os atributos relevantes dos sinais de entrada são muito mais difíceis de se quantificar (p. ex., Iverson e Krumhansl, 1993); a envoltória de amplitude do som, em particular, o ataque e o decaimento são os três elementos mais importantes. A similaridade ou dissimilaridade do timbre depende de atributos mais gerais e estáticos do som complexo, como a frequência centróide (que não muda muito ao longo da duração de um som). A identificação, por outro lado, depende principalmente dos atributos característicos do estabelecimento do som, que contêm informações cruciais sobre o instrumento em questão, como o tipo de mecanismo de excitação (pinçamento, percussão, fricção, sopro). Evidentemente, a identificação ocorre num estágio muito mais elevado da função cerebral cognitiva do que a sensação estática do timbre (veja a próxima Seção).

4.9 IDENTIFICAÇÃO DE SONS MUSICAIS

Um aspecto característico de todos os processos de reconhecimento de padrões é a “perda de informação seletiva”. Isso está relacionado à forma de operar do sistema nervoso: mínimo esforço, máxima eficiência. Para ser capaz de selecionar estímulos portadores de informação significativa dentre a terrível complexidade da entrada sensorial total (e, assim, ser capaz de identificar objetos e as suas inter-relações causais), o sistema deve contar com uma série de “filtros” que ajudam a separar os detalhes que são relevantes daqueles que não o são. Esses “filtros” devem estar “sintonizados” com certas características invariantes dos estímulos e que são consideradas relevantes (por transferência de informação genética, ou como resultado de um processo de aprendizagem). Na percepção sonora, o primeiro e mais “primitivo” elemento de entrada a ser reconhecido pelo sistema nervoso provavelmente é a intensidade, com o seu correlato perceptivo, o volume. Esse

correlato não depende muito de nenhum outro detalhe de estrutura do estímulo, exceto do fluxo total de energia acústica (Seção 3.5). O próximo elemento auditivo de entrada a ser considerado é a periodicidade do som (representada pela distribuição espacial dos máximos de ressonância, ou pela distribuição temporal de pulsos neurais). Desse processo de reconhecimento é extraída a sensação de altura subjetiva. O terceiro nível de refinamento é a consideração do espectro de potência do som, que leva à sensação de timbre.

A percepção de timbre é, na verdade, apenas um primeiro estágio da operação de reconhecimento de fonte sonora — em música, a identificação do instrumento. Desse ponto de vista, a percepção do timbre é o mecanismo pelo qual se extrai informação do sinal auditivo, de modo a prepará-la para:

1. Armazenagem na memória, com um rótulo adequado de identificação.
2. Comparação com informações previamente armazenadas e identificadas.

A primeira operação envolve aprendizado ou condicionamento. Uma criança que aprende a reconhecer um certo instrumento musical ouve repetidamente uma melodia tocada nesse instrumento, e lhe dizem: “Esse é um clarinete”. O seu cérebro extrai a informação pertinente dessa série de estímulos auditivos, rotula essa informação com a qualificação “clarinete” e armazena-a na memória. A segunda operação representa a resposta condicionada ao padrão aprendido: quando a criança ouve um clarinete, após a sua experiência de aprendizado, seu cérebro compara a informação extraída do sinal de entrada (*i.e.*, o timbre) com informações armazenadas e, se encontra uma boa identificação, transmite a resposta: “Um clarinete”. Por outro lado, se ouvimos um som “novo”, p. ex., uma série de notas executadas num sintetizador eletrônico, nosso sistema de extração de informações fornecerá as pistas para o mecanismo de comparação, que tentará desesperadamente comparar a entrada com informação previamente armazenada. Se essa comparação não der certo, um novo “arquivo” de armazenagem será eventualmente aberto para esse novo timbre já identificado. Se o processo for apenas parcialmente bem-sucedido, reagimos

com julgamentos como: “Quase como um clarinete”, ou “como um trombone ganindo”.

Os músicos protestarão e dirão que existe muito mais uma sensação de timbre do que pistas para se descobrir “o que está tocando”. Por exemplo, o que faz com que o som de um instrumento seja mais “bonito” do que o de outro da mesma espécie? Inicialmente, devemos ressaltar que isso está relacionado obviamente a um estágio de sofisticação bem além do mecanismo de identificação já mencionado: podemos aprender a extrair uma quantidade de informação cada vez mais refinada dos padrões de vibração de um instrumento, de tal modo que seremos capazes de distinguir entre diferentes amostras de instrumentos do mesmo tipo. Por que alguns padrões de vibração são mais “bonitos” do que outros, isso não se sabe. Muita pesquisa tem sido feita, por exemplo, para se descobrir quais as características físicas que fazem do violino Stradivarius um grande instrumento (p. ex., Saunders, 1946). Muitas dessas características são dinâmicas em caráter, e a maioria delas parece estar mais relacionada à maior ou menor facilidade que o instrumentista tem de controlar a “cor tonal” desejada (espectro e transientes) do que a um efeito “passivo” sobre o ouvinte. Por exemplo, um aspecto significativo do som do violino parece estar relacionado ao efeito do pequeno espaçamento entre os picos de ressonância (p. ex., Figura 4.15) sobre o volume e o timbre quando a frequência fundamental do som é modulada pelo músico num *vibrato* (Matthews e Kohut, 1973). Nessas circunstâncias, as frequências dos componentes harmônicos oscilam acima e abaixo dos picos de ressonância estreita e irregularmente espaçados. Como resultado, a amplificação de cada componente varia periodicamente, o mesmo ocorrendo com o volume e o timbre do som. Dependendo da microestrutura particular da curva de ressonância do seu instrumento, o músico tem a possibilidade de induzir variações extremamente sutis no volume e timbre acoplados ao seu *vibrato*. Em grande parte, a impressão do ouvinte também está baseada em experiência aprendida: desde o início de nossa formação musical somos treinados a rotular o que é bonito e o que não é! E nosso cérebro aprendeu a construir padrões de leitura correspondentes para identificação, baseado num

processamento extremamente refinado de informação, a fim de reconhecer estruturas sutis, mínimas, praticamente incomensuráveis, do sinal acústico de entrada.

4.10 PROCESSOS COGNITIVOS CEREBRAIS RELATIVOS À PERCEPÇÃO DE SONS INDIVIDUAIS

Concluímos este capítulo com uma breve revisão dos processos cerebrais superiores envolvidos no reconhecimento de sons musicais. Começamos com uma discussão sobre alguns conceitos básicos de vida, informação, e da evolução da função cerebral humana (Roederer, 1978, 1979).

Os sistemas biológicos, desde os microorganismos até os primatas, são “ilhas” de matéria organizada que persistentemente evoluem em direção a uma ordem crescente. Para estar apto a gerar ordem, um sistema orgânico deve se preservar invariante durante um certo intervalo de tempo finito, a despeito de suas interações com o ambiente variável. Ele deve operar de maneira orientada para o futuro, ou seja, manter-se funcional, seguindo linhas de ação que sejam favoravelmente adaptadas às mudanças do ambiente, e deve estar apto a realizar mudanças no ambiente para sua vantagem própria. É essa a diferença individual mais significativa entre os sistemas vivos e não-vivos. Caracteriza-se pelo fato de que as interações dos sistemas vivos entre eles e com o ambiente, sejam simples ou complexas, estão baseadas em processos envolvendo reconhecimento de padrões e processamento de informações.

O reconhecimento de padrões requer a geração consistente de respostas idênticas a diferentes e complexas constelações de entrada que possuam um certo aspecto temporal ou espacial em comum (o padrão que está sendo reconhecido), independentemente do tamanho, posição, orientação ou ordem particular em que esse aspecto apareça entre todos os outros componentes de entrada (p. ex., um arranjo espacial particular de radicais de nucleotídeos na molécula de DNA exercendo um efeito específico desse padrão sobre o seu ambiente químico; os detalhes geo-

métricos do contorno de um objeto cuja imagem óptica é projetada na retina, dando início a operações cognitivas específicas desse padrão no cérebro; o conjunto de harmônicos de um som musical, que gera distintas regiões de ressonância na membrana basilar, dando origem a uma sensação específica de altura). Concluindo, é esse modo de reconhecimento de padrões ou de processamento de informações que permite a um sistema biológico gerar e manter a ordem de maneira consistente e passar informação relevante aos seus descendentes.

Com a evolução das espécies, a informação sobre o ambiente foi gradualmente incorporada e armazenada nas estruturas de memória genética dos organismos. Mudanças muito lentas no ambiente, numa escala de tempo muito maior do que a duração da vida de uma geração da espécie, puderam ser incorporadas nas estruturas de memória dos gens por meio de uma ação combinada de mutação, diversificação e seleção natural. Como as espécies foram se tornando cada vez mais complexas, e cada vez mais e mais características imprevisíveis do ambiente terrestre se tornaram determinantes para a sobrevivência, a capacidade de adaptação ontogênica durante o tempo de vida de um organismo se tornou uma exigência fundamental para a sobrevivência,

Quando surgiu a locomoção nos organismos multicelulares, o número de variáveis ambientais relevantes aumentou drasticamente, com escalas de tempo de mudança de até frações de segundo; tornou-se necessário absorver uma incrível quantidade de informação pelos sentidos. Embora a maior parte da afluência de informação sensorial seja irrelevante, ela leva embutida de uma forma *a priori* imprevisível os sinais dos padrões que são decisivos para a sobrevivência do organismo. O sistema nervoso evoluiu a fim de dotar os organismos superiores com a capacidade de detectar, selecionar, identificar e armazenar informação relevante contida na complexa entrada sensorial, e antecipar e reagir apropriadamente às rápidas mudanças do ambiente. Nesse percurso do desenvolvimento, o que começou como um simples aparelho de transmissão e conversão de sinais ambientais evoluiu para o sistema nervoso central (SNC) dos vertebrados superiores, com sofisticadas habilidades de análise de entradas e planejamento de respostas, com base no estado momentâneo do ambiente, no

estado do organismo, em instruções inatas (instintos) e em informações aprendidas (Figura 4.29). O cérebro emergiu como o processador central, realizando as operações fundamentais necessárias para monitorar e controlar as condições vivas do organismo e otimizar a sua *performance* num ambiente complexo e variável.

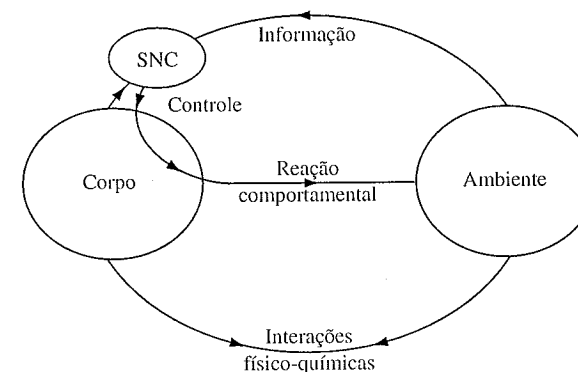


Figura 4.29 Informação e controle por meio do sistema nervoso central (Roederer, 1987).

Um último desenvolvimento crucial veio com o surgimento do cérebro humano. Aristóteles já reconhecia que “os animais têm memória e a capacidade da instrução, mas nenhum animal, exceto o homem, pode recordar o passado quando quiser”. Mais especificamente e em termos modernos, a operação mais fundamental que o cérebro humano, e só ele, pode realizar é recuperar a informação armazenada, assim como as imagens e representações, analisá-las, manipulá-las, e armazenar versões modificadas ou melhoradas delas sem nenhuma entrada sensorial externa. Damos a esses atos de alteração de imagens e de rearmazenagem sem necessidade de entradas externas o nome de “processo de pensamento humano”. Esse processo envolve mais do que o processamento da informação existente: ele representa a geração de informação nova (Roederer, 1978). Ligado a esse desenvolvimento estava o surgimento da fala proposicional e de outros aspectos distintos da linguagem humana. A capacidade de rearran-

jar e alterar imagens levou à possibilidade de previsões e planejamento a longo prazo e à consciência do tempo futuro. Com essa capacidade, tornou-se possível o adiamento de objetivos comportamentais e o replanejamento das prioridades para isso: o corpo passou a servir o cérebro, e não o contrário. Finalmente, a armazenagem planejada de informação no ambiente para uso posterior tornou-se outra característica distintiva da habilidade do cérebro humano (Roederer, 1978); essa “externalização da memória” levou à documentação de eventos e sentimentos por meio da linguagem escrita e das artes e, portanto, ao desenvolvimento da cultura humana como tal²².

Durante as últimas duas décadas houve um grande progresso na compreensão das funções cerebrais. Muito disso foi conseguido pouco a pouco e reunido em algo que ainda parece mais um conjunto de hipóteses científicas do que um modelo unitário e quantitativo, experimentalmente verificado, da função cerebral. Ainda assim, é impressionante a evidência indireta de alguns modos cruciais de operações cerebrais, embora nem sempre livres de controvérsias.

Em primeiro lugar, vamos definir que espécie de informação externa precisa ser reconhecida nos vários sistemas sensoriais. No sentido visual, são as entidades discretas com contornos geralmente bem definidos e superfícies com textura — às quais chamamos de “objetos no espaço” —, suas relações espaciais mútuas e variações no tempo. O sentido auditivo reconhece “objetos no espaço acústico” (que na realidade são “objetos no tempo”), *i.e.*, trens discretos de ondas sonoras com certas marcas bem definidas, as relações entre eles e entre suas fontes no ambiente. Já dissemos várias vezes que a percepção sensorial se desenvolve em vários pas-

22. Os animais, obviamente, também deixam “impressões” de informação no ambiente. Todavia, há uma diferença crucial. Se subitamente destruíssemos todas as impressões animais da Terra (ninhos, reservas de alimento, trilhas etc.), a vida animal ainda teria condições de continuar essencialmente como é hoje: a informação essencial para a sobrevivência e preservação está geneticamente armazenada no organismo; ruínas de “culturas animais” não são usadas por gerações posteriores como fonte de informação. Mas se destruíssemos de uma só vez todas as “impressões” humanas, nossa civilização desaparecería como nós a conhecemos, e as futuras gerações se encontrariam no estado em que a humanidade esteve há milhares de anos.

sos operacionais e fisiológicos bem definidos. Os sinais externos são primeiro convertidos em sinais neurais, que são então transportados às áreas receptoras corticais primárias (p. ex., veja a Figura 2.26 para o sistema auditivo), de onde são transferidos às áreas secundária e de associação, aos lobos frontais e ao resto do cérebro. A Figura 4.30 oferece um esquema hipersimplificado dos principais passos do processamento de informação sensorial pelo sistema visual; o sistema auditivo apresenta estágios semelhantes (Roederer, 1987).

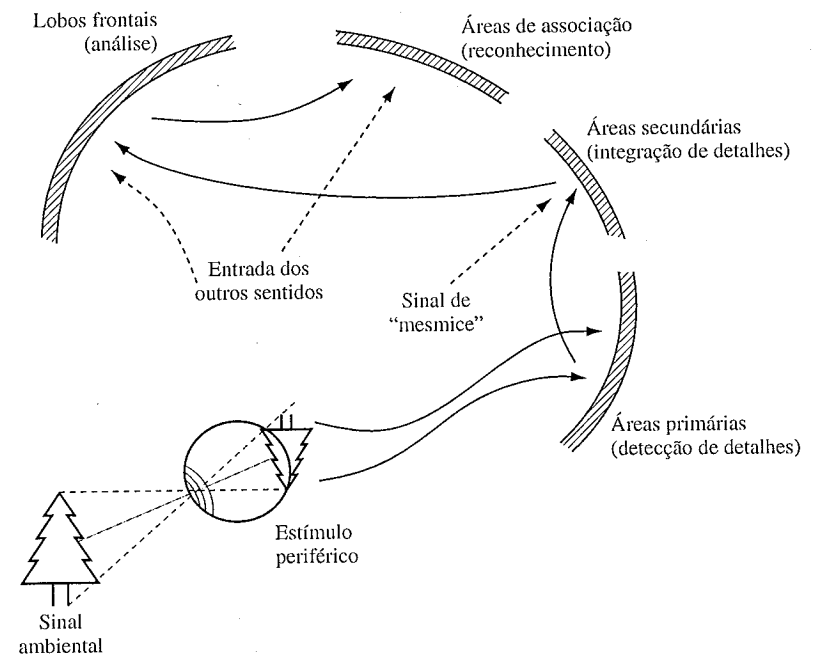


Figura 4.30 Estágios da percepção de um objeto (segundo Roederer, 1987).

O circuito neural nos caminhos periférico e aferente, até a área receptora cortical primária, está quase sempre apto a executar operações básicas de pré-processamento de detecção de detalhes (p. ex., detecção de bordas e movimento, fusão binocular na visão; altura espectral, detecção de

transientes, localização binaural na audição). As medições mostraram que os neurônios das áreas corticais conectados ao sistema de transmissão aferente de um órgão sensorial são na verdade “detectores de características”, respondendo a características físicas bem definidas, mas complexas, do estímulo sensorial original. No córtex visual, por exemplo, foram encontrados neurônios que respondem somente a barras ou bordas escuras ou luminosas numa certa parte do campo visual, inclinadas em um certo ângulo, ou a uma linha que se move de um certo modo (Hubel, 1971). Também no córtex auditivo há neurônios que parecem responder apenas a certos tipos de som complexos (Simmons, 1970). A integração dessas características é a operação necessária para selecionar, numa entrada incrivelmente complexa, os detalhes que pertencem ao mesmo objeto (*i.e.* juntar essas bordas que definem os contornos de um certo objeto; selecionar as regiões de ressonância da membrana basilar que se referem a um certo som complexo). Essa operação também envolve transformações geométricas e topológicas (quase sempre operações aprendidas) que informam o cérebro sobre a “mesmice” ou invariância de um certo objeto visual quando ele é rodado ou transladado no espaço, ou de um objeto acústico, como um som complexo, quando sua altura fundamental ou intensidade mudam (o processo de coincidência ou casamento de modelos mencionado na página 94). Nesse estágio, o cérebro “sabe” que é um objeto; o reconhecimento, *i.e.*, a determinação de qual é realmente o objeto, requer um processo de comparação com uma informação existente, previamente adquirida (identificação do objeto; identificação do instrumento ao qual pertence o som). O reconhecimento pode ser “automático” (lembrança associativa) ou precisar de uma análise de toda a entrada sensorial nos lobos frontais. O estágio final corresponde à integração dos objetos na cena inteiramente percebida (a paisagem sendo vista, a música sendo ouvida).

À medida que vamos para os estágios superiores da Figura 4.30, o processamento de informações se torna menos automático e mais sujeito ao controle central; são necessárias mais ações e decisões controladas por motivação (atenção, seleção de informação irrelevante), e cada vez mais a informação previamente armazenada (aprendida) influenciará o processo. Nos estágios superiores, os aspectos evocativos permeiam todas as

fases do processo cognitivo. A “automação” dos estágios inferiores do processamento de informação dá lugar a um controle por motivação que decidirá se uma certa operação deverá ou não ser executada. Apenas nos seres humanos esse controle por motivação pode, por sua vez, ser regulado ou rejeitado por operações cognitivas ainda mais superiores — mas ele não pode ser descartado do circuito (veja a Seção 5.6). Essa talvez seja a diferença mais importante entre a abordagem do processamento de informação sensorial feita pela “inteligência artificial” dos robôs e as “operações cognitivas reais” do cérebro; estas últimas têm um propósito primordial bem diferente do dos robôs e computadores: otimizar as condições para a sobrevivência!

O que de fato ocorre fisicamente no sistema nervoso quando a informação sensorial é reunida e os objetos são representados, reconhecidos e integrados no panorama ambiental? O que acontece quando os objetos são imaginados? Na Seção 2.9, ressaltamos que a informação sensorial (ou qualquer outro tipo de informação) está codificada no sistema neural sob a forma de distribuição espacial e temporal de impulsos elétricos²³. Definimos como “representação” mental ou cerebral de um objeto (visual, acústico, olfativo, tátil) a distribuição específica de sinais elétricos na rede neural do córtex cerebral que aparece numa correspondência um-a-um, causal, com as características específicas percebidas durante a apresentação desse objeto particular (Roederer, 1979). De acordo com essa definição, a “cognição” nada mais é do que a ocorrência de uma distribuição específica de atividade neural em certas áreas do córtex e que está em correspondência biunívoca com o objeto, evento ou conceito que está sendo reconhecido ou imaginado. Por exemplo, as distribuições de atividade neural provocadas nas áreas corticais intervenientes pela percepção das seguintes coisas — uma grande maçã vermelha, uma pilha de pequenas maçãs verdes, uma macieira, um pomar de macieiras — embora bem dife-

23. É importante lembrar que também há um sistema neuroquímico de informação (peptídeos cerebrais liberados por um mecanismo secretor neural especializado) que regula o estado geral das funções cerebrais (e somáticas), agindo como uma espécie de “controle de volume” sobre centros cerebrais específicos. Esse sistema de neuromoduladores é “lento” e na verdade apenas transmite informações (pela corrente sanguínea) — ele não processa informações.

rentes, devem ter em comum algum subconjunto dessa distribuição de atividade neural: aquela que define e aparece em correspondência com a cognição de “maçã”. Sempre que reconhecemos os sons de um clarinete (p. 218), estará ocorrendo alguma distribuição única de atividade neural que é específica dos sons do clarinete.

O ato da lembrança, ou memória, de um evento sensorial consiste na re-elicitação ou “reprise” daquela distribuição particular de sinais neurais específica do evento sensorial original. Quando esse padrão é gerado externamente enquanto ouvimos um som, lembramos ou reconhecemos que esse som provém, por exemplo, de um clarinete. Quando essa atividade é gerada internamente (por alguma associação ou comando sujeito à vontade), estaremos nos lembrando do som do clarinete na ausência do som externo real. Assim, isso representa a forma mais simples de ativação do mecanismo de imagem acústica. Experimentos com a visão mostraram que, por exemplo, a mera imaginação de uma forma geométrica evoca uma atividade no córtex visual semelhante àquela evocada externamente quando o indivíduo realmente vê essa forma (Herrington e Schneidau, 1968). Hoje, temos evidências experimentais convincentes de que todo o processo esquematizado na Figura 4.30 também funciona no sentido inverso: quando somos solicitados a imaginar uma maçã vermelha, um “estímulo” apropriado dos lobos frontais é transmitido às áreas de recepção e associação visuais, onde é evocada quase a mesma atividade neural específica que ocorreria se estivéssemos realmente vendo a maçã. Um compositor que “imagina” uma seqüência de sons evoca uma atividade neural específica das áreas acústicas que é quase a mesma que seria produzida se esses sons estivessem realmente sendo ouvidos. Em qualquer processo imaginativo, a atividade neural específica também pode ter sido iniciada por um processo alucinatório, durante um sonho, ou pela estimulação eletrofisiológica do cérebro durante uma neurocirurgia.

Enquanto nos estágios inferiores e periféricos do processamento de informação neural, conjuntos relativamente pequenos de neurônios espacialmente contíguos executam funções definidas (p. ex., os detectores de “miau” no trato auditivo do gato), nos níveis superiores do cérebro até mesmo a representação neural mais “simples” aciona centenas de milhões

de neurônios corticais simultaneamente, e não há continuidade espacial da rede neural em atividade nessa distribuição correspondente²⁴. Esses fatos tornam a verificação dos processos de representação e cognição neurais uma tarefa extremamente difícil. Todavia, existem experimentos baseados em registros de microeletrodos, e de potencial evocado, de padrões macroscópicos coerentes de atividade neural no cérebro e, mais recentemente, técnicas de tomografia de pósitrons e de ressonância magnética nuclear que proporcionam uma sustentação sólida a essa descrição.

Como funciona tudo isto? Uma vez que os organismos superiores tiveram que confiar cada vez mais na informação adquirida e armazenada no cérebro durante o seu próprio tempo de vida, a necessidade de armazenagem de informação adequadamente protegida e rapidamente acessível passou a existir desde o início. Isso levou mais à memória distribuída e ao modo de representação holológica (Pribram, 1971; Kohonen, 1988) do que a uma codificação e geração de imagens “fotográficas” das cenas ambientais. Uma representação “fotográfica” seria aquela em que existe uma correspondência “ponto-a-ponto” entre as características do estímulo (objeto) e as características da distribuição de atividade neural (imagem). Uma representação holográfica, “holológica” em nosso caso, é aquela onde a informação sobre cada característica do objeto (ponto) é mapeada em todo o domínio da imagem (como num holograma óptico, onde a informação sobre um ponto está espalhada em todo o domínio do filme fotográfico). Um dano parcial à imagem fotográfica pode levar à perda total de informação (p. ex., a obliteração da face na fotografia de uma pes-

24. Isso significa que quando você imagina, digamos, uma árvore, a “atividade neural específica” citada acima e que ocorre na chamada “tela da sua mente” não é uma área em forma de árvore constituída por neurônios ativados nos córtices visuais dos seus lobos occipitais, mas um conglomerado horrivelmente complexo e descontínuo de milhões de neurônios disparadores de impulsos. É a especificidade dessa complexa distribuição de neurônios que é única e realmente representa a imaginação (ou representação visual real) dessa árvore. Não precisamos nos admirar com o fato de que o “caos altamente organizado” em atividade em nosso cérebro é conhecida como a “simplicidade natural da função cerebral”! Por outro lado, não é preciso se admirar porque, para compreender claramente o processo neural de representação mental, você precisa ter a familiaridade de um físico com as funções de distribuição e suas transformações, processos não-lineares, fenômenos cooperativos e teoria do caos!

soa), enquanto um dano parcial a um holograma em geral apenas produz um embaçamento da imagem restaurada, e não a sua perda completa.

Já afirmamos acima que nos estágios periféricos do trato auditivo existe uma organização tonotópica de sinais neurais em correspondência espacial (e temporal) com as regiões de excitação da membrana basilar (por exemplo, as distintas regiões de ressonância eliciadas pelos primeiros 7 ou 8 harmônicos de um som musical). Mas, à medida que nos dirigimos à área cortical, essa representação “fotográfica” ponto-a-ponto vai se perdendo gradualmente e a representação neural se torna holológica, com a atividade de um ponto da membrana basilar mapeada num amplo e difuso grupo de neurônios.

Uma grande vantagem do modo de armazenagem holológica é o processo de memória associativa (Kohonen, 1988). Na verdade, a reprise de uma atividade neural específica pode ser provocada por outras causas e indícios, além da total reativação sensorial do evento original — uma reativação parcial da atividade neural ocorrida durante o ato da armazenagem é suficiente para liberar toda a configuração da atividade específica (p. ex., o efeito da ausência da fundamental; todas as ilusões de óptica). Os sistemas comuns de armazenagem de informação, como os livros, fitas, discos e filmes, representam o que chamamos de modo “fotográfico” de armazenagem na memória, e para recuperar um determinado bloco de informação devemos conhecer o seu endereço ou perscrutar todo o registro de memória (é por isso que também o chamamos de armazenagem em memória endereçável). Mas, com um sistema holológico de memória de música gravada (que ainda não existe na prática!), poderíamos tocar ou cantar as famosas quatro notas tam-tam-tam-taam e recuperar toda a *Quinta Sinfonia* de Beethoven (uma armazenagem em memória endereçável a um conteúdo distribuído)! Pois é exatamente assim que o nosso cérebro trabalha!

Como realmente acontece no cérebro essa armazenagem e recuperação holológica de informação? Durante certa fase da aprendizagem, as áreas de associação do cérebro e os sistemas de avaliação subjetiva nos lobos frontais exibem padrões de atividade neural provocados por uma série de constelações de estímulos que chegam simultaneamente por

meio de diversos canais diferentes de entrada (por exemplo, as diferentes modalidades sensoriais). Quando a aprendizagem se completa, terão acontecido mudanças sinápticas adequadas no circuito neural, de modo que a constelação de estímulos em apenas um canal de entrada (o estímulo condicionado ou “palavra-chave”) será suficiente para gerar todo o padrão de atividade neural específica do evento original completo. Em outras palavras, o aprendizado não é representado pela armazenagem de informações ou imagens por si, mas por modificações apropriadas da rede de processamento de informação²⁵, e a memória surge nesse esquema como a armazenagem de instruções sobre o processamento de informações. Eis por que o cérebro é chamado de sistema “auto-organizador”. Os aficionados em computador perceberão que na atuação cerebral não há *software*: memória, instruções e operações estão todas baseadas em mudanças apropriadas do *hardware* (a arquitetura e eficiência das conexões sinápticas entre os neurônios). A auto-organização também é o princípio segundo o qual operam os “computadores neurais” de hoje (embora as “mudanças de *hardware*” ainda sejam simuladas com programas de *software* apropriados — p. ex., veja Hinton, 1992). O Apêndice II trata mais desse assunto e mostra um modelo simples de processador de altura central operando holologicamente.

25. Uma demonstração simples mas bem convincente disso é dada pelo fato de que já se verificou que a densidade sináptica nos córtices do gato depende grandemente da complexidade do ambiente onde o animal cresceu e das respostas comportamentais.

SUPERPOSIÇÃO E SEQÜÊNCIAS DE SONS COMPLEXOS E A PERCEPÇÃO DA MÚSICA

No decorrer dos Capítulos 2, 3 e 4, subimos gradualmente a escada do processamento neural de sinais acústicos, dos mecanismos que levam à percepção da altura primária, altura subjetiva, volume e timbre, até o reconhecimento de um instrumento musical. No lado físico, analisamos como as características sonoras que levam a essas sensações são realmente geradas nos instrumentos musicais. Esses atributos psicológicos são ingredientes necessários, porém não suficientes, da música. A música é feita de seqüências e superposições de sons que carregam certos significados que podem ser analisados, armazenados e comparados no cérebro. Todavia, é muito difícil realizar um estudo quantitativo da neuropsicologia desses processos de alto nível.

5.1 SUPERPOSIÇÃO DE SONS COMPLEXOS

A música polifônica consiste em superposição de sons complexos. Mesmo que uma única melodia esteja sendo executada em música monofô-

nica, o que atinge nossos ouvidos é uma superposição de sons reverberantes, levando à superposição de sons complexos. O estudo psicofísico dos efeitos das superposições de sons complexos está ainda muito incompleto. Isso se aplica particularmente à compreensão de como o cérebro é capaz de decifrar a “confusão” de freqüências pertencentes a sons complexos diferentes soando simultaneamente, de modo a manter separadas as sensações desses sons.

Quando dois sons complexos de alturas diferentes são superpostos, uma das duas situações seguintes pode ocorrer: a freqüência fundamental do som mais agudo é igual a um dos harmônicos do som mais grave, ou não. No primeiro caso, o som mais agudo irá reforçar certos harmônicos do som mais grave. Por que então nós simplesmente não detectamos uma variação de timbre no som mais grave, em vez de selecionar o som mais agudo e até manter distintos os timbres de ambos os sons? Um problema semelhante surge no segundo caso, onde cada som produz a sua própria multiplicidade de regiões de ressonância na membrana basilar. De que forma nosso cérebro seleciona, dessa mistura resultante, qual seqüência pertence a qual som? Por exemplo, considere a superposição de dois sons complexos, digamos lá₂ (110 Hz) e dó₃# (140 Hz). Os harmônicos de ambas as notas estão na Figura 5.1, numa escala linear de freqüências. Para cada uma dessas freqüências existe uma região de ressonância correspondente na membrana basilar (Figura 2.8). A menos que haja pequenas modificações na intensidade ou na altura (e isso ocorre na música real — veja adiante), as células sensoriais não têm nenhuma informação sobre a qual nota pertence cada região de ressonância. Portanto, essa discriminação deve ser realizada num centro ainda mais superior no sistema neural auditivo.

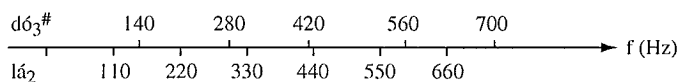


Figura 5.1

A altura das duas notas é discriminada pelo processador central de altura (Seção 2.9 e Apêndice II).

Uma capacidade muito impressionante do sistema neural auditivo é a de discriminar o timbre de dois sons complexos soando simultaneamente. A música real não seria possível sem essa capacidade. Por exemplo, suponha que você esteja ouvindo, monauralmente, com fones de ouvido, o instrumento 1 tocando exatamente a nota lá₄, e o instrumento 2 tocando exatamente a nota lá₅, com níveis de intensidade quase iguais. A Figura 5.2 mostra a superposição hipotética. O comprimento total das barras verticais representa a intensidade total de cada harmônico que realmente chega ao ouvido. Como nosso cérebro procede para manter separados os dois sons, em termos de timbre? Esse mecanismo de discriminação ainda não é conhecido; parece que um elemento temporal exerce o papel principal. Inicialmente, o começo (ataque) dos dois sons “simultâneos” nunca está exatamente sincronizado, nem a formação de cada som se desenvolve da mesma maneira, especialmente se os sons vierem de direções diferentes (efeito estéreo). Como mencionamos antes, o ataque de um som é o atributo mais importante para a identificação do timbre (Iverson e Krumhansl, 1993). Durante esse período transiente, o mecanismo de processamento em nosso cérebro parece capaz de se concentrar em certos detalhes característicos do padrão de vibração de cada instrumento e de se manter na pista desses detalhes, mesmo que eles sejam perturbados ou eclipsados pelo sinal do outro instrumento. Também numa apresentação, ligeiras excursões em freqüência e intensidade (o chamado efeito *chorus*) e que são coerentes para toda a série harmônica de cada som são usadas pelo mecanismo de processamento auditivo (Seções 4.9 e 4.10).

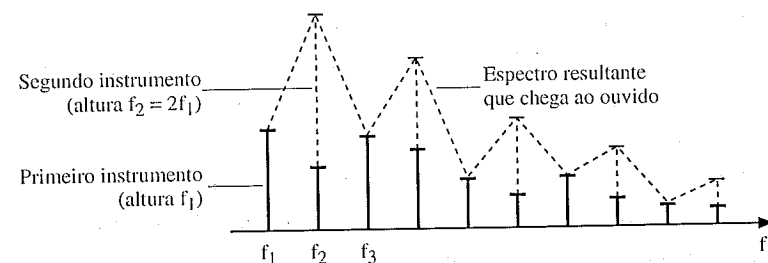


Figura 5.2

Outras variações mais pronunciadas, que aparentemente fornecem pistas importantes ao mecanismo para discriminar timbres, são as variações periódicas de altura (*vibrato*) e intensidade (*tremolo*, *vibrato*) que podem ser evocadas (voluntária ou involuntariamente) nos sons contínuos de muitos instrumentos musicais. A superposição de sons complexos totalmente destituídos dessas pequenas perturbações coerentes dependentes do tempo — como acontece quando se combinam muitos registros na música de órgão — é na verdade muito mais difícil de ser decifrada em termos de timbre.

O que provavelmente mais ajuda o mecanismo de discriminação de alturas e de timbres de sons complexos é a informação obtida de uma seqüência de superposições de sons. Nesse caso, podem ser usadas não só as pistas “primárias” já mencionadas, fornecidas pelas flutuações de cada nota em tempo e freqüência, mas também a informação secundária ou de “ordem superior” extraída das linhas melódicas (as “mensagens” musicais) tocadas por cada instrumento (ver também as pp. 216-217).

O mecanismo de discriminação de sons complexos tem o seu equivalente (e, provavelmente, a sua raiz primordial) no mecanismo que orienta a nossa percepção auditiva quando acompanhamos a conversa de uma determinada pessoa entre muitas conversas diferentes travadas simultaneamente em níveis sonoros semelhantes. Essa habilidade tem sido chamada de “efeito festa”, e possivelmente usa as mesmas pistas, primárias e secundárias, que o mecanismo de discriminação de sons complexos. Finalmente, é provável que esse mesmo mecanismo nos possibilite discernir os sons individuais dentre a confusa superposição de sons em uma sala muito reverberante (Seção 4.7). Nesse caso, novamente é o efeito tempo que parece desempenhar o papel decisivo: a primeira chegada do som direto (Figuras 4.26 e 4.27) fornece as pistas principais, nas quais o nosso sistema de percepção se concentrará para definir a sensação real do som e a sua discriminação. Esse é o chamado efeito de precedência em acústica ambiental.

Como acontece com dois sons puros, existe uma diferença mínima nas freqüências fundamentais que dois sons complexos devem ter para que possam ser ouvidos separadamente. Quando a diferença em altura entre dois sons complexos é menor do que o limite de discriminação de sons,

podem surgir batimentos de primeira ordem entre todos os harmônicos (Seção 2.4). Se, por exemplo, ambos os sons complexos são uníssonos desafinados, com freqüências fundamentais f_1 e $f_1 + \varepsilon$, respectivamente, todas as regiões de ressonância da membrana basilar vão se superpor e produzir batimentos de diferentes freqüências. As fundamentais terão batimentos com freqüência ε , os segundos harmônicos com freqüência 2ε , e assim por diante. Apenas os primeiros harmônicos são importantes; em geral, os batimentos existentes entre as fundamentais são os mais evidentes (ε). Os batimentos de primeira ordem entre os harmônicos correspondentes também vão aparecer se nós superpusermos dois sons complexos que formam outros intervalos musicais desafinados. Estes são bem diferentes (embora iguais em freqüência) dos batimentos de segunda ordem que surgem em intervalos desafinados entre sons puros (Seção 2.6).

5.2 A SENSACÃO DA CONSONÂNCIA E DA DISSONÂNCIA MÚSICAIS

A consonância e a dissonância são sensações subjetivas associadas a dois (ou mais) sons soando simultaneamente, e de uma natureza bem menos definida do que as variáveis psicofísicas de altura e volume, e até mesmo timbre. Todavia, a música tonal de todas as culturas aparentemente indica que o sistema auditivo humano possui um senso para certos intervalos especiais de freqüência — a oitava, a quinta, a quarta etc. O mais significativo é que esses intervalos são “avaliados” quase na mesma ordem em que eles aparecem na série harmônica (veja Figura 2.19).

Quando dois sons complexos soam em uníssonos ou oitava perfeitamente afinados, todos os harmônicos de uma nota casarão com os da outra; a segunda nota não introduzirá freqüências “intermediárias”. Essa propriedade coloca a oitava numa situação muito especial como intervalo musical (além dos argumentos relativos à simplicidade do padrão de vibração apresentados na análise das superposições de sons puros, na Seção 2.6). Essa situação muda quando produzimos uma quinta justa formada com sons complexos (Tabela 5.1).

Tabela 5.1

Tônica	Quinta justa
f_1	$f'_1 = 3/2 f_1$
$2f_1$	
$3f_1$ -----	$f'_2 = 3 f_1$
$4f_1$ -----	
$5f_1$ -----	$f'_3 = 9/2 f_1$
$6f_1$ -----	$f'_4 = 6 f_1$

Todos os harmônicos ímpares da quinta têm frequências situadas entre os harmônicos da tônica; apenas os harmônicos pares coincidem. Em particular, o terceiro harmônico da quinta, com frequência $9/2f_1$ encontra-se “perigosamente próximo” às frequências do quarto e quinto harmônicos da tônica: as suas regiões de ressonância na membrana basilar podem se superpor, e portanto surgirão batimentos, ou “asperezas” (Seção 2.4), mesmo que o intervalo de frequências fundamentais esteja perfeitamente afinado. Construindo tabelas semelhantes à Tabela 5.1, o leitor poderá verificar que, para outros intervalos como quartas, terças e sextas, a proporção de harmônicos que “colidem” aumenta rapidamente e vai descendo na ordem dos harmônicos. Historicamente, acreditava-se que esse efeito fosse a causa das sensações de consonância e dissonância.

Na verdade, desde os tempos de Helmholtz a dissonância era associada ao número, intensidade e frequência dos harmônicos em batimento — e a consonância à ausência deles. Em outras palavras, assumia-se que, por algum motivo não definido, o nosso sistema auditivo “não gosta de batimentos”. Conseqüentemente, ele prefere o unísono perfeito e a oitava perfeita, porque nesses intervalos todos os harmônicos da nota superior casam com os da tônica. Na quinta, de acordo com a Tabela 5.1, o terceiro harmônico da nota superior pode entrar em batimentos com o quarto e quinto harmônicos da tônica. Essa proporção crescente de pares de harmônicos em batimento que aparecem à medida que vamos analisando os

intervalos de quarta, terça, sexta, sétima e segunda explicaria, assim, a decrescente consonância — ou crescente dissonância — desses intervalos. Essa suposição era particularmente interessante porque — como se pode demonstrar matematicamente com facilidade —, para maximizar o número de harmônicos que casam em dois sons complexos (minimizando assim os que não casam), é necessário que as suas frequências fundamentais f_1 e f'_1 tenham entre si uma razão de números inteiros, e que esses números sejam os menores possíveis. Na verdade, se

$$\frac{f'_1}{f_1} = \frac{n}{m}, \quad \text{com } n \text{ e } m \text{ inteiros} \tag{5.1}$$

então, o n -ésimo harmônico de f'_1 terá a mesma frequência que o n -ésimo harmônico de f_1 : $mf'_1 = nf_1$ (e assim acontecerá com o $(2 \times n)$ -ésimo e o $(2 \times n)$ -ésimo etc.). Nenhum dos outros harmônicos casará, e, portanto, dará origem a batimentos se as suas frequências estiverem muito próximas umas das outras. A Tabela 5.2 mostra todos os intervalos dentro de uma oitava que podem ser formados com números pequenos m e n e que são aceitos na cultura ocidental como consonâncias (em ordem decrescente de “perfeição”):

Tabela 5.2

Razão de Frequência (n/m)	Intervalo
Consonâncias “perfeitas”	1/1 unísono
	2/1 oitava
	3/2 quinta
	4/3 quarta
Consonâncias “imperfeitas”	5/3 sexta maior
	5/4 terça maior
	6/5 terça menor
	8/5 sexta menor

Com base em experimentos monaurais e binaurais sofisticados (Plomp e Levelt, 1965) sobre os julgamentos de consonâncias envolvendo pares de sons puros e sons complexos não-harmônicos, tornou-se evidente que os batimentos entre os harmônicos não podem ser o principal fator determinante na percepção da consonância. Dois sons puros a um intervalo de uma oitava ou menos foram apresentados a vários indivíduos musicalmente leigos (não-treinados), aos quais foi solicitado que classificassem a superposição quanto à sua “consonância” ou “agradabilidade”. Obteve-se um padrão contínuo, que não revelou preferência por nenhum intervalo musical em particular. Um exemplo é visto na Figura 5.3. Sempre que a distância entre os sons puros era menor que uma terça menor, aproximadamente, os intervalos eram classificados como “dissonantes” (com exceção do uníssono); os intervalos iguais ou maiores que uma terça menor foram classificados como mais ou menos consonantes, não importando a sua razão real de frequências¹. A forma da curva realmente depende da frequência absoluta da nota fixa. Tudo isso está relacionado com a sensação de “aspereza” dos uníssonos desafinados e à banda crítica, discutidas na Seção 2.4. Os resultados desses experimentos podem ser resumidos da seguinte maneira:

1. Quando as frequências de dois sons puros se encontram fora da banda crítica, o intervalo correspondente é classificado como “consonante”.
2. Quando elas coincidem, ele é julgado como “perfeitamente” consonante.
3. Quando as frequências diferem de uma quantidade que fica entre 5% e 50% da largura de banda crítica correspondente, eles são classificados como “não-consonantes”. Chamaremos de dissonância básica um intervalo entre dois sons puros nesta última condição.

Vamos agora voltar ao caso musicalmente mais significativo de dois sons complexos soando simultaneamente e aplicar os resultados acima a cada par de harmônicos vizinhos. Se o número total de pares que são mais

1. Os músicos treinados foram excluídos desse experimento porque eles teriam uma forte tendência a identificar consonâncias com base no treinamento.

ou menos consoantes (ver (1) acima) e os que são perfeitamente consoantes (2) for ponderado com os basicamente dissonantes (3), pode-se obter um “índice de consonância” para cada intervalo de sons complexos (Plomp e Levelt, 1965; Kameoka e Kuriyagawa, 1969). Pode-se mostrar que esse índice obtém valores de pico para sons cujas frequências fundamentais satisfazem a condição (5.1): a altura dos picos (“grau” de consonância) segue aproximadamente a ordem decrescente dada na Tabela 5.2. Além disso, em vista da dependência que a largura de banda crítica tem da frequência (Figura 2.13), um determinado intervalo musical tem um grau de consonância que varia ao longo da gama de frequências. Em particular quando caminha para as baixas frequências, o intervalo se torna cada vez menos consonante — um fato muito conhecido na música polifônica, na qual, no registro grave, são usadas principalmente oitavas e, eventualmente, algumas quintas.

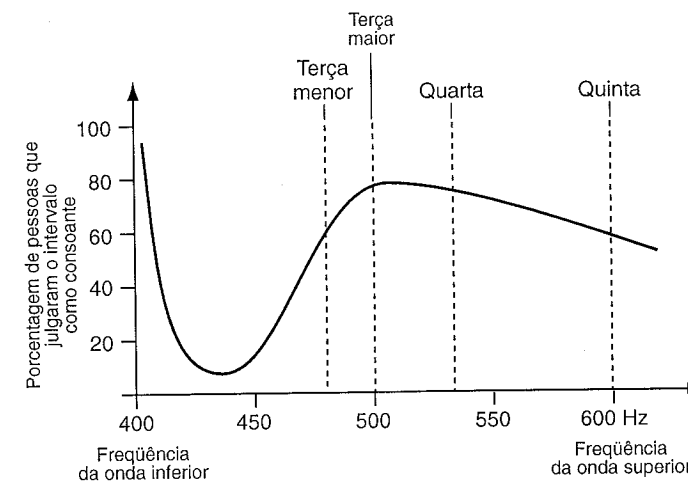


Figura 5.3

O grau de consonância também depende do espectro ou qualidade de cada som componente, *i.e.*, da intensidade relativa dos harmônicos dissonantes *versus* harmônicos consoantes. Isso também é bastante conhecido em

música: existem combinações de instrumentos que se “fundem” melhor do que outras na música polifônica. Mesmo a ordem em que dois instrumentos definem um intervalo é relevante. Por exemplo, se um clarinete e um violino tocam uma terça maior, com o clarinete tocando a nota mais grave, o primeiro par dissonante de harmônicos será o 7^o harmônico do clarinete com o 6^o harmônico do violino (porque só estão presentes os harmônicos ímpares do clarinete, Seções 4.4, 4.5). Esse intervalo soa macio. Se, por outro lado, o clarinete estiver tocando a nota superior, o 3^o harmônico deste colidirá com o 4^o harmônico da nota do violino, e o intervalo soará “áspero”.

A teoria da percepção de consonância de Terhard (1974) postula que a música tonal está baseada essencialmente em mecanismos de reconhecimento de padrões que atuam no sistema auditivo (Seções 2.9 e 4.8). Um desses mecanismos — o processador central de alturas —, responsável pela extração de uma única sensação de altura da complexa distribuição de atividade provocada por um som musical, adquire o conhecimento das relações específicas que existem entre os máximos de ressonância e os focos de atividade neural resultantes, evocados pelos seis ou oito harmônicos mais baixos de um determinado som musical (Seção 4.8). Assim, os intervalos de altura primária correspondentes (oitava, quinta, quarta, terça maior e terça menor) tornam-se “familiares” ao processador central do sistema auditivo e conferem significados tonais a todos os estímulos externos cujas frequências (fundamentais) mantêm essas relações (Apêndice II)². De acordo com

2. De modo geral, se for comprovadamente correta, a hipótese de que o processador central de alturas é uma unidade neural que precisa *aprender* a extrair informação significativa dos complexos sinais de entrada por meio de uma repetitiva exposição aos sons naturais (Terhardt, 1972; Seção 2.9) poderá ter um impacto de amplas dimensões. Na música, por exemplo, alguém pode se propor a tentar reaprender todo um novo conjunto de características “invariantes” pertencentes a uma certa classe de sons não-harmônicos — com a possibilidade de construir a partir daí escalas e esquemas tonais inteiramente novos (Terhardt, 1974). De um lado mais prático, essa capacidade intrínseca de aprendizagem traria novas esperanças aos esforços atuais de desenvolvimento de próteses eletrônicas para os surdos baseadas na implantação de microeletrodos no nervo acústico. Embora o padrão de ativação espacial desses implantes seja extremamente difícil de se determinar, a interpretação dos padrões de excitação provocados pode ser perfeitamente aprendida pelo processador central do paciente.

essa teoria, tanto a aspereza mínima quanto o significado tonal desempenham papel determinante na sensação de consonância subjetiva. Contudo, em vista do fenômeno da mudança de altura primária dos componentes harmônicos percebidos individualmente (p. 216), ambos os princípios podem fornecer instruções conflitantes e, em situações musicais reais, podem forçar o processador central a buscar um meio termo (Terhardt, 1974). A afinação “distendida” que se prefere para os pianos (em comparação com a escala de temperamento equivalente, Seção 5.3) e o fato observado de que a nota superior de um intervalo melódico é, por tendência, entoada mais agudo (Seção 5.4) podem ser o resultado desse “meio-termo”.

Existem fatores mais complicados que influenciam a sensação de consonância, mais destacadamente a experiência e treinamento e o preconceito que daí resulta (*i.e.*, a tradição musical). É interessante observar que, historicamente, os intervalos musicais como ingredientes harmônicos explícitos têm sido gradualmente “aceitos” na civilização ocidental, numa ordem próxima à da Tabela 5.2.

Isso parece indicar uma tolerância gradual da nossa habilidade de processamento auditivo, que agora está receptiva a “digerir” uma entrada muito mais complicada (como são os padrões de vibração das dissonâncias) do que há mil anos atrás, quando, provavelmente, uma quinta era considerada uma ousada inovação. Naturalmente, isso não foi resultado de evolução biológica, mas, antes, de uma sofisticação das experiências de aprendizagem a que os seres humanos estiveram expostos com o passar do tempo. Esse desenvolvimento, como o da civilização como um todo, foi gradual, em saltos “quânticos”: sempre foi necessária a mente de um gênio revolucionário para introduzir inovações ousadas cuja compreensão requeria novas e mais complexas operações de processamento de informações no cérebro, e a grandeza do gênio foi necessária para persuadir a humanidade a aprender, e assim aceitar, essas inovações ousadas.

Até agora estivemos considerando intervalos musicais menores ou iguais a uma oitava. Para os intervalos maiores (p. ex., dó₃ – sol₅), costuma-se descer a nota superior em oitavas (sol₅, sol₄, sol₃), até que se obtenha um intervalo menor que uma oitava (dó₃ – sol₃). Então, o grau de consonância desse intervalo é considerado “equivalente” ao do inicial. Essa

propriedade cíclica dos intervalos e que se repete em cada oitava recebe o nome de *croma* das notas musicais. É uma propriedade básica que atribui uma qualificação equivalente para todas as notas cuja altura difira em uma ou mais oitavas. Quem é responsável por esse curioso caráter cíclico das notas musicais, que se repete a cada oitava (sempre que a frequência dobra)? Não há equivalente em nenhuma das outras modalidades sensoriais. Isso obviamente está relacionado à propriedade básica da poderosa oitava: a de ter todos os seus harmônicos coincidentes com os harmônicos da nota mais grave. Não existe outro intervalo musical com essa propriedade (exceto o uníssono, naturalmente). De modo genérico, a existência da *croma*, *i.e.*, o fato de que alturas que diferem de uma oitava têm um grau de “semelhança” que é considerado idêntico ao do uníssono, indica que o processo de reconhecimento de padrões em nosso sistema auditivo deve responder de uma forma “especial” e talvez simplificada quando aparecem oitavas. Lembremos mais uma vez que a oitava é o primeiro intervalo da série harmônica, e que a frequência de repetição associada é idêntica à da nota mais grave. Qualquer outro intervalo musical consonante (quinta, quarta etc.) tem uma frequência de repetição fundamental associada (relações 2.7) que não está presente no estímulo original de dois sons. Se lembrarmos como opera o processador de alturas (Seção 2.9 e Apêndice II), veremos que, sempre que o processador deparar com dois sons complexos cujas frequências fundamentais f_1 e f_2 formem um intervalo musical, a sua saída deverá conter dois sinais prominentes representando a frequência de cada som (correspondendo a f_1 e f_2), mais outros sinais menos evidentes representando a frequência de repetição (2.7) correspondente ao par de primeiros harmônicos f_1 e f_2 e seus múltiplos (Apêndice II). Sob condições normais, esses sinais adicionais são descartados como sensação de altura, um processo que requer uma operação adicional de “filtragem”. Observe, porém, que essa operação adicional não será necessária quando se tratar de uma oitava, porque o terceiro sinal de saída não estará presente (Figura AII.2, Apêndice II)! Na verdade, o “significado tonal” mencionado acima pode estar fortemente relacionado ao número, intensidade e posição de sinais “parasitas” na saída do processador de alturas. Quanto mais complexa for a multiplicidade desses sinais

(*i.e.*, quanto mais complexo for o padrão de vibração sonora), “menor” será o significado tonal da superposição de sons original.

Finalmente, quando três ou mais notas soam simultaneamente, costuma-se analisar o acorde resultante em pares de notas, considerando os seus “valores” individuais de consonância. É óbvio que, quanto mais e mais notas forem acrescentadas, surgirá uma configuração mais complicada de regiões de ressonância na membrana basilar. A superposição aumentará até o limite em que as células sensoriais que cobrem grande parte da membrana basilar responderão todas ao mesmo tempo. À luz das várias teorias de altura (Seções 2.9, 4.8 e Apêndice II), podemos afirmar que, também nesse caso, o grau de consonância (ou dissonância) pode estar relacionado tanto à proporção de harmônicos em batimento quanto ao número, intensidade e posição dos sinais “parasitas” na saída do processador de alturas. Observe que a tríade maior é uma combinação de três notas cujos componentes, tomados dois a dois, sempre fornecem frequências de repetição que diferem da tônica apenas em oitavas (*i.e.*, têm a mesma *croma*).

Contudo, existe um limite para a inteligibilidade de muitos sons. Quando os padrões de vibração se tornam aleatórios (*i.e.*, a sua periodicidade é destruída), ou quando a sua complexidade excede um certo limiar, o mecanismo de processamento neural simplesmente pára: não é possível estabelecer nenhuma sensação de timbre ou altura definida. A sensação resultante é chamada de ruído. Qualquer oscilação de pressão não-periódica leva a uma sensação de ruído. Porém, o ruído pode ser altamente organizado. Assim como uma oscilação periódica pode ser analisada como uma superposição discreta de oscilações harmônicas puras com frequências que são múltiplos inteiros da frequência fundamental (Seção 4.3), as vibrações não-periódicas podem ser analisadas como uma superposição contínua de vibrações puras com todas as frequências possíveis. Dependendo de como a intensidade está distribuída ao longo de todas essas frequências, obteremos diferentes espectros de ruído. O ruído desempenha um papel fundamental na formação das consoantes na fala, mas também tem seu papel na música: a importância dos componentes de ruído nos instrumentos de percussão é óbvia. Já se demonstrou que a concentração de ruído detectado durante os primeiros décimos de segundo em sons de piano e cravo é um elemento fundamental para o pro-

cesso de reconhecimento. O efeito dos ruídos de espectros eletronicamente controlados sobre nossa percepção auditiva está sendo estudado amplamente com sintetizadores de notas e de ruídos. Um vasto e novo território das sensações auditivas (música??) está se revelando (veja também a Seção 5.7).

5.3 CONSTRUINDO ESCALAS MUSICAIS

Concluimos este capítulo com uma discussão mais ou menos convencional sobre escalas musicais. Vamos aceitar o fato de que as escalas existem e, por uma finalidade puramente prática, vamos definir uma escala simplesmente como um conjunto discreto de alturas arranjadas de tal modo que forneçam o máximo número possível de combinações consonantes (ou o mínimo número possível de dissonâncias) quando duas ou mais notas do conjunto soam simultaneamente. Com essa definição, e mediante os dados da Tabela 5.2, é possível gerar, a princípio, duas escalas, de modo praticamente inequívoco, dependendo se todos os intervalos consonantes devem ser levados em conta ou se apenas serão consideradas as consonâncias “perfeitas”. No primeiro caso, temos a escala justa; no segundo, temos a escala pitagórica³.

I. A Escala Justa

Começamos com uma nota de frequência f_1 , que chamaremos de $dó^4$. A primeira coisa mais óbvia a ser feita é introduzir a oitava acima, que chamaremos de $dó'$. Isso nos fornece o intervalo mais consonante de todos. A segunda coisa óbvia é acrescentar a quinta de frequência $3/2f_1$, que chamaremos de sol. Isso fornece dois novos intervalos consonantes

3. Consideramos aqui uma escala como um conjunto de sons com relações de frequência matematicamente definidas. Devemos distinguir isso dos vários modos de escala, definidos pela ordem particular em que tons e semitons se sucedem.
4. A notação do solfêjo, $dó-ré-mi-fá-sol-lá-si-dó$, é usada aqui para indicar a posição relativa numa escala (*i.e.*, a *croma*), e não a altura real.

além da oitava, com razões de frequência $3/2$ ($dó-sol$) e $4/3$ (a quarta $sol-dó'$). Como próximo passo, existem duas opções, se quisermos manter um número máximo de intervalos consonantes. Essas opções são as notas $5/4f_1$ e $6/5f_1$, que chamaremos de mi e mi-bemol, respectivamente. Escolheremos a primeira, mi, porque ela nos garantirá um número de consonâncias de mais alto “grau”. A Figura 5.4 mostra os intervalos resultantes, todos consonantes. As notas $dó-mi-sol$ constituem a tríade maior, o elemento básico da harmonia da música ocidental (a nossa segunda escolha, $6/5f_1$ ou mi-bemol, teria fornecido tríade menor).

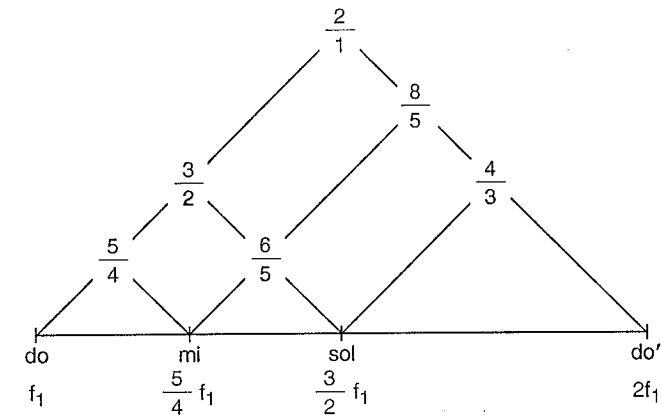


Figura 5.4

Podemos continuar a introduzir notas, tentando, a cada passo, manter mínimo o número de dissonâncias e máximo o de consonâncias (Tabela 5.2). Acabamos encontrando a escala diatônica justa de sete notas dentro da oitava (Figura 5.5). Essas sete notas podem ser projetadas em oitavas para cima e para baixo, de modo a formar uma escala diatônica completa sobre toda a extensão de alturas audíveis. Observe, na Figura 5.5, os dois intervalos com razões de frequência semelhantes $9/8$ e $10/9$, que representam tons inteiros. O intervalo $16/15$ define um semitom. Com as notas

dessa escala, tomadas aos pares, podemos formar 16 intervalos consonantes (Tabela 5.4), 10 intervalos dissonantes (sétimas maiores e menores, quinta diminuta, tons inteiros, semitons) e — infelizmente — duas consonâncias “desafinadas”: a terça menor 1,5% mais alta ré-fá ($32/27$) e a quinta 1,9% mais baixa ré-lá ($40/27$). Finalmente, e muito importante, com a escala diatônica justa podemos formar quatro tríades maiores justas: dó-mi-sol, mi-sol-si, fá-lá-dó, e sol-si-ré; duas tríades justas menores: mi-sol-si e lá-dó-mi e uma tríade menor “desafinada”: ré-fá-lá.

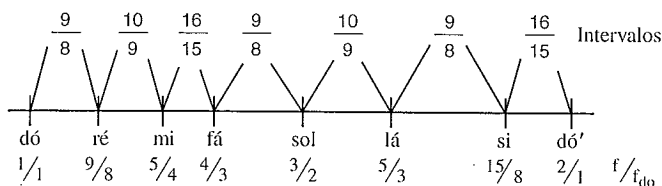


Figura 5.5

Considerando a existência de espaçamento desigual entre notas vizinhas, é possível ainda aperfeiçoar essa escala dividindo os espaços maiores (tons inteiros) em dois semitons cada um. Infelizmente, os novos intervalos resultantes vão ficando cada vez mais complexos (p. ex., vários tipos de semitons, consonâncias desafinadas), as opções não são únicas, e encontramos valores diferentes de frequência para os chamados equivalentes enarmônicos dó#-réb, ré#-mi b etc. Procurando conservar máxima a proporção de consonâncias possíveis, são introduzidas as seguintes notas: mi b ($6/5f_1$), si b ($9/5f_1$), sol# ($25/16f_1$) (ou lá b, ($8/5f_1$), dó# ($25/24f_1$) e fá# ($45/32f_1$). O resultado é uma cromática justa, com 12 notas dentro da oitava.

2. A Escala Pitagórica

Vamos agora nos restringir às chamadas consonâncias perfeitas, a quinta e a quarta justas (e a oitava, naturalmente), e construir nossa escala apenas com base nesses intervalos. Podemos proceder da seguinte maneira. Depois

de introduzir sol nós descemos uma quinta justa a partir de dó' (e introduzimos fá ($2/3 \times 2f_1 = 4/3f_1$). Descemos agora uma quarta justa a partir de sol e obtemos ré ($3/4 \times 3/2f_1 = 9/8f_1$), e uma quinta acima desta última, obtendo lá ($3/2 \times 9/8f_1 = 27/16f_1$). Finalmente, preenchemos os espaços restantes descendo uma quarta de lá para obter mi ($3/4 \times 27/16f_1 = 81/64f_1$), e uma quinta a partir daí, até si ($3/2 \times 81/64f_1 = 243/128f_1$). O resultado é a chamada escala pitagórica (Figura 5.6). Observe que existe apenas um intervalo de tom inteiro, o tom inteiro pitagórico, com razão de frequência $9/8$ (igual ao tom inteiro “pequeno” da escala justa). O intervalo $265/243$ é o semitom pitagórico diatônico.

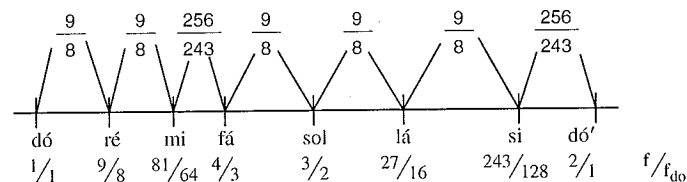


Figura 5.6

Podemos transformar essa escala numa escala cromática se continuarmos a saltar quartas ou quintas para cima e para baixo. Obteremos então fá# (uma quarta abaixo de si), dó# (uma quarta abaixo de fá#), sol# (uma quinta acima de dó#), si b (uma quarta acima de fá) e mi b (uma quinta abaixo de si b). Dessa forma surgirá um novo semitom (p. ex., fá-fá#), definido pela singular razão de $2187/2048$, chamado semitom pitagórico cromático. Todo este procedimento nos leva também aos equivalentes enarmônicos de frequências diferentes. Em particular, se continuarmos a subir ou descer em saltos de quartas e quintas justas, eventualmente retomaremos à nossa nota inicial dó — mas não exatamente). Em outras palavras, chegaremos ao equivalente enarmônico “si#”, cuja frequência não é igual à de dó' ($= 2f_1$).

Assim, com base em alguns princípios “lógicos”, geramos duas escalas. Cada uma tem a sua própria coleção de problemas. O mais sério, sem dúvida, é o fato, comum a ambas, de que apenas um grupo muito limitado de tonalidades pode ser tocado com essas escalas sem que se tenha problemas com consonâncias desafinadas. Em outras palavras, ambas as escalas impõem restrições muito sérias de transposição e modulação. Já no século XVII, esse problema havia sido reconhecido. Não há dúvida, entretanto, de que ambas as escalas revelam um caráter bastante específico quando se executa música em instrumentos afinados em uma delas. Mas o tipo de música que pode ser tocado é extremamente limitado.

3. Escala Temperada

Assim, tornou-se evidente que era necessário um novo tipo de escala que, com base num compromisso razoável, abrindo mão de um pouco de “justeza” dos intervalos, levasse a intervalos igualmente espaçados, a despeito da tonalidade em particular. Em outras palavras, um semitom teria sempre a mesma razão de frequência, fosse ele dó-dó#, mi-fá, ou lá-sib, e uma quinta seria sempre a mesma, fosse ela fá-dó' ou dó#-sol#. Isso foi conseguido com a escala temperada, entusiasticamente patrocinada por nada menos que J. S. Bach, que compôs uma série de Prelúdios e Fugas (“Das wohltemperierte Clavier”*) com o propósito específico de aproveitar ao máximo as novas regiões descobertas pelas possibilidades irrestritas de mudança de tonalidade.

Na escala temperada, a razão de frequência é a mesma para todos os 12 semitons entre dó e dó'. Vamos chamar essa razão de s . Isso significa que

$$f_{do} = s f_{do\#}, f_{re} = s f_{do\#} = s^2 f_d, \dots, f_{do'} = s^{12} f_{do}$$

Como nós sabemos que $f_{do'} = 2 f_{do}$ (a oitava é o único intervalo “justo”!), a décima segunda potência de s deve ser igual a 2, Ou

$$s = \sqrt[12]{2} = 1,0595 \tag{5.2}$$

* “O cravo bem-temperado” (N. da T.).

Esta é a razão de frequência de um semitom temperado. As frequências resultantes para as notas da escala cromática temperada são potências inteiras de s vezes f_{do} . A Tabela 5.3 mostra as razões de frequências dos intervalos consonantes nas três escalas.

Tabela 5.3

Intervalo	Escala justa		Escala pitagórica		Escala temperada	
	razão	cents	razão	cents	razão	cents
Oitava	2000	1200	2000	1200	2000	1200
Quinta	1500	702	1500	702	1498	700
Quarta	1333	498	1333	498	1335	500
Terça maior	1250	386	1265	408	1260	400
Terça menor	1200	316	1184	294	1189	300
Sexta maior	1667	884	1687	906	1682	900
Sexta menor	1600	814	1580	792	1587	800

É conveniente introduzir uma subdivisão-padrão do intervalo básico da escala temperada, para que se possa expressar numericamente as pequenas diferenças entre os intervalos pertencentes a diferentes escalas. Essa subdivisão, embora não tendo aplicação por si na música, é normalmente usada para descrever pequenas variações de frequência (*vibrato*), variações de entonação (altura) e o grau de desafinação de notas e intervalos. Muitas formas de subdividir a escala de frequências foram propostas e usadas no passado. O procedimento mais aceito hoje é dividir o semitom temperado em cem intervalos iguais ou, o que é equivalente, dividir a oitava em 1 200 partes iguais. Como o que define um intervalo musical é a razão entre as frequências fundamentais das notas componentes (e não a sua diferença), o que é necessário fazer aqui é dividir a razão de frequência do semitom s (5.2) em cem fatores c iguais:

$$\underbrace{c \times c \times c \times \dots \times c}_{100 \text{ vezes}} = c^{100} = s$$

Pela relação (5.2), o valor de c é

$$c = \sqrt[100]{1,0595} = 1,000578 \quad (5.3)$$

A unidade dessa subdivisão é chamada *cent*. Para sabermos quantos *cents* estão “contidos” num certo intervalo com uma determinada razão de frequência r , precisamos determinar quantas vezes c deve ser multiplicado por ele mesmo para obtermos r :

$$c^n = r \quad (5.4)$$

n será então o “valor” de r expresso em *cents*. Por definição, um semitom temperado tem 100 *cents*, um tom inteiro temperado (s^2) tem 200 *cents*, uma quinta temperada (s^7) tem 700 *cents* etc. Para descobrirmos o valor em *cents* de qualquer outro intervalo precisamos usar logaritmos. Levando em conta as propriedades descritas na Seção 3.4, tomamos o logaritmo da relação (5.4): $n \log c = \log r$. Então:

$$n = \frac{\log r}{\log c} = 3986 \log r \quad (5.5)$$

Usando essa relação, encontramos os valores em *cents* dos vários intervalos consonantes dados na Tabela 5.3.

5.4 A ESCALA-PADRÃO E O PADRÃO DE ALTURA

A escala temperada tem sido usada há mais de 200 anos e transformou-se efetivamente na escala-padrão na qual são afinados todos os ins-

trumentos de notas fixas. Desde o seu advento, todavia, ela tem sido atacada em várias ocasiões — até hoje. O alvo principal desses ataques é a “injusteza” dos intervalos consonantes da escala temperada, em particular as terças e sextas (Tabela 5.3), que de fato soam um pouco desafinados quando ouvidos cuidadosa e persistentemente, em especial no registro grave.

Vamos comparar criticamente as escalas discutidas na seção anterior. Não há dúvida de que, para uma determinada tonalidade, a escala justa é a escala “teoricamente” perfeita, fornecendo uma possibilidade máxima de combinações de intervalos justos ou “puros” (*i.e.*, sem batimentos). Por essa razão, a escala justa deveria ser tomada como uma espécie de escala de referência; eis exatamente por que a apresentamos em primeiro lugar. Mas a grande questão é: o nosso sistema auditivo realmente se importa com intervalos justos, sem batimentos? E então: será que abriríamos mão das possibilidades de transposição e modulação de tonalidades a fim de obtermos esses intervalos “puros”? Uma história da música com 200 anos já respondeu, sem titubear, a essas questões, com um sonoro e enfático NÃO! Assim, a escala justa está descartada.

A escala pitagórica pode estar um passo à frente na direção certa (enquanto as quintas e quartas são conservadas como intervalos justos, as terças e sextas são ligeiramente desafinadas, Tabela 5.3), mas ela ainda não permite possibilidades limitadas de modulação e transposição. Existem outras escalas, apresentadas como pequenas modificações da escala pitagórica, e que todavia nem vamos mencionar aqui. Nenhuma delas teve sucesso em evitar que a escala temperada fosse aceita universalmente.

Houve tentativas de se estabelecer experimentalmente a questão de qual escala é realmente preferida (deixando de lado os argumentos de capacidade de modulação). Existem duas abordagens possíveis:

1. Usar instrumentos de frequências fixas (piano, órgão) e comparar cuidadosamente as impressões subjetivas de uma determinada peça executada sucessivamente em dois instrumentos do mesmo tipo, afinados em escalas diferentes. A peça, obviamente, deve ser bem simples, sem modulações para tonalidades distantes. E o instrumento não

deve apresentar decaimento nas notas (como o órgão), para que se perceba mais claramente os batimentos ou aspereza.

2. Medir experimentalmente as frequências médias de entonação escolhidas por cantores ou executantes de instrumentos com alturas variáveis (cordas) e determinar se eles preferem uma escala ou outra.

A segunda abordagem é mais apropriada para fornecer resultados quantitativos. A instrumentação eletrônica tornou possível medições de frequência instantâneas muito precisas diretamente nos artistas. Os intervalos musicais observados mais atentamente são a terça maior e a sexta maior, para os quais as diferenças entre as escalas são mais pronunciadas (Tabela 5.5). Observe, em particular, que a nota superior em ambos os intervalos é baixa na escala justa e alta na escala pitagórica (em relação à escala temperada). Os instrumentos experimentais mostram, muito convincentemente, que, em média, os cantores e os instrumentistas de cordas executam as notas superiores desses intervalos numa entonação alta (Ward, 1970). Isso parece indicar uma preferência pela escala pitagórica. Todavia, não devemos tirar conclusões apressadas: os mesmos experimentos revelaram que também as quintas e quartas e até a poderosa oitava eram executadas mais alto, em média!⁵ Em vez de revelarem uma preferência por uma certa escala (a pitagórica), esses experimentos indicam a existência de uma inesperada tendência universal de tocar ou cantar mais alto todos os intervalos musicais. Essa entonação “distendida” pode ser causada pelo desvio na altura primária dos componentes harmônicos de um dado som musical (Seções 4.3 e 5.2), o que deixa um registro “ligeiramente errado” no processador central de alturas (para uma discussão detalhada, veja Hartmann, 1993). Além disso, um resultado talvez mais significativo desses experimentos é que as flutuações individuais da altura de uma determinada nota durante a execução são muito grandes. Isso inclui o *vibrato*, bem como as variações da altura média de uma certa nota quando ela aparece várias vezes numa mesma peça. Nessas varia-

5. Existe um efeito recíproco: intervalos melódicos justos são consistentemente considerados baixos (Terhardt e Zick, 1975).

ções de altura de uma certa nota escrita, varre-se uma faixa de frequências que vai bem além das diferenças de frequência entre escalas diferentes — na verdade, isso faz com que estas últimas se tornem totalmente irrelevantes! De modo geral, todos esses resultados apontam para o fato de que os intervalos musicais são percebidos de um modo categórico, sendo as flutuações reais facilmente ignoradas pelo mecanismo de processamento primário.

Até agora estivemos lidando com intervalos, *i.e.*, razões de frequência. E quanto às próprias frequências absolutas? Desde que uma escala tenha sido adotada, é suficiente definir a frequência de apenas uma nota; pode ser qualquer uma, não faz diferença. Todavia, se se deseja que os instrumentos musicais de frequências fixas tenham trânsito fácil por todo o mundo, isso precisa ser feito com base num acordo internacional. Atribuiu-se ao “lá central” do piano ($l\acute{a}_4$) uma frequência fundamental de 440 Hz. Diferentes frequências-padrão “regionais” têm sido usadas desde o início do emprego dos diapasões no século XVII. Mas, nos dois últimos séculos, tem havido um aumento gradual da frequência padrão de 415 Hz até 461 Hz⁶. Esperamos agora que o padrão atual se mantenha efetivamente constante.

Na escala temperada, todos os intervalos da mesma espécie (p. ex., todas as quintas, todas as terças maiores etc.) são exatamente “a mesma coisa”, exceto pela altura real de seus componentes. Não existe nenhuma diferença entre uma melodia tocada em dó maior e a mesma melodia tocada em ré maior (exceto pela faixa de frequências coberta). As “cores” absolutas e os diferentes “climas” das tonalidades não têm fundamento psicoacústico, como os experimentos já demonstraram (Corso, 1957). Porém, pode haver pequenas diferenças nos sons de várias tonalidades devido a circunstâncias físicas: a maior ocorrência de teclas pre-

6. Isso tem uma conseqüência séria para instrumentos históricos famosos e que ainda estão em uso hoje em dia. Por exemplo, um violino Stradivarius, construído originalmente para uma altura-padrão de, digamos, $l\acute{a}_4 = 415$ Hz, atualmente tem que ser afinado mais alto, o que significa uma tensão maior nas cordas (relação 4.3). Isso altera a qualidade (espectro) do som. Um órgão barroco, também construído para $l\acute{a}_4 = 415$ Hz, precisará ter seus tubos parcialmente cortados para encurtar o seu comprimento efetivo (relação 4.6).

tas (que são percutidas de uma forma ligeiramente diferente) no piano, para certas tonalidades⁷, a maior ocorrência de cordas soltas em certas tonalidades nos instrumentos de corda, ou o efeito da faixa de frequência fixa dos formantes (Seção 4.4) em caixas de ressonância e outros corpos ressonantes.

Uma palavra final é oportuna quanto à percepção absoluta de alturas*. As poucas pessoas dotadas com a habilidade de identificar ou reproduzir vocalmente uma determinada nota de maneira absoluta (também chamada de “altura perfeita”) são normalmente muito admiradas. Acontece que fomos treinados a prestar atenção e armazenar em nossa memória de longo prazo apenas os intervalos de altura relativa, porque esta é a informação mais relevante numa “mensagem musical”. Em outras palavras, nosso cérebro foi treinado para interpretar e armazenar uma melodia como uma seqüência de transições de altura, em vez de “valores” de altura; a informação sobre a altura absoluta, embora chegando ao nosso cérebro, é descartada como não-essencial no processo cognitivo. Todavia, ela pode ser retida por qualquer pessoa normal durante pequenos intervalos de tempo, que variam de 10 segundos até alguns minutos (Rakowski, 1972). É bem possível que as “alturas absolutas” possam ser aprendidas num estágio precoce da educação mental e então retiradas, assim como se “aprende” as cores absolutas ou o reconhecimento de cheiros. Essas considerações e as surpreendentes experiências sobre alturas detectadas eletronicamente na música ao vivo, e que parecem passar despercebidas pelo ouvinte, sugerem a necessidade de uma redefinição operacional de altura como o correlato subjetivo de cada um dos eventos acústicos numa seqüência de notas musicalmente significativa (Houtsma e Goldstein, 1972). Não deve ser surpresa para ninguém que essa definição tem sido sempre usada — implicitamente ou explicitamente — por todos os músicos!

7. Por exemplo, a música para piano de Chopin!

* Também conhecida por “ouvido absoluto” (N. da T.).

5.5 POR QUE EXISTEM ESCALAS MUSICAIS?

Nosso ouvido é sensível a ondas sonoras dentro de uma ampla faixa de frequências. Detectamos variações de frequência muito pequenas; a diferença no limite do perceptível é de apenas 0,5% ou menos (Figura 2.9). Porém, a nossa música ocidental (e a de muitas outras culturas) está baseada em escalas, *i.e.*, transições e superposições de sons que diferem uns dos outros em mais de 20 vezes o limiar de nossa capacidade de resolução de frequências. Por que não fazemos música com alturas continuamente variáveis, que se assemelham, por exemplo, às “canções” das baleias e dos golfinhos (que possuem um sistema de comunicação acústica altamente sofisticado, baseado em “deslizamentos” contínuos de frequência)? Por que a altura sempre tem que saltar em passos discretos?

Não existe uma resposta simples para esta pergunta. Em primeiro lugar, vamos lembrar que uma nota musical tem que durar um certo tempo mínimo para que possa ser totalmente processada pelo cérebro (Seção 3.4). Isso tem impedido que os sons “deslizantes” se tornem elementos básicos e definitivos na música. Em segundo lugar, observemos que as diferentes culturas musicais usam ou usaram diferentes escalas — assim, as escalas estão de alguma forma relacionadas ao treinamento e à tradição, ou foram influenciadas por eles. Em terceiro lugar, a maioria dos instrumentos musicais primitivos tinha alturas fixas.

A existência de escalas também foi justificada com base na consonância; isso implicaria que as escalas apareciam relacionadas à música polifônica. Todavia, as escalas já haviam sido usadas quando as melodias eram apenas cantadas (e/ou acompanhadas) monofonicamente em uníssono (ou, no máximo, em oitavas ou quintas).

Talvez a razão neuropsicológica subjacente para a existência das escalas seja porque é mais fácil para o cérebro processar, identificar e armazenar em sua memória uma melodia formada por uma seqüência temporal de valores discretos de altura que ostentem uma certa relação entre si, proporcionada, de certa forma, pela série harmônica, do que por um padrão temporal de alturas que deslize continuamente para cima e para baixo por todas as frequên-

cias possíveis, cujo processamento, identificação e armazenagem na memória exigiriam muito mais *bits* de informação do que uma seqüência discreta.

A explicação da existência de escalas, *i.e.*, seqüências discretas de notas, também foi tentada sobre uma base dinâmica de relações temporais nota-nota. Essa linha de raciocínio está baseada num campo musicalmente muito importante, mas psicofisicamente quase inexplorado, das sensações de “direção” ou expectativa de uma seqüência de duas (ou mais) notas, de “dominância” de uma dessas notas e de “retorno” a essa altura principal (também chamado “finalidade”). Por exemplo, temos a tendência de atribuir uma direção “natural” a uma seqüência de duas notas, que será ascendente (em altura) se a distância entre elas for de um semitom, e descendente se a distância for de um tom inteiro. Em ambos os casos, atribuímos uma dominância à segunda nota; o senso “natural” será equivalente à direção que se aproxima da dominante, e o senso antinatural à direção que se afasta dela. Da mesma forma, uma seqüência como dó-sol-dó-sol-dó-sol-... “quer” terminar em dó, enquanto que a seqüência dó-fá-dó-fá-dó-fá-... “implora” para terminar em fá. E se ouvirmos mi-sol-mi-sol-mi-sol-... nenhuma das duas notas é satisfatória como final, e nós desejamos ouvir dó! Toda a orientação diatônica ao se ouvir a música ocidental é baseada nesses efeitos, e assim também é o “significado” (a *Gestalt tonal*) das mensagens musicais (Seção 1.3).

No início deste século, Meyer (1900) e Lipps (1905) tentaram explicar a preferência por certos finais melódicos e dominâncias tônicas em termos das propriedades numéricas das razões de freqüências de um “melodia” de duas notas. Nos exemplos acima, a nota dominante é aquela cuja freqüência correspondente a uma potência de dois na razão de números inteiros. Por exemplo, $f_{sol}/f_{do} = 3/2$ (a nota grave predomina); $f_{fa}/f_{do} = 4/3$ e $f_{do}/f_{si} = 16/15$ (a nota aguda predomina). Entretanto, pesquisas posteriores passaram a atribuir esses efeitos principalmente ao condicionamento cultural. Porém, a questão permanece: por que surgiram essas, e não outras preferências? Quanto a isso, é interessante notar que, quando os intervalos musicais dó-sol, dó-fá e mi-sol dos exemplos acima são considerados como vizinhos de uma série harmônica, a nota fundamental dessa série é que acaba impondo a altura predominante, determinada

pelo senso de retorno (dó, fá e dó, respectivamente). Mais uma vez, essa expectativa pode ser ditada pela “familiaridade” com as relações harmônicas adquirida pelo nosso processador central de alturas (Seções 2.9, 4.8, 5.2 e Apêndice II) ou, num nível cognitivo mais elevado, pela familiaridade adquirida com a exposição à cultura musical do Ocidente (Bharucha, 1994).

Outro fenômeno perceptivo de importância para a música, relacionado a seqüências temporais de sons, é o da *fissão* da melodia. Se se tocar uma melodia na qual as notas se sucedam rapidamente em intervalos melódicos de vários semitons alternando-se para cima e para baixo, a coerência se perderá e serão percebidas duas linhas melódicas diferentes (ou mais). Nesse caso, nosso cérebro tende a agrupar as notas de acordo com a sua proximidade em altura, em vez de sua proximidade no tempo. Esse efeito foi usado abundantemente, especialmente durante o período barroco, para tornar possível a execução de uma música a várias vozes em um único instrumento que toque apenas uma nota de cada vez*. Para análises abrangentes desse e de outros fenômenos relativos a seqüências temporais, veja Van Noorden (1975), Deutsch (1982) e Sundberg (1992).

Podemos tentar prever que todos esses efeitos surgem provavelmente pela maneira de o sistema neural operar, segundo a “lei do mínimo esforço”: no processo de identificação das mensagens musicais (e de outras mensagens sensoriais), o sistema reserva um mínimo de informações, descartando todo o resto. Se a identificação for bem-sucedida, ele continua com a próxima mensagem. Caso contrário, ele volta e procura pistas adicionais. Isso se aplica não só à entrada de um único som, mas também à mensagem musical como um todo: o sistema nervoso procura usar quaisquer informações que estejam à mão, oriundas de experiências anteriores (p. ex., mensagens armazenadas por muito tempo na memória) para facilitar e até antecipar o processo de identificação da informação nova que chega. Essa estratégia de processamento neural é bem conhecida na percepção da fala e pode assinalar uma relação natural entre a música e a linguagem (veja a próxima seção).

* Em música, esse efeito é chamado de “polifonia implícita”. (N. da T.).

5.6. PROCESSOS CEREBRAIS COGNITIVOS E AFETIVOS NA PERCEÇÃO MUSICAL: POR QUE REAGIMOS EMOCIONALMENTE À MÚSICA?

A percepção musical, como a percepção da fala, envolve tarefas cognitivas complexas, em que a informação levada por sinais acústicos é analisada, armazenada, recuperada, intercomparada e interpretada — embora todo o propósito da música pareça estar bem mais relacionado aos estados afetivos que ela nos evoca do que ao conteúdo informativo básico das mensagens musicais que ela transmite. Talvez a demonstração mais convincente disso seja o fato de gostarmos de uma gravação musical em particular todas as vezes que a ouvimos, não importando o nosso grau de familiaridade com a obra.

Para compreender a percepção musical como processo cognitivo, é necessário analisar em termos gerais o papel da motivação e da emoção nas tarefas cognitivas. Esta ainda é uma área um pouco desprezada: o estudo do comportamento inteligente se direciona principalmente para o conteúdo, para a lógica e para a organização do processamento de informação pelo cérebro e a resposta comportamental resultante, com uma atenção menos explícita aos impulsos somáticos que motivam a aquisição da informação, seu processamento e armazenagem, e os estados afetivos geralmente descritos como emoção. Todavia, esta última é justamente o que diferencia a “função cerebral” da “computação neural” (Apêndice II): ao que se sabe, os computadores, por mais sofisticados que sejam, não estão preocupados com a sobrevivência (por enquanto)! Em todos os atos perceptivos conscientes é sempre alguma motivação que conduz a cognição; a cognição leva a algum estado afetivo, que por sua vez conduz, reforça ou modifica a motivação e a atenção perceptivas. Nos animais, essa relação circular é iniciada e controlada pela combinação de uma entrada ambiental de momento e uma necessidade somática corrente. Nos humanos, porém, ela também pode ser iniciada e controlada exclusivamente por uma atividade cerebral interna, sendo a informação sobre o ambiente e o organismo inteiramente resgatada da memória, em vez de adquirida no momento real (Seção 4.10).

As funções cognitivas do cérebro são controladas principalmente pelas redes corticais frontal e de associação, enquanto o impulso motivador e a resposta emocional são controlados pelo sistema límbico — mas essas estruturas cerebrais trabalham inextricavelmente juntas, como está fisiologicamente manifesto pelas interconexões neurais inatas entre elas. O sistema límbico é uma parte filogeneticamente antiga do cérebro, que compreende várias estruturas que se situam entre as áreas de associação cortical e o hipotálamo (hipocampo, amígdala, vários núcleos talâmicos). Em conjunção com o hipotálamo (a parte do cérebro que integra as funções do sistema nervoso autônomo e regula o sistema de informação neuroquímico), o sistema límbico “polícia” a entrada sensorial, dirige seletivamente a armazenagem na memória, de acordo com a relevância da informação, e mobiliza a saída motora (Figura 5.7) com o objetivo específico de assegurar uma resposta que seja a mais benéfica para a preservação do organismo e a perpetuação da espécie num ambiente complexo. Ao contrário das redes corticais que controlam o comportamento inteligente, boa parte do circuito neural do sistema límbico já está, no nascimento, “pré-conectada” a funções programadas durante o lento curso da evolução filogenética. A motivação e a emoção são manifestações integrais da função límbica, manifestações do princípio-guia do sistema límbico de assegurar que todos os processos corticais sejam realizados para o máximo benefício do organismo (p. ex., Panksepp, 1982, e as referências aí contidas). Nisso o sistema límbico trabalha maravilhosamente, dispensando as sensações de recompensa (ou punição) sempre que o organismo se engaja em atos comportamentais que, de acordo com comandos genéticos, são avaliados para melhorar (ou agir contra) as chances de sobrevivência do organismo e/ou espécie e para incrementar (ou agir contra) a diversificação.

O sistema límbico é ativado por entradas somáticas e ambientais (Figura 5.7) (Roederer, 1987). Nos humanos, porém, ele igualmente pode responder de modo exclusivo a imagens evocadas internamente e exibidas no córtex durante o processo de pensamento (Seção 4.10). Em outras palavras, a motivação e a emoção nos seres humanos podem ser provocadas sem nenhuma relação com o estado instantâneo do ambiente ou do organismo. Da mesma forma, os humanos podem deliberadamente invalidar as ordens do sistema límbico.

Vamos voltar à música. Parece que as mensagens musicais não transportam informações biologicamente relevantes, como no caso da fala, das elocuições animais e dos sons ambientais — embora as pessoas em todas as culturas reajam às mensagens musicais. O que, na evolução humana, pode ter levado a isso? Será a música uma vantagem da raça humana, específica da espécie? Obviamente, essa questão deve ser considerada parte de uma questão mais abrangente, relativa ao surgimento da motivação, resposta e criatividade estéticas. Em busca dessas respostas (Roederer, 1984), devemos primeiro nos voltar à habilidade individual mais característica dos seres humanos: a fala. A fala envolve o sistema auditivo e redes neurais centrais correlatas — não existe equivalente na visão ou em outro sentido. (A linguagem escrita é relativamente recente na história humana — algumas culturas até hoje não possuem linguagem escrita — e não requer o surgimento de centros corticais especializados nesse tipo de processamento de informação óptica.)

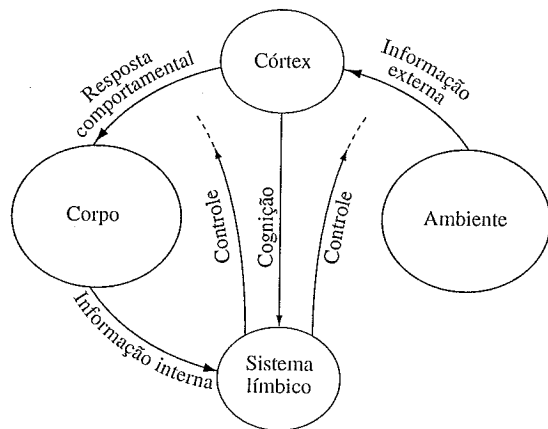


Figura 5.7 O papel do sistema límbico na aquisição de informação, na cognição e no comportamento (Roederer, 1987).

Na percepção da fala humana, o sistema auditivo é levado ao seu limite de percepção acústica e interpretação. Portanto, é concebível que, com a evo-

lução da linguagem humana e o surgimento de áreas corticais especializadas na percepção da fala, uma diretriz tenha surgido para treinar o sentido acústico no reconhecimento de padrões sonoros sofisticados, como parte de um instinto humano inato de adquirir a linguagem desde o momento do nascimento. Durante os estágios finais do desenvolvimento intra-uterino, o sentido acústico do feto começa a registrar passivamente o ambiente sonoro intra-uterino. No nascimento existe uma transição súbita para uma resposta comportamental ativa em que a comunicação acústica com a mãe ou seu substituto tem um papel fundamental. Desse modo se estabelece um ciclo de realimentação de comunicação acústica que pode contribuir tanto para a relação emocional com a mãe como para o impulso motivador para a aquisição da linguagem. Os sons “musicais” simples e as sucessões rítmicas de sons (como as que são vocalizadas pela mãe) despertam a atenção da criança para ouvir, analisar e armazenar esses sons como prelúdio para a aquisição da linguagem. Isso pode ter levado ao surgimento da motivação para ouvir, analisar, armazenar e também vocalizar sons “musicais”, bem como da reação emocional ou recompensa límbica quando esses atos são realizados.

A motivação para descobrir simetrias e regularidades, para extrapolar, prever, lidar com a surpresa de uma mudança súbita ou a familiaridade da redundância e da repetição, e o impulso para explorar, diversificar e priorizar também contribuem com os elementos afetivos da música, que vão desde os instantâneos ou de curta duração até as sensações subjetivas de timbre, consonância, expectativa tonal, senso de retorno tonal e estruturas de longa duração ligadas às linhas melódicas. Esses elementos afetivos podem ser manifestações de recompensas límbicas em busca de um conteúdo fonético ou fonêmico do som e da identificação da organização gramatical e do conteúdo lógico dos sinais acústicos. O afortunado fato de que eles são irreprímíveis e ocorrem o tempo todo encontra-se como verdadeiro fundamento da moderna teoria musical (p. ex., Lerdahl e Jackendoff, 1983; Bharucha, 1994).

De modo geral, é possível que o efeito combinado da complexidade da identificação da mensagem sonora (*i.e.*, o número total de operações neurais necessárias por unidade de tempo) e as associações provocadas em outros centros cerebrais sejam a “causa” definitiva das sensações musi-

cais evocadas por uma certa mensagem musical. Por exemplo, quando a tonalidade de uma peça musical é modulada, o sistema nervoso auditivo deve rapidamente elaborar uma nova “lista de checagem” para as operações de identificação tonal que virão. Essa rápida elaboração, uma carga extra de milhares de operações individuais, forneceria a sensação musical peculiar relacionada à mudança de tonalidade. Os indivíduos não-musicais, que não conseguem experimentar essas sensações, são provavelmente pessoas cujo mecanismo de identificação de mensagens musicais não teve chance de desenvolver totalmente sua capacidade. Assim, embora eles ouçam tudo o que uma pessoa altamente musical ouve, o sistema auditivo central desses indivíduos não desenvolveu a habilidade de extrair informação musicalmente relevante de seqüências e superposições de sons não relacionados com a fala. Eventos musicais como mudanças de tonalidade são ouvidos, mas não interpretados, e por isso não provocam nenhuma resposta afetiva.

Desde o início da vida, a maioria das pessoas está exposta a um conjunto limitado de estímulos musicais. O condicionamento cultural rapidamente se impõe, e a resposta emocional começa a ser influenciada por fatores externos, alguns fortuitos, como o estado emocional experimentado por uma pessoa durante a primeira audição de uma certa obra musical ou um certo trecho; alguns mais controláveis, como o grau de repetição de formas musicais características pertencentes a um certo estilo musical; alguns induzidos pelo impulso inato de diversificar as possibilidades do empenho humano, que, no caso da música, tira proveito do desenvolvimento tecnológico, como o surgimento dos instrumentos de teclado ou, mais recentemente, dos sintetizadores eletrônicos. Isso pode determinar afinal por que um determinado estilo ou tipo de música é preferido. O que permanece invariante quanto aos instintos originais é:

1. O fato de que existe motivação para prestar atenção nos sons e formas musicais.
2. O fato de que uma reação emocional pode ser provocada.
3. O fato de que existem alguns componentes da música que são comuns a todas as culturas musicais.

Mas, mesmo num desenvolvimento posterior, culturalmente condicionado, podemos procurar outros elementos representativos de um valor perene da música.

Como um bom discurso público, a música pode ter sucesso em despertar e manter a atenção de grandes quantidades de pessoas, neutralizando os seus impulsos límbicos “normais” por longos períodos de tempo. Todos nós sabemos o que é o conteúdo informativo da linguagem falada, mas o que é transmitido na música? A música transmite informação sobre estados afetivos. Ela pode contribuir para a equalização dos estados emocionais de um grupo de ouvintes, assim como um discurso pode contribuir para a equalização do estado intelectual (conhecimento) do público. Infelizmente, não temos paradigmas quantitativos para descrever estados emocionais, como podemos fazer com os estados informativos relacionados com a linguagem. Mas o papel da música em rituais sexuais e de superstição, na religião, no proselitismo ideológico, nos movimentos militares, e mesmo em comportamentos anti-sociais, demonstra claramente o valor da música como meio de se alcançar uma coerência comportamental em grandes grupos de pessoas. Num passado distante, isso pode realmente ter tido um valor de sobrevivência muito grande, já que o ambiente humano cada vez mais complexo demandava ações coletivas coerentes por parte de grandes grupos da sociedade humana.

Até agora, deixamos de lado qualquer referência ao ritmo, um dos componentes fundamentais da música. Esta é uma omissão particularmente crítica, porque parece que o surgimento do ritmo tem sido sempre o primeiro passo na evolução de uma cultura musical. A propagação, pelo tecido cerebral, de um fluxo ciclicamente variável de sinais neurais disparados por padrões sonoros rítmicos pode de alguma forma entrar em “ressonância” com os “relógios” naturais do cérebro, que controlam as funções do corpo e as respostas comportamentais. Esses relógios provavelmente trabalham com base na atividade neural que se propaga em circuitos fechados, ou engramas, ou quaisquer outros sistemas de interligação neural que têm períodos naturais de resposta cíclica. Ainda é necessária muita pesquisa antes que possamos nos aventurar a dar uma resposta precisa à questão formulada no título desta seção; por que respondemos emocionalmente à música?

5.7. A ESPECIALIZAÇÃO DA FALA E O PROCESSAMENTO DA MÚSICA NOS HEMISFÉRIOS CEREBRAIS

No capítulo introdutório, referimo-nos rapidamente à notável divisão de tarefas encontrada entre os hemisférios cerebrais direito e esquerdo dos seres humanos (Seção 1.5). Concluiremos este último capítulo com uma discussão desse fenômeno, especialmente no que se refere à sua relevância para a música (p. ex., veja Scheid e Eccles, 1974; Bradshaw e Nettleton, 1981).

O corpo dos animais vertebrados apresenta uma simetria bilateral, especialmente com relação aos órgãos ligados à interação motora e sensorial com o ambiente. Essa simetria se estende aos hemisférios cerebrais, com o córtex esquerdo ligado ao lado direito do corpo e vice-versa. Esse cruzamento se verifica principalmente nos sistemas capazes de perceber a dimensão direcional, como a visão e a audição (p. ex., veja a Figura 2.25), e no controle motor referente a pernas e braços. Ele provavelmente se desenvolveu pela necessidade de manter no mesmo hemisfério cortical os mecanismos de interação, conectando a informação de entrada e as instruções motoras de saída relativas a eventos de um mesmo semi-campo espacial do ambiente. A imagem óptica é fisicamente invertida pelas lentes dos olhos, projetando o campo panorâmico direito na metade esquerda da retina e vice-versa, em cada olho. As metades esquerdas de ambas as retinas estão conectadas ao córtex visual esquerdo a fim de reunir num único hemisfério cerebral toda a informação relativa ao mesmo semi-campo espacial.

Conforme foi mencionado na Seção 2.9, ambos os hemisférios estão conectados entre si pelos 200 milhões de fibras do *corpus callosum* (e cerca de um milhão de fibras da junção anterior), o que restabelece a unidade global da representação do ambiente no cérebro. Nas vias auditivas aferentes existem conexões de mais baixo nível entre os canais de ambos os lados (p. ex., Figura 2.26), por meio das quais os sinais dos lados esquerdo e direito podem interagir para fornecer informações sobre a direção do som.

Na evolução do cérebro humano, as imensas exigências de processamento de informação suscitadas pelo desenvolvimento da comunicação

verbal levaram ao surgimento da especialização hemisférica. Nessa divisão de tarefas, as funções analíticas e sequenciais da linguagem tornaram-se a meta do hemisfério “dominante” (no lado esquerdo, em 97% das pessoas — Penfield e Roberts, 1959). O hemisfério menor se definiu como mais adaptado à percepção das relações holísticas, globais e sintéticas⁸. Há mais de cem anos sabe-se que os centros de fala estão localizados num certo hemisfério. Isso foi a conclusão de estudos de autópsia em pacientes com deficiências de fala e comunicação verbal (afasia, alexia, anomia, agrafia) adquiridas após uma hemorragia cerebral (derrame) no hemisfério esquerdo (p. ex., Geschwind, 1972). Por outro lado, verificou-se que lesões no hemisfério direito causavam uma deficiência no reconhecimento de padrões visuais (Kimura, 1963) e a perda da memória tonal e de timbres (Milner, 1967). De modo geral, todas as tarefas auditivas não-verbais são reduzidas nesses pacientes. Mais recentemente, exemplos convincentes foram documentados em estudos de pacientes com “cérebro seccionado”, cujo *corpus callosum* havia sido cortado transversalmente por razões terapêuticas (p. ex., Gazzaniga, 1970). Esses pacientes não conseguem descrever verbalmente nenhum objeto, palavra escrita ou evento localizado em seu campo visual esquerdo porque a informação sensorial pertinente, originalmente exposta no córtex visual do lado direito, não pode ser transferida aos centros de fala devido ao corte no *corpus callosum*. Uma técnica usada em pacientes sem traumas físicos no cérebro é a injeção de um barbiturato numa artéria carótida, anestesiando brevemente um hemisfério (este procedimento é usado às vezes para confirmar a localização dos centros de fala no lado esquerdo ou direito); uma série de testes com esses pacientes (veja o resumo em Borchgrevink, 1982) confirmou que a altura e a tonalidade na música (mas não na fala) estão vinculadas ao hemisfério direito, enquanto a produção e a compreensão da linguagem falada normal e do ritmo musical são tarefas do hemisfério esquerdo. Têm-se desenvolvido técnicas de

8. Suspeita-se que o processamento da fala esteja localizado no hemisfério que também controla a mão predominantemente usada (p. ex., o hemisfério esquerdo para a mão direita). Mas o canhotismo é muito mais freqüente do que os 3% de processamento de fala do lado direito.

não-invasão que podem ser usadas em pessoas normais e sadias para testar a especialização hemisférica. Uma dessas técnicas, chamada teste de audição dicótica, é baseada no fato, mencionado na Seção 2.9, de que a informação auditiva que flui pelos principais canais auditivos contralaterais (Figura 2.26) ignora qualquer informação conflitante que se propague pela via ipsilateral. Assim, quando ambos os ouvidos recebem informações conflitantes entre si, o córtex auditivo esquerdo prestará mais atenção à entrada do ouvido direito (embora também receba informações do ouvido esquerdo — Figura 2.26), e vice-versa. Na verdade, uma vantagem do ouvido direito foi constatada em tarefas de reconhecimento de fala, e o ouvido esquerdo leva vantagem nos testes de melodia (p. ex., Kimura, 1963). Finalmente, a eletroencefalografia, a tomografia de emissão de pósitrons e a ressonância magnética nuclear, que fornecem mapas da localização espacial de qualquer atividade elétrica ou metabólica cerebral intensificada durante uma certa operação mental, confirmam em termos gerais os resultados acima. A Tabela 5.4 (baseada numa análise de Bradshaw e Nettleton, 1981, e em suas referências) resume alguns dos aspectos básicos da especialização hemisférica em tarefas auditivas.

Tabela 5.4 Lista comparativa da especialização hemisférica nas tarefas auditivas (segundo Bradshaw e Nettleton, 1981)

Hemisfério esquerdo	Hemisfério direito
Consoantes explosivas	Vogais contínuas
Atributos fonológicos, sintaxe	Atributos estereotipados, rimas poéticas
Compreensão da fala	Entonação da fala, sons ambientais e animais
Fala proposicional	Conteúdo emocional da fala
Análise de sons falados sem sentido	Altura, timbre, tonalidade, harmonia
Texto falado (conteúdo verbal)	Texto cantado (conteúdo musical e fonético)
Ritmo, seqüências curtas de sons melódicos	Melodia holística
Memória verbal	Memória tonal

Por que razão essa curiosa dicotomia da função hemisférica surge ao longo da evolução humana?⁹ A razão mais plausível para esse desenvolvimento foi a necessidade de conservar o mais próximo possível as áreas responsáveis pelo processamento dos sinais de entrada da fala e pelo direcionamento das saídas vocal, gestual e mímica, a fim de minimizar os atrasos de transmissão entre as redes atuantes. As complexas operações seqüenciais de processamento de fala simplesmente não podem permitir o tempo necessário para a transmissão de sinais neurais de um hemisfério cerebral para o outro (aproximadamente 50 milissegundos)!¹⁰ Como resultado desse desenvolvimento, o substancial “espaço de processamento” no hemisfério esquerdo tornou-se inviável para as outras tarefas, mais lentas, de natureza holística e integrativa, que, assim, foram assumidas, “à revelia”, pelo hemisfério direito. É importante observar que a especialização dos hemisférios cerebrais é de natureza muito mais básica, envolvendo dois modos operacionais bem diferentes. Um modo envolve análise seqüencial de subpartes (subpartes temporais) de informação, como as necessárias ao processamento da linguagem. O outro envolve a integração espacial ou síntese de padrões momentâneos de atividade neural, para cumprir a determinação das qualidades holísticas dos estímulos de entrada (p. ex., Papçun e outros, 1974). Todavia, ambos os modos devem coexistir e cooperar a fim de processar a informação sobre o complexo ambiente humano e programar a resposta do organismo a ele¹¹. Em particular, podem ser necessárias tarefas seqüenciais (como a varredura visual) para o reconhecimento de padrões e a construção de imagens e, inversamente, podem ser necessárias imagens holísticas como operação paralela à programação seqüencial.

9. Nossos ancestrais primatas não apresentam tal especialização hemisférica distinta (embora isso ainda seja controverso: alguns animais apresentam uma assimetria hemisférica operacional quando se trata de tarefas de processamento seqüencial *versus* processamento holístico. Veja Denenberg, 1981.
10. Observe que é a extensão espacial que limita a velocidade de processamento dos computadores atuais pela mesma razão!
11. Por “ambiente humano” queremos dizer um ambiente contendo outros humanos que se intercomunicam.

Como a música é preferencialmente tratada pelo hemisfério menor, poderá isso significar que a música envolve principalmente operações sintéticas de reconhecimento de qualidade holística? Quanto ao reconhecimento de sons complexos, isso realmente parece estar de acordo com as teorias de “casamento de modelos” na percepção de alturas (Seção 4.8 e Apêndice II). A quantidade holística num estímulo musical é a distribuição espacial momentânea de atividade neural (correspondente aos máximos de ressonância na membrana basilar), levando à altura de som complexo (Seção 4.8), à discriminação de sons múltiplos (Seção 5.1), à consonância (Seção 5.2) e à expectativa de retorno tonal (Seção 5.5). Outra quantidade é a distribuição relativa da *quantidade* de atividade, dada pelo espectro de potência, e que leva à identificação de timbre e fonte sonora (Seção 4.9). Podemos fazer aqui uma boa analogia formal com a visão: o padrão (temporal) sonoro de entrada é “projetado” como um padrão espacial na membrana basilar, e o resultado é uma imagem *espacial*, na forma das imagens espaciais projetadas na retina. Daí para a frente, ambos os sistemas atuam sobre suas respectivas entradas em analogia formal, levando eventualmente a sensações musicais e pictóricas.

Um aparente paradoxo surge quando consideramos *melodias* e a dependência temporal das mensagens musicais. Exigirão elas um seqüenciamento, *i.e.*, operações do hemisfério dominante? Não necessariamente. Nosso cérebro reconhece as mensagens tipicamente musicais como de natureza holística, padrões temporais de longa duração, em vez de seqüências curtas. O fenômeno da fissão melódica (p. 261) é um exemplo bem convincente. Em outras palavras, parece que a música é identificada pelo nosso cérebro como a representação de imagens auditivas holísticas integrais (a estrutura harmônica), cuja seqüência temporal (de longa duração) ostenta em si um valor de *gestalt holístico* (os contornos melódicos). Embora inicialmente a atenção humana aos sons e mensagens musicais possa ser o resultado de algum instinto primitivo de se preparar para a aquisição de uma linguagem (Seção 5.6), a gramática e a estrutura da música são bem diferentes das da linguagem (Jackendoff e Lerdahl, 1982).

Conforme foi discutido acima, o lobo temporal esquerdo é especializado no processamento de entradas “verbais”, enquanto o direito lida com as

entradas “musicais”. Um resultado interessante é o seguinte: mesmo uma fala sem sentido, reproduzida do fim para o começo num gravador, é processada preferencialmente pelo hemisfério dominante (Scheid e Eccles, 1974). Isso só confirma o caráter de curto prazo das operações de seqüenciamento relevantes à fala, mencionadas acima: elas têm a ver com um estágio no processamento da informação acústica que vem antes do reconhecimento do conteúdo conceitual. Numa linha semelhante, os testes de audição dicótica (Bever e Chiarello, 1974) mostraram que, quando indivíduos musicalmente treinados se concentram nas estruturas finas de curta duração duma melodia, as estratégias correspondentes de processamento analítico são traçadas principalmente pelo hemisfério dominante. Uma situação inversa surge com um texto cantado. Verificou-se que pacientes com deficiências graves de linguagem (afasias) são capazes de cantar com palavras nitidamente compreensíveis uma canção aprendida antes de acontecer o trauma, mas não são capazes de *falar* essas mesmas palavras. Isso sugere que a linguagem, quando serve à expressão musical, é processada preferencialmente no hemisfério menor. Nessas situações (e provavelmente também na poesia), a linguagem tem uma qualidade bastante holística e sintética (Scheid e Eccles, 1974).

Tudo isso é bem apropriado à compreensão da evolução da música ocidental. Num sentido amplo, podemos descrever essa evolução como uma transição gradual entre duas configurações extremas. Em um dos extremos, encontramos os padrões sonoros altamente estruturados, claramente definidos, enfaticamente repetidos, tanto espaciais (harmônicos) quanto temporais (melódicos), e cada um porta um valor como um todo não analisado (p. ex., um certo acorde e uma certa voz ou progressão de acordes, respectivamente). No outro extremo (para o qual estamos caminhando na música contemporânea), identificamos as formas sonoras cujo valor fundamental é reconhecido no estado momentâneo das impressões sonoras temporais de curta duração. À luz do que dissemos acima sobre especialização hemisférica, podemos especular que essas duas configurações extremas estão intimamente relacionadas às duas estratégias distintas do cérebro humano para processamento. Só o futuro dirá se as tendências atuais da música representam meramente um esforço mais ou menos aleatório para

simplesmente “romper” com as formas tradicionais (que, em parte, surgiram naturalmente como resultado de propriedades físicas do sistema auditivo humano), ou se essas tendências podem ser canalizadas para uma exploração premeditada e subsequente utilização das vastas e ainda não testadas capacidades de processamento do sistema nervoso central.

Finalmente, devemos retornar à questão da simetria geral dos corpos e cérebros dos vertebrados. Gur e outros (1980) relataram que no córtex humano havia mais massa cinzenta (consistindo de neurônios compactamente interconectados) do que massa branca (axônios de transmissão mielinizados) no hemisfério esquerdo, em comparação com o direito. Isso pode indicar que o hemisfério esquerdo tende a mediar o processamento dentro das regiões corticais, e o hemisfério direito mediará o processamento entre elas. Outra assimetria notável foi identificada¹², mostrando que o *planum temporale* — uma área cortical que desempenha um papel importante na compreensão da linguagem — é significativamente maior no lobo temporal esquerdo do que no direito, em 65% de todos os casos examinados (Geschwind e Levitzky, 1968). Em 24% dos casos, não foi encontrada uma assimetria significativa e nos 11% restantes, a assimetria era inversa. Tal assimetria não foi encontrada em cérebros de primatas não-humanos. É altamente sugestivo que essa assimetria tenha alguma relação com a assimetria das tarefas realizadas pelos dois hemisférios. Embora a quantidade de 65% seja bem menor do que a porcentagem de representação de fala no hemisfério esquerdo (cerca de 97%), é bem possível que o *planum temporale* direito, bem maior, nos 32% restantes, seja uma indicação de uma capacidade inata bem maior de processamento de sons não-verbais. Além disso, foi sugerido (Scheid e Eccles, 1974) que o aumento do *planum temporale* direito pudesse ser uma medida da habilidade musical inata. Essa hipótese, se confirmada estatisticamente por meio de exames pós-morte sistemáticos¹³, daria um fundamento anatômico à transmissão hereditária da musicalidade.

12. Constatadamente existe até mesmo em bebês e fetos (Scheid e Eccles, 1974).

13. Músicos: doem seus restos mortais a uma causa nobre!

APÊNDICE II

ALGUNS ASPECTOS QUANTITATIVOS DO MECANISMO DE MOVIMENTO DO ARCO

Vamos considerar uma situação ideal: uma corda muito, muito longa friccionada num ponto A por um arco infinitamente fino (Figura AI.1). O arco move-se para cima com velocidade b . Além disso, vamos assumir que, desde o início, a corda adere ao arco. Isso significa que o ponto de contacto A (deveríamos dizer superfície de contato) também se move para cima com a mesma velocidade b . O resultado será uma deformação da corda na forma de uma onda transversal que se afasta do ponto A , como mostra a Figura AI.1 em diferentes instantes t_1, t_2, t_3 . Como as ondas transversais se propagam com uma velocidade V dada pela relação (3.3), que é muito maior do que qualquer velocidade de arco b razoável, a declividade b/V das porções deformadas da corda, AP e AQ , será, na verdade, extremamente pequena. Sob essas condições, a força transversal F aplicada ao arco (não confundir com a pressão de arco, que teria uma direção perpendicular ao papel) mantém o balanço com as projeções de ambas as forças de tensão sobre OA . Isso significa que $F = 2T b/V$. A fim de termos realmente um regime de atrito estático, com a corda aderente ao arco, a força F deve ser menor que um certo limite F_s , chamado de limite de atrito estático. Os resultados experimentais

mostram que esse limite é proporcional à pressão de arco: $F_s = \mu_s P \cdot \mu_s$ (a letra grega *miú*) é o coeficiente de atrito estático; ele depende da “aspereza” das superfícies em contato (nesse caso, a quantidade de resina nas cerdas do arco). A condição para aderência é então: $F = 2T b/V < \mu_s P$. Da mesma forma, a condição para deslizar é $F = 2T b/V > \mu_s P$. Como as quantidades V , T e μ_s são parâmetros constantes para uma dada corda, podemos resumir ambas as expressões da seguinte forma, mais física:

Quantidade controlável pelo músico	Quantidade fixa para cada corda	Tipo de movimento da corda em relação ao arco	
$\frac{b}{P}$	{	$< \frac{\mu_s V}{2T}$	(adere)
		$> \frac{\mu_s V}{2T}$	(desliza)

(AI.1)

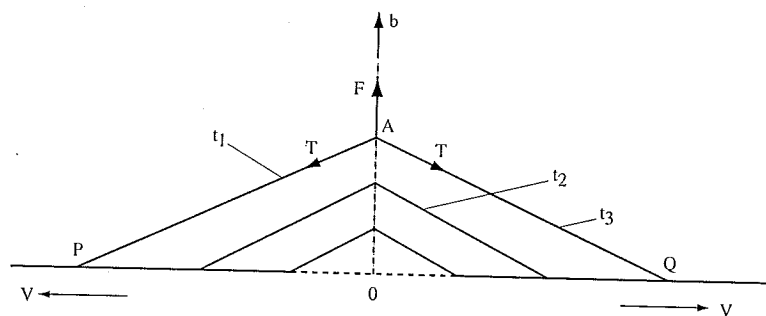


Figura AI.1

Nessas relações (AI.1), observe que o que importa é a razão entre a velocidade do arco e a pressão do arco, e não b ou P sozinhos. A razão b/P , portanto, define a natureza do movimento da corda friccionada.

1. < significa “menor que”; > significa “maior que”.

E se a corda deslizar desde o início (relação inferior de AI.1)? Nesse caso, a velocidade v do ponto de contato A da corda será menor que (e talvez esteja até na direção oposta a) a velocidade do arco, b . Teremos um regime de atrito dinâmico, onde se verifica que a força $F = 2T b/V$ (Figura AI.1) é proporcional a P , mas também depende da velocidade relativa $b-v$ entre o arco e a corda (velocidade de deslizamento). Escrevemos isso na forma $F = \mu_D P$, onde μ_D é o coeficiente de atrito dinâmico, que depende da velocidade relativa $b-v$ (é uma função de $b-v$). Assim, durante o regime de deslizamento:

$$\frac{b}{P} = \mu_D \frac{V}{2T} \tag{AI.2}$$

Se soubéssemos como μ_D depende da velocidade de deslizamento $b-v$, poderíamos usar a expressão (AI.2) para determinar a velocidade v do ponto de contato A da corda. Aqui também esta relação é governada pela razão b/P . Mas observe atentamente: o que é determinado aqui nesta razão é a diferença $b-v$, *i.e.*, a velocidade da corda vista pelo arco. Quanto maior for b , maior será v , para um dado valor de b/P . Enquanto esta razão determina a natureza do movimento da corda (aderências \times deslizamento, relações AI.1), a velocidade do arco determina a velocidade real da corda (para um certo b/P). Assim, se aumentarmos a velocidade do arco, mas ao mesmo tempo aumentarmos a pressão do arco, de modo a mantermos a razão entre essas grandezas constantes, a natureza do movimento da corda não mudará — apenas a sua velocidade aumentará linearmente com b . Isso levará a um aumento de amplitude, *i.e.*, da intensidade do som, na situação real. Em outras palavras: a amplitude de vibração de uma corda friccionada (volume do som) é controlada unicamente pela velocidade do arco, mas, para manter constante a natureza ou tipo de movimento da corda (timbre), é necessário manter a pressão do arco proporcional à velocidade do arco.

Vamos agora considerar um caso um pouco mais realista: uma corda de comprimento finito L , friccionada com um arco infinitamente fino no ponto central O (Figura 4.12). Nessa figura mostramos esquematica-

mente a forma da corda quando começamos a friccioná-la (aqui também as declividades estão bastante exageradas). v é a velocidade do ponto médio — tanto podemos ter deslizamento ($v < b$) como grude ($v = b$). Observe que, no instante $t_1 = L/2V$, a primeira “onda” (com declividade v/V) atinge os pontos extremos da corda. Aí, a onda é refletida e superpõe-se à onda propagante inicial, revelando o aspecto deformado que é visto nos tempos t_5 a t_7 . Então, em $t_8 = L/V$, algo novo acontece (Figura AI.2): a declividade muda subitamente no ponto de atrito. Isso modifica a expressão da força F e pode surgir um novo regime (p. ex., deslizamento, se antes estava ocorrendo). Não podemos prosseguir nossa discussão sem entrar em complexidades matemáticas consideráveis (Keller, 1953). Observe que essas variações cruciais na forma (sempre que a onda é refletida nos pontos extremos fixos) ocorrem sempre em momentos que são múltiplos inteiros de L/V , uma quantidade que é completamente independente do mecanismo de arco. Na verdade, o inverso de L/V aparece na expressão da frequência fundamental da corda vibrante (4.3). O leitor já pode então prever como essa frequência (e todos os harmônicos superiores) pode realmente ser excitada (e mantida) pelo mecanismo de arco e inferir da Figura AI.2 (com um pouco mais de imaginação) que, em seu movimento vibratório real, uma corda friccionada sempre apresenta uma forma instantânea que é formada por seções de linhas retas; esse resultado já foi verificado experimentalmente há muito tempo.

Assim como no caso de uma corda percutida ou dedilhada, a mistura particular de modos harmônicos de vibração dependerá da posição do ponto de fricção. A fricção próxima à ponte realçará os harmônicos superiores, tornando o som mais “brilhante”; a fricção próxima ao ponto médio das cordas (próximo ao início do braço do instrumento) reduz consideravelmente a intensidade dos harmônicos superiores, e o som será mais “macio”.

Nessa discussão, assumimos tacitamente que o arco se desloca em direção exatamente perpendicular à corda. Se houver um pequeno componente de movimento paralelo, os modos longitudinais de vibração da corda poderão ser excitados. As suas frequências são muito maiores do

que as frequências fundamentais dos modos transversais; eles são responsáveis pelos sons rangentos que se ouvem quando um principiante está tocando.

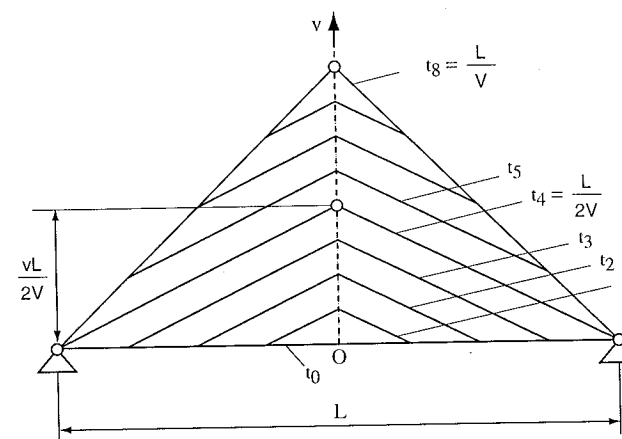


Figura AI.2

APÊNDICE II

ALGUNS ASPECTOS QUANTITATIVOS DE MODELOS DE PROCESSADOR DE ALTURA CENTRAL

Neste apêndice, continuamos nossa discussão sobre a percepção da altura de sons complexos que apresentamos nas Seções 2.9, 4.8 e 4.10. Em particular, mostraremos, com um pouco de álgebra, de que modo a teoria de “casamento de modelos” (Goldstein, 1973) pode explicar algumas características quantitativas da percepção da altura de sons complexos. Na segunda parte, vamos especular sobre os modelos neurais que podem realizar as funções de um processador central de altura e outras operações cognitivas.

O modelo de Goldstein (Seção 2.9) está baseado na suposição de que a informação neural nas posições espaciais dos máximos de ressonância da membrana basilar não está claramente definida, flutuando estatisticamente em torno de valores médios. O casamento do “modelo” que finalmente levará à sensação de altura subjetiva é tal que minimiza as diferenças entre os “valores”-modelo já incorporados e os do sinal real. Para ilustrarmos como poderia funcionar essa estimativa de semelhança de máximos, vamos aplicá-la aos experimentos de Smoorenburg (1970) (Seção 2.7). Consideremos um estímulo de dois sons com frequências $f_a = 1000$ Hz e

$f_b = 1200$ Hz. Essas são exatamente as frequências do quinto e do sexto harmônicos de uma fundamental $f_1 = 200$ Hz. Nosso “modelo” é representado por um conjunto harmônico de frequências $f_1, 2f_1, \dots, nf_1, \dots$, cuja fundamental f_1 pode ser mudada arbitrariamente. O processo de casamento consiste em encontrar uma frequência f_1 para a qual dois harmônicos sucessivos nf_1 e $(n+1)f_1$ coincidam ou estejam o mais próximo possível das frequências do som de entrada, f_a e f_b . Neste estágio, é irrelevante a ordem desses dois harmônicos (qual é o valor de n), desde que o casamento seja o melhor possível (que as diferenças $[nf_1 - f_a]$ e $[(n+1)f_1 - f_b]$ sejam as menores possíveis). Em nosso exemplo, apenas um casamento é ideal: para $n = 5$ e $f_1 = 200$ Hz, ambas as diferenças de frequência são exatamente zero — o casamento é perfeito. O leitor facilmente se dará conta de que não existe outra frequência fundamental f_1 , nem qualquer outro valor de n , que possa proporcionar um casamento perfeito. Observe atentamente que um determinado casamento requer a determinação ou estimativa de dois valores: a ordem harmônica n e a frequência fundamental f_1 . A tese da teoria é que essa frequência corresponderia à sensação de altura subjetiva, real e única evocada por um estímulo de dois sons. Tudo isso deveria acontecer de modo semelhante quando estivessem presentes mais de dois sinais pertencentes a uma série perfeitamente harmônica (como num som complexo), e deveria fornecer a frequência f_1 correspondente à altura do som complexo (não importando se a fundamental f_1 está realmente presente no estímulo original).

Como nos experimentos de Smoorenburg, vamos agora mudar as frequências para o par não-harmônico $f_a = 1050$ Hz e $f_b = 1250$ Hz. Não existe série harmônica da qual essas frequências sejam componentes vizinhas. Como funciona o casamento de modelos neste caso? De acordo com a teoria acima (numa versão hipersimplificada), devemos encontrar um par de valores n e f_1 tais que os harmônicos vizinhos nf_1 e $(n+1)f_1$ minimizem o erro quadrático médio relativo ε :

$$\varepsilon^2(n) = \frac{1}{2} \left[\left(\frac{f_a - nf_1}{f_a} \right)^2 + \left(\frac{f_b - (n+1)f_1}{f_b} \right)^2 \right] \quad (\text{AII.1})$$

Para cada valor de n haverá uma frequência f_1 que minimiza o valor do erro $\varepsilon(n)$ e que pode ser facilmente encontrada com cálculo elementar¹:

$$f_1 = \frac{f_a}{n} \frac{1 + \frac{n+1}{n} \frac{f_a}{f_b}}{1 + \left(\frac{n+1}{n} \frac{f_a}{f_b} \right)^2} \quad (\text{AII.2})$$

Inserindo esse valor de f_1 na expressão (AII.1), obtemos o valor correspondente do erro $\varepsilon(n)$:

$$\varepsilon^2(n) = \frac{1}{2} \frac{\left(1 - \frac{n+1}{n} \frac{f_a}{f_b} \right)^2}{1 + \frac{n+1}{n} \frac{f_a}{f_b}} \quad (\text{AII.3})$$

Em vez de trabalhar com o erro $\varepsilon(n)$, podemos usar o seu inverso

$$Q(n) = \frac{1}{\varepsilon(n)} \quad (\text{AII.4})$$

que chamaremos de “qualidade de ajuste”. Um casamento perfeito ($\varepsilon = 0$) implica $Q = \infty$ (“ajuste infinitamente bom”). Para diferentes valores de n obteremos diferentes frequências fundamentais f_1 e diferentes valores de qualidade de ajuste $Q(n)$. Chegaremos então a uma série $Q(n), Q(n+1), \dots$. Se, dentre esses valores, um se sobressair como o maior, os correspondentes f_1 e n proporcionarão o melhor casamento — e f_1 representará a frequência que é ouvida! Em nosso exemplo de dois sons, com $f_a = 1050$ Hz e $f_b = 1250$ Hz, obtemos, para $n = 4, f_1 = 256,0$ e $Q(4) = 41$; para $n = 5, f_1 = 209,2$ e $Q(5) = 251$; para $n = 6, f_1 = 176,8$ e $Q(6) = 99$. Obviamente, a escolha de

1. O leitor verificará que assim pode obter a resposta certa para o exemplo harmônico acima, com $n = 5$ e $f_1 = 200$ Hz, para os quais $\varepsilon = 0$ (ajuste perfeito).

$n = 5$ nos dá o maior valor de Q . A frequência correspondente (209,2 Hz) é na verdade a altura subjetiva mais facilmente identificada quando esse complexo não-harmônico de dois sons é apresentado (veja Smoorenburg, 1970). Observe que $n = 6$ e, em menor extensão, $n = 4$ também fornecem uma qualidade de ajuste não desprezível; isso explica o fato observado que as frequências fundamentais correspondentes $f_1 = 176,8$ Hz e 256,0 Hz podem ser identificadas como sensações de altura “secundárias”, embora com muito maior dificuldade. Assim, esse modelo explica quantitativamente as sensações de altura múltiplas ou ambíguas provocadas por sons não-harmônicos. Se afastarmos ainda mais da harmonicidade as frequências do estímulo de dois sons (mas sempre conservando a mesma diferença de frequência de 200 Hz), obteremos os resultados vistos na Tabela AII.1.

Tabela AII.1. Valores de f_1 (AII.2) e $Q(n)$ (AII.4) para um complexo de dois sons com f_a e f_b .

f_b/f_a (Hz)	$n = 4$	$n = 5$	$n = 6$	$n = 7$
1 200/1 000	$f_1 = 240,0$	200,0	169,0	146,0
	$Q(n) = 49$	∞	91	41
1 250/1 050	256,0	209,2	176,8	153,0
	41	251	99	49
1 300/1 100	263,1	218,3	184,5	159,7
	36	131	155	60
1 350/1 150	278,2	227,4	192,3	166,5
	32	91	323	75
1 400/1 200	289,3	236,6	200,0	173,2
	29	91	∞	97

Observe que, para o par 1 300/1 100, dois casamentos se evidenciam para as ordens harmônicas $n = 6$ e $n = 5$. Repare também, pela Tabela AII.1, como a ordem harmônica n na qual é obtido o maior valor de Q muda de 5 para 6 quando a frequência central do par f_a e f_b aumenta, enquanto a altura de melhor ajuste diminui de um valor acima de 200 Hz para um abaixo de 200 Hz. De modo geral, como f_a e f_b aumentam continuamente (man-

tendo-se a diferença $f_b - f_a$ constante), a principal sensação de altura subjetiva “oscila” em torno da frequência de repetição dada por $f_b - f_a$ (200 Hz), coincidindo com esta última nas posições harmônicas. Alturas múltiplas ou ambíguas aparecem mais distintamente sempre que um estímulo de dois sons se situe aproximadamente a meio caminho entre duas posições harmônicas. Nesse contexto, assumimos que a qualidade de ajuste Q está relacionada à “clareza” ou inteligibilidade da sensação de altura correspondente. É importante ressaltar que os valores teóricos obtidos com esse procedimento de casamento de modelos não se ajustam aos resultados experimentais quando a ordem n for maior que aproximadamente 7. Sons resultantes dos tipos (2.5) e (2.6) (Seção 2.5) desempenham seu papel e devem ser levados em consideração como sinais adicionais de altura baixa, antes de se aplicar o procedimento do casamento de modelos (p. ex., Plomp, 1976).

Esse método de ajuste pode ser estendido a estímulos de vários sons. É bastante ilustrativo utilizar essa versão estendida para descobrir a altura ou alturas previstas para combinações “estranhas”, como, por exemplo, sons complexos superpostos, que é o que acontece nos acordes musicais² (veja mais adiante). É interessante observar que essa teoria de extração de altura via casamento de modelos funciona em analogia formal com o mecanismo de ajuste de alturas num instrumento de sopro (p. 200).

Uma conseqüência de importância musical é a seguinte. As relações (AII.2) e (AII.3) podem ser usadas para mostrar que, quando um certo harmônico numa certa vibração complexa desafina ($f_n = nf_1 + \epsilon$), o efeito sobre a altura subjetiva é muito pequeno (da ordem de $\epsilon/2n$ ou menor). Assim, as inarmonicidades menores dos harmônicos que surgem nas cordas vibrantes (p. 162) terão apenas um efeito de segunda ordem sobre a altura subjetiva resultante.

Vamos agora especular sobre um modelo neural para o processador central de alturas como uma continuação da “matriz de aprendizagem” apresentada por Terhardt (1974). Faremos isso aqui meramente como um exercício acadê-

2. Deve-se tomar muito cuidado ao se estender as relações (AII.1) e (AII.2) a um caso de muitos sons — a simples soma no numerador e no denominador não fornece o resultado correto, a menos que todos os componentes da entrada sejam quase harmônicos.

mico, sem a pretensão de representar outra teoria de altura central. Esboçamos, na Figura AII.1, um esquema de ligação neural capaz de realizar as operações necessárias para a extração de altura e o rastreamento de fundamentais, com base na matriz de aprendizagem de Terhardt (1974). Assumimos que as fibras horizontais conduzam os sinais neurais combinados de ambas as cócleas para o processador de altura primária (espectral). Essa saída leva informação sobre cada componente harmônico de um som complexo, mas assumimos que ela seja normalmente ignorada nos estágios superiores do processamento de sons musicais. Esses axônios horizontais são interceptados por uma cadeia vertical de dendritos neuronais (Seção 2.8), como se vê na Figura AII.1. Assumimos que, inicialmente (no nascimento), as conexões sinápticas ativas estão distribuídas como se mostra para os neurônios K , L e M . Assumimos ainda que, para chegar ao limite e ao disparo, cada neurônio vertical deve ser ativado em muitos contatos sinápticos quase simultaneamente. Por esse quadro, fica claro que, num cérebro acusticamente virgem, a distribuição de atividade de saída dos neurônios verticais (ao longo da dimensão x) seria quase idêntica à das fibras horizontais (ao longo da dimensão y).

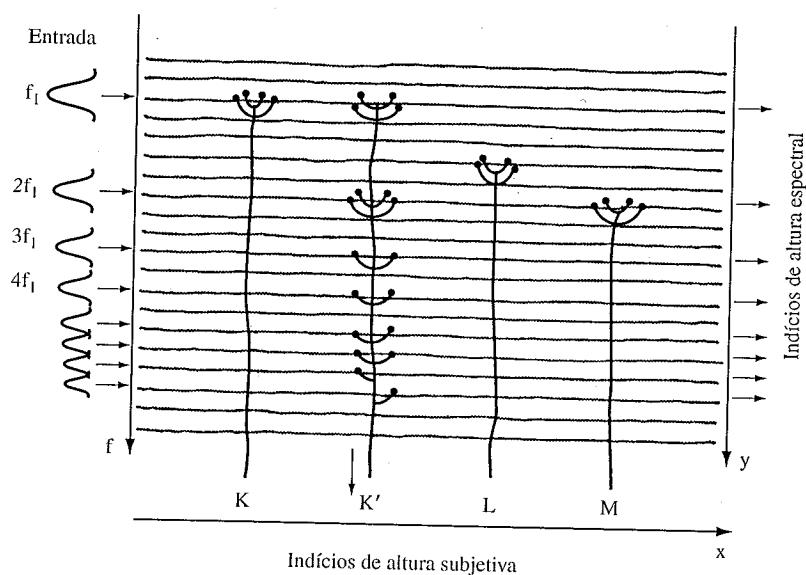


Figura AII.1

Nossa próxima suposição, em concordância com a teoria de Terhardt, é que, quando o ouvido é exposto repetidamente a sons harmônicos, os contatos sinápticos entre um neurônio vertical e entre todos os axônios horizontais que provavelmente estejam disparando ao mesmo tempo também são ativados (a “essência” do processo de aprendizagem no sistema nervoso). Quando um som harmônico complexo cuja fundamental seja f_1 é ouvido, o neurônio K (Figura AII.1) inicialmente responderá apenas a essa fundamental. Mas, quando esse complexo sonoro é repetido, o dendrito de K desenvolverá contatos sinápticos ativos com todas as fibras horizontais cujas melhores freqüências correspondam a harmônicos desse som (p. 82). Como resultado, o neurônio vertical fica “afinado” com toda a série harmônica de f_1 , como é exemplificado na Figura AII.1 para o neurônio K' . Assim, os neurônios verticais representariam fisicamente uma primeira aproximação dos “modelos” postulados pela teoria de Goldstein (veja acima). A resposta do modelo será maior, sempre que um melhor casamento local (máximo local de excitação do neurônio vertical) for conseguido. Finalmente, assumimos que a localização dos máximos de atividade de saída dos neurônios verticais (ao longo da dimensão x , Figura AII.1) leva à sensação de altura subjetiva ou de periodicidade. Após o processo de aprendizagem, essa saída será na verdade bem diferente daquela das fibras horizontais (na dimensão y). Por exemplo, se uma quantidade suficiente de sinapses for ativada pelos harmônicos de f_1 , o neurônio K' responderá, mesmo que esteja faltando a fundamental f_1 no som original. Isso representa o mecanismo de rastreamento de fundamentais. Quanto maior a ordem dos harmônicos, menos definida será a entrada “horizontal”, por causa da proximidade dos respectivos máximos de excitação (Figura AII.1). Assim, os neurônios verticais podem ser levados a responder ao sinal de entrada “errado” (um que não corresponda à freqüência fundamental à qual a sua árvore dendrítica está sintonizada). Múltiplas alturas são possíveis, portanto, como mostramos quantitativamente na primeira parte deste Apêndice.

Contudo, nosso modelo precisa de aperfeiçoamentos. Como se vê na Figura AII.1, o neurônio “afinado” K' também responderia a todos os sons complexos cujas freqüências fundamentais sejam múltiplos inteiros

de f_1 . Ele também pode disparar quando apenas um harmônico estiver presente. Para evitar que esse efeito indesejável aconteça, podemos introduzir um conjunto intermediário de neurônios verticais capazes de detectar coincidências entre harmônicos vizinhos. É bem simples escrever um programa para um computador digital que simule a operação desse tipo de modelo neural³.

Alguns resultados quantitativos são vistos na Figura AII.2 (para suposições simples, mas realistas sobre a distribuição de excitação primária em torno de cada harmônico e sobre a degradação da resposta com o aumento da ordem harmônica). No topo de cada figura encontra-se o espectro de potência de entrada primária (em escalas lineares), correspondendo à superposição de dois sons complexos formando uma oitava, uma quinta e uma terça menor, respectivamente. Os gráficos inferiores representam a distribuição de atividade neural computada ao longo da dimensão x (cf. Figura AII.1). Repare nos picos pronunciados correspondendo às frequências fundamentais de cada som complexo original. Supomos que esses picos sejam reconhecidos num estágio superior do processamento neural e levem às duas sensações de altura principais correspondentes num complexo de dois sons. A posição e, em grande parte, a forma desses picos primários são independentes dos espectros de potência reais dos sons componentes, dependendo apenas das suas frequências fundamentais f_a e f_b . Observe também que, seguindo a ordem de consonância decrescente (Seção 5.2), picos “parasitas” aparecem nas posições da frequência de repetição r e de seus múltiplos. Esses picos parasitas (que devem ser deliberadamente inibidos em algum estágio superior) estão ausentes na oitava. Além disso, existe uma atividade de fundo ou “nível de ruído” abaixo do som de frequência mais baixa e que aumenta com a diminuição da ordem de consonância.

3. Esse programa não faz suposições quanto a transformações de Fourier ou funções de auto-correlação. O programa simplesmente realiza um “modelo” de frequências harmônicas dentro de um domínio determinado de frequências fundamentais e conta, para cada posição, o número de “excitações” simultâneas em posições harmônicas vizinhas. O número total de pares de excitações simultâneas representa a saída (a intensidade ou probabilidade de ativação do neurônio vertical correspondente).

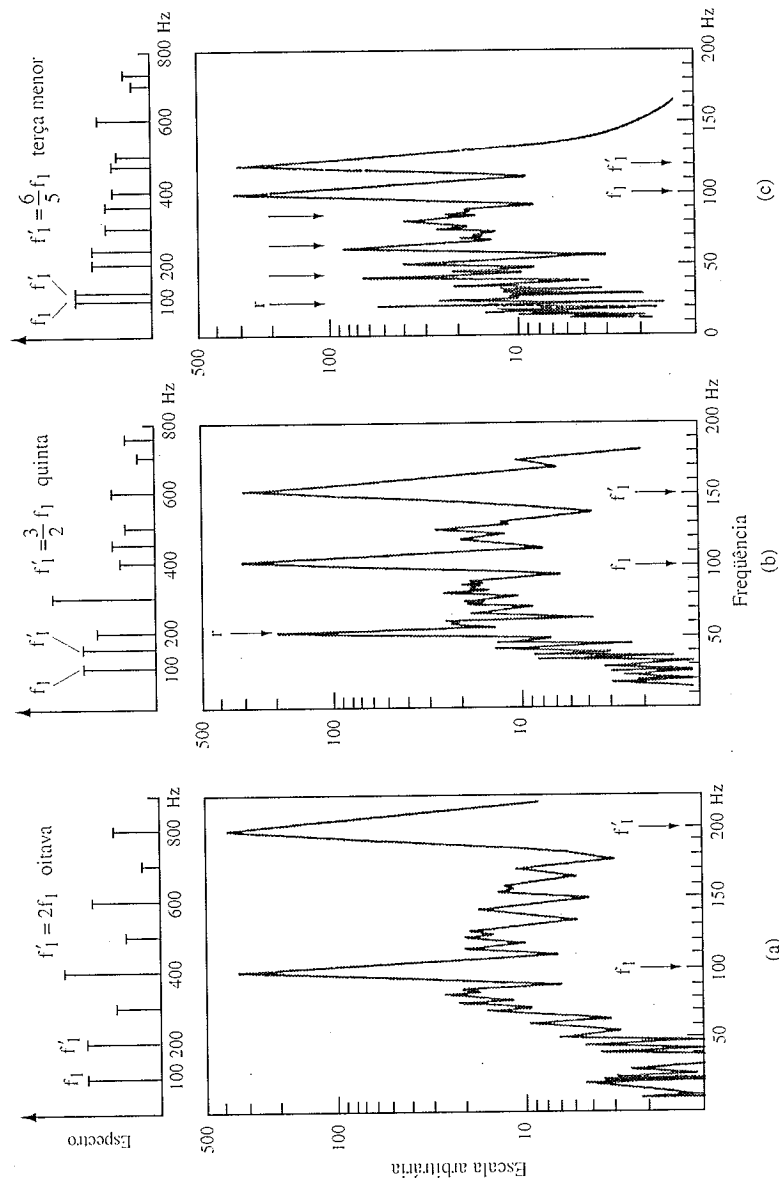


Figura AII.2

O modelo neural discutido acima é bem primitivo e, por seu esquema, está um tanto quanto afastado da realidade fisiológica — exceto pelo que chamamos de “essência” do processo de aprendizagem: o estabelecimento de novas sinapses ou a mudança de eficiência das sinapses existentes entre os neurônios (Seção 4.10). Um grande progresso tem sido conseguido em anos recentes no desenvolvimento de modelos mais realistas de circuitos neurais (p. ex., veja Hinton, 1992, ou, para maiores detalhes, Arbib, 1987, e Kohonen, 1988), em parte pelo esforço de cientistas da computação e engenheiros da robótica interessados no projeto de *computadores neurais*. A computação neural está começando a ser usada no estudo da percepção de altura e outras tarefas musicais cognitivas. Todavia, uma discussão detalhada iria muito além do âmbito deste livro, de modo que podemos apenas apresentar resumidamente alguns conceitos básicos.

Os modelos de circuito neural são formados por unidades interconectadas representando neurônios-modelo. Cada conexão comporta um “peso” modificável, que determina a eficiência da transmissão da “frequência de disparo” ou atividade que passa por essa conexão (um peso negativo corresponderia a uma sinapse inibidora). Cada unidade recebe atividades de muitas outras, multiplica cada uma pelo fator de peso correspondente a essa sinapse, soma tudo e transforma essa entrada total ponderada numa atividade única de saída (veja a Figura 2.21). Os tipos comuns de modelo de circuito têm suas unidades dispostas em várias camadas distintas. A primeira camada recebe a entrada “do mundo exterior”, representando, na percepção de altura, por exemplo, a distribuição de atividade na membrana basilar (veja o eixo vertical na Figura AII.1). A camada de saída deve fornecer distribuições de atividade neural específicas de um padrão que o circuito neural está projetado para reconhecer (p. ex., um “sinal” de altura separado para cada padrão sonoro complexo de entrada, como na Figura AII.2, ou, num caso bem mais complexo, um “sinal cognitivo” separado para cada instrumento musical dentro de um som orquestral que chega ao ouvido).

Inicialmente, os pesos sinápticos são todos pré-atribuídos (p. ex., todos iguais, ou escolhidos aleatoriamente). Esses pesos são então gradualmente modificados em etapas de “treinamento”; em cada etapa, a saída real é

comparada com a saída desejada, e os pesos são modificados de acordo com alguns registros numéricos complexos para produzirem uma melhor aproximação. Esse procedimento é repetido incontáveis vezes até que se consiga uma precisão preestabelecida. Após se completar a “aprendizagem”, freqüentemente se verifica que o sistema é capaz de reconhecer corretamente padrões incompletos (a fundamental ausente!) ou certos padrões relativos que *não foram fornecidos* ao sistema durante o processo de treinamento (Kohonen, 1988). Essa capacidade de interpretar situações novas (não-aprendidas) é uma propriedade fundamental dos modelos avançados de circuito neural e essencial para a descrição dos modos de processamento distribuídos, paralelos e holológicos da função cerebral (Seção 4.10), que permite a um sistema que tenha evoluído para realizar um conjunto específico de tarefas realizar operações “novas” que não tinham nenhuma função na evolução genética, como a escrita e a leitura (Kosslyn e Koenig, 1992) — e música...? (Veja Seção 5.5).

De que modo os modelos neurais correntes se adaptam à realidade neural? Primeiro, há a questão do que se chama “propagação de volta”, *i.e.*, efetuar mudanças no circuito que o ajustem à saída desejada. Isso exige que a informação de saída realmente o circuito neural na direção oposta à da atividade “preferencial” da camada de entrada — a arquitetura neural do córtex exibe na verdade muitas vias que aparentemente justificam isso! Segundo, esses modelos precisam de um “professor” para fornecer a saída desejada, checar a qualidade do ajuste e comandar as alterações apropriadas nos pesos sinápticos — ou seja, esse processo está incorporado àquilo que chamamos de “sinal de mesmice” na Figura 4.30⁴. Por outro lado, o modelo primitivo na Figura AII.1 é um exemplo e circuito de aprendizagem “não-supervisionada”, em que o surgimento repetitivo de um aspecto específico no sinal de entrada leva automaticamente a mudanças na arquitetura sináptica após muitas exposições, fornecendo uma saída específica em seus detalhes após o exercício de treinamento ter se completado. Finalmente, existem no sistema nervoso central sistemas pré-

4. Uma criança olhando um objeto que segura na mão recebe um “sinal de mesmice” pelo sentido do tato, enquanto a imagem visual do objeto muda quando ele é deslocado ou rodado.

conectados processando informação sensorial cujas arquiteturas sinápticas surgiram pela evolução.

Muito ainda resta por ser feito no desenvolvimento da computação neural. Por exemplo, até agora a maioria dos modelos de circuito neural é projetada para lidar com a análise de padrões espaciais (mesmo os sistemas de reconhecimento de voz operam principalmente com base na análise de espectros sonoros “instantâneos”). Todavia, o cérebro precisa lidar com configurações sequenciais, trabalhando com distribuições espaciais e *temporais* de atividades neurais (veja as Seções 2.8 e 4.10). Assim, é necessário projetar modelos onde a atividade circule em *loops*, ou circuitos fechados. Uma questão bem mais difícil é compreender a importante capacidade de “aprendizagem imediata” do nosso cérebro, que pode armazenar informação *sem* exposição reiterada⁵ e implantá-la com sucesso nos modelos de circuito neural. Finalmente, é preciso desenvolver a matemática para uma compreensão abrangente e uma manipulação quantitativa de todos os algoritmos envolvidos (p. ex., Arbib, 1987). Quando tudo isso for conseguido, teremos completado uma importante parte do estágio meramente inicial do desenvolvimento da computação neural!

Devido ao seu enorme impacto em potencial sobre as neurociências, as ciências do comportamento, a robótica, a matemática e a informática, e às suas muitas aplicações na física, economia, ecologia etc., o desenvolvimento de modelos de circuito neural e de computadores neurais é um dos mais desafiadores empenhos interdisciplinares da ciência e tecnologia. Poderá a música se beneficiar disso? Um olhar ao passado nos mostra que, desde o início dos tempos, a música sempre acompanhou prontamente os mais recentes avanços da ciência e da tecnologia. São testemunhas os trompetes, flautas e harpas do Egito antigo! São testemunhas as maravilhas acústicas e mecânicas dos órgãos do século XV! São testemunhas o uso dos computadores na teoria e composição musicais e a refinada instrumentação eletroacústica de hoje!

5. Esse estudo se concentra no *hipocampo*, uma estrutura cerebral que desempenha um importante papel na armazenagem e recuperação de informações da memória.

APÊNDICE III

ALGUMAS OBSERVAÇÕES SOBRE O ENSINO DA FÍSICA E PSICOFÍSICA DA MÚSICA

Não seria realista fazer recomendações elaboradas sobre como organizar um curso verdadeiramente interdisciplinar sobre este assunto. A principal razão está na constituição quase imprevisível de uma classe formada por estudantes interessados e que se habilitariam a fazer esse curso com as suas formações bem diferentes e o amplo espectro de interesses. Supondo que esse curso seja aberto a todo o corpo discente da universidade (e assim deveria ser), ele poderá conter cinco populações principais: estudantes de:

1. Música.
2. Psicologia.
3. Ciências humanas em geral.
4. Comunicações de massa e engenharia.
5. Física e matemática.

A dificuldade mais geral é tornar o curso igualmente interessante, útil e facilmente compreensível para todos. Isso impõe três exigências gerais:

1. Minimizar o uso da matemática — mas fazê-lo de modo a não ridicularizar a apresentação do assunto aos olhos dos estudantes de engenharia e ciências exatas. (Sugestão: use o curso para mostrar a esses estudantes explicitamente como ensinar ciência sem matemática!)
2. Explicar tudo “desde o princípio”, tanto um assunto de física, psicofísica quanto de música — mas fazê-lo de modo a não parecer condescendente com os respectivos “especialistas”. (Sugestão: use o curso para mostrar explicitamente aos “especialistas” como apresentar resumos abrangentes e concisos de tópicos de suas próprias áreas!)
3. Realizar experimentos e demonstrações em aula e promover testes, ensaios e problemas que sejam “significativos”, *i.e.*, conduzidos de tal modo que o estudante (a despeito de sua formação) possa responder o seguinte questionário sem hesitar: o que eu aprendi presenciando essa demonstração, realizando esse experimento, ou resolvendo esse problema? (Sugestões: nos experimentos não permita que os estudantes assistam a medições bobas ou as façam apenas por fazê-las — mostre a eles como as magnitudes se relacionam umas com as outras na natureza, como elas se modificam umas em relação às outras e como estão conectadas por meio de relações físicas de causa-e-efeito. Nos problemas, não deixe que eles apenas “resolvam equações” — mais uma vez, mostre como uma determinada relação liga uma ou mais quantidades “dinamicamente” em toda uma gama de variabilidade, induza-os a um “sentimento” intuitivo das relações quantitativas entre as magnitudes; mostre a eles como as relações matemáticas podem ser usadas para prever o comportamento de um sistema. Nos testes, leve-os a pensarem intuitivamente, mas respondendo com precisão científica.)

Uma séria dificuldade é que muitos estudantes de música (e de outras artes) têm um “pavor” inerente ao rigor científico, uma vez que assumem *a priori* que “não compreenderão”. Isso é simplesmente um bloqueio mental que pode ser dissipado com sucesso por meio da persuasão, paciência e dedicação à pessoa por parte do professor. (Sugestão: convença-os de que, se também são capazes de controlar suas despesas mensais, serão também capazes de compreender o pouco de matemática necessário neste curso!)

A inclusão da psicoacústica num curso introdutório de acústica musical apresenta diversos outros desafios ao professor. Primeiro, há o mais óbvio: como encaixar tudo no tempo disponível. Não importando se o tempo é muito ou pouco, será necessário tomar decisões difíceis sobre quais tópicos deixar de lado e quais incluir. Segundo, a psicoacústica e os tópicos da neuropsicologia que lhe são relativos são assuntos talvez ainda menos familiares do que a física para os estudantes de acústica musical. Isso torna necessário restringir esses tópicos apenas aos mais relevantes e interessantes.

Será útil apresentar, logo no início do curso, alguns aspectos relevantes da psicoacústica. Por exemplo, alerte os estudantes que as recentes perscrutações no campo da percepção musical podem ser incorporadas na criação de novas fronteiras para a composição musical. Ressalte que muitas das falácias existentes sobre a *performance* musical têm origem em modos particulares de processamento de informação acústica no ouvido e no cérebro. Mostre que muitos dos requisitos técnicos de equipamentos eletroacústicos de alta qualidade estão diretamente relacionados a aspectos particulares do processamento de sinais no sistema nervoso. Ressalte que a compreensão da percepção musical deve ser interesse não somente de músicos, mas também de neuropsicólogos, que podem se beneficiar do conhecimento sobre certas funções cerebrais, e dos psicólogos, que podem obter informações quantitativas de relevância para a terapia musical.

Uma dificuldade geral é o fato de que as demonstrações experimentais sofisticadas relativas a esse assunto requerem um equipamento muito caro. Contudo, é possível sair-se bem com um mínimo básico de equipamentos, que provavelmente poderão ser tomados de empréstimo de outros cursos ou departamentos. Esse mínimo básico, com o qual se pode demonstrar a maioria dos experimentos citados neste livro, está resumido adiante.

EXPERIMENTAÇÃO PSICOACÚSTICA

1. Dois geradores de ondas senoidais, um amplificador de boa qualidade, bons fones de ouvido para todos, dois alto-falantes *hi-fi*.

2. Um osciloscópio para cada grupo de quatro a seis estudantes, se possível com duplo feixe e memória de traço.
3. Um sintetizador eletrônico (uma versão portátil é suficiente!).

Com esse equipamento é perfeitamente possível demonstrar quase tudo o que está mencionado nas pp. 39, 42, 50, 58, 60, 64, 67, 71-72, 139, 216, 242, 255-256. Se a classe tiver acesso a um órgão de tubos num bom ambiente acústico, podem-se fazer ainda outras demonstrações úteis (p. ex., veja as pp. 74, 138-139, 140-141, 211, 216-217, 237). Em todos esses experimentos e demonstrações, deve-se ter como regra que tudo o que estiver sendo ouvido deve simultaneamente ser mostrado na tela do osciloscópio.

EXPERIMENTAÇÃO ACÚSTICA

1. Um “sonômetro”² por grupo de quatro a seis estudantes, com estroboscópio e circuito apropriado para realizar os experimentos descritos nas páginas 162-163. Essa configuração também prepara para a discussão experimental das relações (4.2), (4.3) e dos mecanismos de arco e de pinçagem.
2. Um piano é útil para a realização das demonstrações simples das páginas 162-163.
3. Tubos de órgão avulsos (em geral encontráveis nos departamentos de física) para explorar curvas de ressonância, como as da Figura 4.24 e as relações (4.5) e (4.6), usando um pequeno alto-falante de boa qualidade “implantado no tubo”.
4. Tanques de ondas e filmes em *loop*, em geral facilmente encontráveis, são extremamente úteis para demonstrar ondas progressivas, ondas estacionárias e “óptica” acústica em geral.

1. Veja a excelente descrição de experimentos selecionados e demonstrações de classe feita por Hartmann (1975).
2. Um nome tolo para uma única corda montada sobre uma caixa de ressonância, com tensão mensurável e controlável.

5. A experimentação acústica deve ser acompanhada por tarefas de casa significativas. Uma ótima coleção delas pode ser encontrada em Savage (1977).

Além de tudo isso, é aconselhável propor aos estudantes que realizem individualmente estudos específicos e escrevam ensaios sobre um instrumento musical de sua escolha, o que, obviamente, necessitará de referências à literatura apropriada.

Para resumir, este é um curso desafiador e agradável de se ministrar — talvez o mais interdisciplinar de todos os que uma universidade possa oferecer num nível para principiantes. Ele oferece uma chance, tanto ao professor como ao aluno, de voar alto com a imaginação — dentro das rígidas fronteiras da ciência!

BIBLIOGRAFIA

- ALLEN, J. B. & NEELY, S. T. 1992. "Micromechanical Models of the Cochlea". *Phys. Today*, July 1992, 40.
- ANDO, Y. 1995. *Concert Hall Acoustics*. Berlin, Springer-Verlag.
- ARBIB, M. A. 1987. *Brains, Machines, and Mathematics*. New York, Springer-Verlag.
- ASKENFELT, A. & JANSSON, E. 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. Amer.*, 56:1266.
- BACKUS, J. & HUNDLEY, T. C. 1971. "Harmonic Generation in the Trumpet". *J. Acoust. Soc. Amer.*, 49:509.
- BENADE, A. H. 1971. "Physics of Wind Instrument Tone and Response". In: *Symposium on Sound and Music*, December 1971. Washington, D. C., American Association for the Advancement of Science.
- _____. 1973. "The Physics of Brasses". *Sci. Amer.*, 229(1):24.
- _____. 1976. *Fundamentals of Musical Acoustics*. New York-London-Toronto, Oxford University.
- BEVER, T. G. & CHIARELLO, R. J. 1974. "Cerebral Dominance in Musicians and Nonmusicians". *Science*, 185:537.
- BHARUCHA, J. J. 1994. "Tonality and Expectation", In: AIELLO, R. (ed.). New York-Oxford, *Musical Perceptions*. Oxford University Press, 213.

- BILSEN, F. A. & GOLDSTEIN, J. L. 1974. "Pitch of Dichotically Delayed Noise and Its Possible Spectral Basis". *J. Acoust. Soc. Amer.*, **55**:292.
- BORCHGREVINK. 1982. "Prosody and Musical Rhythm Are Controlled by the Speech Hemisphere". In: CLYNES, M. (ed.). *Music, Mind, and Brain*. New York-London, Plenum Press, 151.
- BRADSHAW, J. L. & NETTLETON, N. C. 1981. "The Nature of Hemispheric Specialization in Man". *Behavioral and Brain Sci.*, **4**:51.
- BREDBERG, G.; LINDERMANN, H. H.; ADES, H. W.; WEST, R. & ENGSTRM, H. 1970. "Scanning Electron Microscopy of the Organ of Corti". *Science*, **170**:861.
- BRODAL, A. 1969. *Neurological Anatomy*. New York, Oxford University Press.
- CORSO, J. F. 1957. "Absolute Judgments of Musical Tonality". *J. Acoust. Soc. Amer.*, **29**:138.
- DALLOS, P. 1992. "The Active Cochlea". *J. of Neuroscience*, **12**:4575.
- DAMASKE, P. 1971. "Head-related Two-channel Stereophony with Loud-speaker Reproduction". *J. Acoust. Soc. Amer.*, **50**:1109.
- DAVIS, H. 1962. "Advances in the Neurophysiology and Neuroanatomy of Cochlea". *J. Acoust. Soc. Amer.*, **34**:1377.
- DEBOER, 1983. "No Sharpening? A Challenge to Cochlear Mechanics". *J. Acoust. Soc. Amer.*, **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: CLYNES, M. (ed.) *Music, Mind, and Brain*. Plenum Press, London, New York, 119
- _____. (ed.). 1983. *The Psychology of Music*. New York, Academy Press.
- EGAN, J. P. & HAKE, H. W. 1950. "On the Masking Pattern of a Simple Auditory Stimulus". *J. Acoust. Soc. Amer.*, **22**:622.
- FLANAGAN, J. L. 1972. *Speech Analysis, Synthesis and Perception*. 2. ed. New York, Springer-Verlag.
- FLETCHER, H. & MUNSON, W. A. 1933. "Loudness, Its Definition, Measurement and Calculation". *J. Acoust. Soc. Amer.*, **5**:82.
- FRIEDLANDER, F. G. 1953. "On the Oscillations of a Bowed String". *Cambridge Phil. Soc. Proc.*, **49**:516.
- GAZZANIGA, M. S. 1970. *The Bisected Brain*. Meredith Corporation.
- GELFAND, S. A. 1990. *Hearing*. New York, Marcel Dekker Inc.
- GESCHWIND, N. 1972. "Language and the Brain". *Sci. Amer.*, **226**(4):76.
- GESCHWIND, N. & LEVITZKY, W. 1968. "Human Brain: Left-right Asymmetries in the Temporal Speech Region". *Science*, **161**:186.
- GOLDBERG, J. M. & BROWN, P. B. 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: PLOMP, R. & SMOORENBURG, G. F. *Frequency Analysis and Periodicity Detection in Hearing*. Leiden, A. W. Suijthoff.
- _____. 1973. "An Optimum Processor Theory for the Central Formation of the Pitch of Complex Tones". *J. Acoust. Soc. Amer.*, **54**:1496.

- GREY, J. M. & GORDON, J. W. 1978. "Perceptual Effects of Spectral Modifications of Musical Timbres". *J. Acoust. Soc. Amer.*, **63**:1493.
- GUR, R. C.; PACKER, I. K.; HUNGERBUHLER, J. P.; REIVICH, M.; OBRIST, W. D.; AMARNEK, W. S. & SACKEIM, H. A. 1980. "Differences in Distribution of Gray and White in Human Cerebral Hemispheres". *Science*, **207**:1226.
- HALL, D. & ASKENFELT, A. 1988. "Piano String Excitation V: Spectra for Real Hammers and Strings". *J. Acoust. Soc. Amer.*, **83**:1627.
- HARTMANN, W. M. 1975. "The Electronic Music Synthesizer and the Music". *Amer. J. Phys.*, **43**:755.
- _____. 1993. "On the Origin of the Enlarged Melodic Octave". *J. Acoust. Soc. Amer.*, **93**:3400.
- HERRINGTON, R. N. & SCHNEIDAU, P. 1968. "The Effect of Imagery on the Waveshape of the Visual Evoked Response". *Experientia*, **24**:1136.
- HINTON, G. E. 1992. "How Neural Networks Learn from Experience". *Sci. Amer.*, September 1992, 145.
- HOSTMA, A. J. M. 1970. "Perception of Musical Pitch". *J. Acoust. Soc. Amer.*, **48**:88(A).
- HOSTMA, A. J. M. & GOLDSTEIN, J. L. 1972. "Perception of Musical Intervals: Evidence for the Central Origin of the Pitch of Complex Tones". *J. Acoust. Soc. Amer.*, **51**:520.
- HUBEL, H. 1971. "The Visual Cortex of the Brain". In: *Contemporary Psychology. Readings from Scientific American*. San Francisco, W. H. Freeman and Company.
- HUDSPETH, A. J. 1985. "The Cellular Basis of Hearing: The Biophysics of Hair Cells". *Science*, **230**:745.
- _____. 1989. "How the Ear's Works". *Nature*, **341**:397.
- HUCHINS, C. M. & FIELDING, F. L. 1968. "Acoustical Measurements of Violins". *Phys. Today*, **21**(7):34.
- IVERSON, P. & KRUMHANS, C. L. 1993. "Isolating the Dynamic Attributes of Musical Timbre". *J. Acoust. Soc. Amer.*, **94**:2595.
- JACKENDOFF, R. & LERDAHL, F. 1982. "A Grammatical Parallel between Music and Languages". *Music, Mind, and Brain*. CLYNES, M. New York-London, Plenum Press.
- JANSSON, E.; MOLIN, N.-E. & SUNDIN, H. 1970. "Resonances of a Violin Body Studied by Hologram Interferometry and Acoustical Methods". *Physica Scripta*, **2**:243.
- JOHNSTONE, B. M.; PATUZZI, R. & SELICK, P. 1983. "Comparison of Basilar Membrane, Hair Cell and Neural Responses". In: KLINKE, R. & HARTMANN, R. (eds.). *Hearing - Psychological Bases and Psychophysics*. Berlin-Heidelberg-New York-Tokyo, Springer-Verlag, 46.
- JOHNSTONE, B. M.; PATUZZI, R. & YATES, G. K. 1986. "Basilar Membrane Measurements and the Traveling Wave". *Hearing Res.*, **22**:147.
- KACHAR, B.; BROWNELL, W. E.; ALTSCHULER, W. E. & FEX, J. 1986. "Electrokinetic Shape Changes of Cochlear Outer Hair Cells". *Nature*, **322**:365.
- KAMEOKA, A. & KURIYAGAWA, M. 1969. "Consonance Theory. Part II: Consonance of Complex Tones and Its Calculation Method". *J. Acoust. Soc. Amer.*, **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. Amer.*, **64**:1386.

- KIANG, N. Y.-S.; WATANABE, T.; THOMAS, E. C. & CLARK, L. F. 1965. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MIT Press.
- KIMURA, D. 1963. "Right Temporal Lobe Damage". *Arch. Neurol.*, **8**:264.
- KLEIN, W.; PLOMP, R. & POLS, L. C. W. 1970. "Vowel Spectra, Vowel Spaces and Vowel Identification". *J. Acoust. Soc. Amer.*, **48**:999.
- KOHONEN, T. 1988. *Self-Organization and Associative Memory*. Berlin, Springer-Verlag.
- KOSSLYN, S. M. & KOENIG, O. 1992. *Wet Mind: The New Cognitive Neuroscience*. The Free Press.
- LERDAHL, F. & JACKENDOFF, R. 1983. *A Generative Theory of Tonal Music*. Cambridge, MIT Press.
- LIBERMAN, M. C. 1978. "Auditory-nerve Response from Cats Raised in a Low Noise Chamber". *J. Acoust. Soc. Amer.*, **63**:442.
- LICKLIDER, J. C. R. 1959. "The Auditory Theories". In: KOCH, S. (ed.). *Psychology: A Study of a Science*. New York, McGraw-Hill Book Company, vol. 1.
- LIPPS, T. 1905. *Psychologische Studien*. Leipzig, Durr'sche Buchhandlung.
- MATTHEWS, M. V. & KOHUT, J. 1973. "Electronic Stimulation of Violin Resonances". *J. Acoust. Soc. Amer.*, **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: MILLIKAN, C. H. & DARLEY, F. L. (eds.). *Brain Mechanisms Underlying Speech and Language*. New York-London, Grune & Stratton.
- MILNER, B.; TAYLOR, L. & SPERRY, R. W. 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 Tones of Different Frequencies". *Percept Psychophys.*, **14**:1.
- _____. 1974. "Psychophysical Verification of Predicted Interaural Differences in Localizing Distant Sound Sources". *J. Acoust. Soc. Amer.*, **55**:139.
- MOORE, B. C. J. 1973. "Frequency Difference Limens for Short-duration Tones". *J. Acoust. Soc. Amer.*, **54**:610.
- PANKSEPP, J. 1982. "Toward a General Psychobiological Theory of Emotions". *Behavioral and Brain Sci.*, **5**:407.
- PAPCUN, G.; KRASHEN, S.; TERBEEK, R.; REMINGTON, R. & HARHMAN, R. 1974. "Is the Left Hemisphere Specialized for Speech, Language and/or Something Else?". *J. Acoust. Soc. Amer.*, **55**:319.
- PATTERSON, B. 1974. "Musical Dynamics". *Sci. Amer.*, **231**(5):78.
- PENFIELD, W. & ROBERTS, L. 1959. *Speech and Brain Mechanisms*. Princeton, Princeton University Press.
- PIERCE, J. R. 1983. *The Science of Musical Sound*. New York-San Francisco, Scientific American Books, W. H. Freeman and Co.
- PLOMP, R. 1964. "The Ear as a Frequency Analyser". *J. Acoust. Soc. Amer.*, **36**:1628.
- _____. 1965. "Detectability Threshold for Combination Tones". *J. Acoust. Soc. Amer.*, **37**:1110.
- _____. 1967. "Beats of Mistuned Consonances". *J. Acoust. Soc. Amer.*, **42**:462.
- _____. 1967a. "Pitch of Complex Tones". *J. Acoust. Soc. Amer.*, **41**:1526.
- _____. 1970. "Timbre as a Multidimensional Attribute of Complex Tones". In: PLOMP, R. & SMOOREMBURG, F. G. (eds.). *Frequency Analysis and Periodicity Detection in Hearing*. Leiden, A. W. Suihoff.
- _____. 1976. *Aspects of Tone Sensations*. London, New York-San Francisco, Academic Press.
- PLOMP, R. & STEENEKEN, H. J. M. 1973. "Place Dependence of Timbre in Reverberant Sound Fields". *Acustica*, **28**:50.
- PLOMP, R. & BOUMAN, M. A. 1959. "Relation between Hearing Threshold and Duration for Tone Pulses". *J. Acoust. Soc. Amer.*, **31**:749.
- PLOMP, R. & LEVELT, W. J. M. 1965. "Tonal Consonance and Critical Bandwidth". *J. Acoust. Soc. Amer.*, **38**:548.
- PRIBRAM, K. H. 1971. *Languages of the Brain*. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.
- RAIFORD, C. A. & SCHUBERT, E. D. 1971. "Recognition of Phase Changes in Octave Complexes". *J. Acoust. Soc. Amer.*, **50**:559.
- RAKOWSKI, A. 1971. "Pitch Discrimination at the Threshold of Hearing". In: *Proc. 7th Int. Congr. Acoust.*, Budapest, 3:373.
- _____. 1972. "Direct Comparison of Absolute and Relative Pitch". In: *Proc. Symp. Hearing Theory*, IPO Eindhoven.
- RATLIFF, F. 1972. "Contour and Contrast". *Sci. Amer.*, **226**(6):91.
- REINICKE, W. & CREMER, L. 1970. "Application of Holographic Interferometry to Vibrations of the Bodies of String Instruments". *J. Acoust. Soc. Amer.*, **48**:988.
- RHODE, W. S. & ROBLES, L. 1974. "Evidence from Mössbauer Experiments for Nonlinear Vibration in the Cochlea". *J. Acoust. Soc. Amer.*, **55**:588.
- RICHARDS, A. M. 1977. "Loudness Perception for Short-duration Tones in Masking Noise". *J. Acoust. Soc. Amer.*, **20**:684.
- RITSMA, R. J. 1967. "Frequencies Dominant in the Perception of the of Complex Sounds". *J. Acoust. Soc. Amer.*, **42**:191.
- ROEDERER, J. G. 1978. "On the Relationship between Human Brain Functions and the Foundations of Physics, Science, and Technology". *Found. of Physics.*, **8**:423.
- _____. 1979. "Human Brain Functions and the Foundations of Science". *Endeavour*, **3**:92.
- _____. 1984. "The Search for a Survival Value of Music". *Music Perception*, **1**:350.
- _____. 1987. "Neuropsychological Processes Relevant to the Perception of Music: An Introduction". In: SPRINTGE, R. & DROH, R. (eds.). *Music in Medicine*. Berlin-Heidelberg-New York, Springer-Verlag, 81.
- ROSE, J. E.; BRUGGE, J. F.; ANDERSON, D. J. & HIND, J. E. 1969. "Some Possible Neural Correlates of Combination Tones". *J. Neurophys.*, **32**:402.
- RUGGERO, M. A. & RICH, N. 1991. "Application of a Commercially-manufactured Doppler-shift Laser Velocimeter to the Measurement of Basilar Membrane Motion". *Hearing Res.*, **51**:215.
- SACHS, M. B. & ABBAS, P. J. 1974. "Rate versus Level Functions for Auditory-nerve Fibers in Cats: Tone-burst Stimuli". *J. Acoust. Soc. Amer.*, **56**:1835.
- SAUNDERS, F. A. 1946. "The Mechanical Action of Instruments of the Violin Family". *J. Acoust. Soc. Amer.*, **17**:69.

- SAVAGE, W. R. 1977. *Problems for Musical Acoustics*. New York, Oxford University Press.
- SCHARF, B. 1983. "Loudness Adaptation". In: TOBIAS, J. V. & SCHUBERT, E. D. (eds.). *Hearing Research and Theory*. New York, Academic Press, vol. 2.
- SCHEID, P. & ECCLES, J. C. 1975. "Music and Speech: Artistic Functions of the Human Brain. In: *Psychology of Music* (no prelo).
- SCHELLENG, J. C. 1973. "The Bowed String and the Player". *J. Acoust. Soc. Amer.*, **53**:26.
- 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: TOBIAS, J. V. (ed.). *Foundations of Modern Auditory Theory*. New York, Academy Press.
- SMALL, A. M. 1970. "Periodicity Pitch". In: TOBIAS, J. V. (ed.). *Foundations of Modern Auditory Theory*. New York, Academy Press.
- SMOORENBURG, G. F. 1970. "Pitch Perception of Two-frequency Stimuli". *Acoust. Soc. Amer.*, **48**:924.
- _____. 1972. "Audibility Region of Combination Tones". *J. Acoust. Soc. Amer.*, **52**:603.
- SOKOLICH, W. G. & ZWISLOCKI, J. J. 1974. "Evidence for Phase Oppositions between Inner and Outer Hair Cells". *Acoust. Soc. Amer.*, **55**:466.
- SOMMERHOFF, G. 1974. *Logic of the Living Brain*. London-New York-Sydney-Toronto, John Wiley and Sons.
- STEVENS, S. S. 1955. "Measurement of Loudness". *J. Acoust. Soc. Amer.*, **27**:815.
- _____. 1970. "Neural Events and the Psychophysical Law". *Science*, **170**:1043.
- STEVENS, S. S. ; VOLKMANN, J. & NEWMAN, E. B. 1937. "A Scale for the Measurement of Psychological Magnitude Pitch". *J. Acoust. Soc. Amer.*, **8**:185.
- SUNDBERG, J. (ed.). 1992. *Gluing Tones*. Royal Swedish Academy of Music (com CD).
- TERHARDT, E. 1971. "Pitch Shifts of Harmonics of the Octave Enlargement Phenomenon". In: *Proc. 7th Int. Congr. Acoust.*, Budapest, **3**:621.
- _____. 1972. "Zur Tonhöhenwahrnehmung von Klängen, I, II". *Acustica*, **26**:173, 187.
- _____. 1974. "Pitch, Consonance and Harmony". *J. Acoust. Soc. Amer.*, **55**:1061.
- TERHARDT, E. & FASTL, H. 1971. "Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen". *Acustica*, **25**:53.
- TERHARDT, E. & ZICK, M. 1975. "Evaluation of the Tempered Tone in Normal, Stretched and Contracted Intonation". *Acustica*, **32**:268.
- VAN NOORDEN, L. A. P. S. 1975. *Temporal Coherence in the Perception of Tone Sequence*. Eindhoven, Institute for Perception Research (com gravação fonográfica demonstrativa).
- VON BÉKÉSY, G. 1960. *Experiments in Hearing*. New York, McGraw-Hill Book Company.
- VON HELMHOLTZ, H. 1863. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Trad. ingl., 1954. New York, Dover Publications.
- 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: TOBIAS, J. V. (ed.). *Foundations of Modern Auditory Theory*. New York, Academic Press.

- WHITFIELD, I. C. 1967. *The Auditory Pathway*. London, Edward Arnold Ltd.
- WIGHTMAN, F. L. 1973. "The Pattern-transformation Model of Pitch". *J. Acoust. Soc. Amer.*, **54**:407.
- YOST, W. A. & WATSON, C. S. (eds.). 1987. *Auditory Processing of Complex Sounds*. New Jersey-London, Lawrence Erlbaum Associates, Publ. Hillsdale.
- ZWICKER, E. & FASTL, H. 1990. *Psychoacoustics*. Berlin-Heidelberg-New York, Springer-Verlag.
- ZWICKER, E. & SCHARF, B. 1965. "A model of Loudness-summation". *Psych. Rev.*, **72**:3.
- ZWICKER, E.; FLOTTORP, G. & STEVENS, S. S. 1957. "Critical Bandwidth in Loudness Summation". *J. Acoust. Soc. Amer.*, **29**:548.
- ZWISLOCKI, J. J. 1965. "Analysis of Some Auditory Characteristics". In: LUCE, R. D.; BUSH, R. R. & GALANTER, E. (eds.). *Handbook of Mathematical Psychology*. New York, Wiley.
- _____. 1969. "Temporal Summation of Loudness: An Analysis". *J. Acoust. Soc. Amer.*, **46**:431.
- ZWISLOCKI, J. J. & SOKOLICH, W. G. 1973. "Velocity and Displacement Responses in Auditory-nerve Fibers". *Science*. **182**:64.

ÍNDICE REMISSIVO

A

- Absorção (do som) 207
- Afinação 62, 245
- Altura 21, 44, 48, 74, 146, 165, 197, 215
- Altura absoluta 258
- Altura de periodicidade 75, 96, 147, 216, 236, 244
- Altura subjetiva, *ver* Altura de periodicidade.
- Amplitude 41, 54, 68, 71, 91, 171, 179, 277
- Análise (sonora) 176, 213
- Análise de Fourier 177
- Antinodos, *ver* Ventres.
- Ápice 45
- Aprendizado (como processo neural) 29, 221, 225, 244, 258
- Aspereza (de uma superposição de dois sons) 57, 240, 242
- Ápice 174
- Atuação do arco 230
- Autocorrelação (de sinais neurais) 92

B

- Banda crítica 58, 62, 137, 145, 180, 219, 243
- Batimentos de primeira ordem 56, 63, 68, 239
- Batimentos de segunda ordem 69
- Batimentos subjetivos, *ver* Batimentos de segunda ordem.

C

- Caixa de ressonância do violino 181
- Células capilares 47, 85, 144, 150, 151
- Cent (subdivisão de frequência) 254
- Ciclos por segundo, *ver* Hertz.
- Circuito neural (auto-organizador) 229
- Clarinete (instrumentos do tipo) 192, 199, 201
- Cóclea 19, 44, 149
- Codificação (no sistema nervoso auditivo) 48, 89, 143
- Coefficiente de absorção 210
- Coluna de ar (aberta) 188
- Coluna de ar (fechada) 191

Comprimento de onda 114
 Comprimento de onda longitudinal 119
 Comprimento de onda transversal 116
 Condicionamento cultural 24, 30, 34, 222, 260
 Consonância 23, 239, 248, 259
 Cor (das tonalidades) 257
 Cordas 157, 163, 164, 168, 172, 275
 Corte (de picos de ressonância) 204
 Córtex cerebral 24, 228, 268
 Cromá 246
 Curva de resposta, *ver* Curva de ressonância.
 Curva de ressonância 185, 199, 201

D

Decaimento (de uma vibração) 170, 208
 Decibel 131
 Detecção de direcionabilidade (do som) 22, 92, 93
 Detectores de detalhes 85, 227
 Diferença no limite do observável (de frequência) 50
 Diferença no limite do observável (de intensidade) 129, 132
 Difração (do som) 211
 Discriminação de frequência (ou altura) 56, 58, 59, 236
 Discriminação do timbre 237
 Dissonância 23, 239
 Dominância (de uma nota musical numa seqüência) 260
 Duração (efeito sobre o volume) 141, 145

E

Efeito *chorus* 237
 Efeito "festa" 238
 Efeito de precedência 238
 Emoção 141
 Energia 107
 Engramas neurais 267
 Escala justa 248, 253
 Escala musical 24, 259, 248
 Escala pitagórica 250, 253

Escala temperada 252, 253, 254
 Espectro (sonoro) 21, 179, 186, 196, 206, 211
 Evolução (da percepção musical, processamento neural etc.) 25, 34, 245, 264

F

Fala, *ver* Linguagem.
 Fase 42, 69, 179, 211, 218
 Física 25
 Física clássica 26
 Física quântica 27
 Flauta (instrumentos do tipo) 190, 199, 202
 Fon (unidade de volume) 135
 Força 105
 Formantes 187, 206
 Frequência 43, 119
 Frequência central 58
 Frequência de disparo neural 85
 Frequência de repetição (*ver também* Frequência fundamental) 75, 89, 177
 Frequência de ressonância 185, 215
 Frequência fundamental 21, 160, 165, 177, 190, 195, 196, 201, 241
 Frequência padrão 257
 Fundamental ausente (*ver também* Altura de periodicidade) 75, 205

H

Harmônicos 76, 77, 161, 164, 176, 179, 200, 215, 236, 241
 Harmônicos aurais 67, 140
 Hertz (unidade de frequência) 44

I

Identificação (de instrumentos musicais), *ver* Reconhecimento.
 Impulso neural 82, 87, 145
 Inarmônico (modo, vibração) 163, 193, 201, 204
 Instrumentos "belos" 222
 Instrumentos de metal 197, 204
 Instrumentos (do som) 21, 121, 128, 207

Intervalos musicais 71, 241, 246, 248, 256

J

Joule (unidade de energia) 107

L

Linguagem 34, 264
 Logaritmo 130

M

Magnitude psicofísica 29, 52, 179, 219
 Mascaramento 133, 138, 140, 141
 Mecanismo de decodificação (no cérebro) 229, 238
 Mecanismo de depuração 62, 149
 Mecanismo de excitação 18, 20, 166, 194, 196, 198
 Medição psicofísica 28, 133
 Meio 19, 112
 Melodia (linha melódica) 23, 216, 238, 265, 272
 Membrana basilar 46, 85, 143, 215
 Memória (como processo neural) 221, 226, 230, 258, 259
 Mensagem musical (*ver também* Melodia) 23, 258, 261, 264, 272
 Modos de vibração 162, 166, 167, 182, 188, 193
 Movimento 36
 Movimento harmônico 41, 108
 Movimento periódico 38, 177
 Música 33, 235

N

Não-harmônicos (modos de vibração) - *ver* Inarmônicos.
 Neurônio 48, 82, 144, 228
 Newton (unidade de força) 105
 Nível de mascaramento 133, 139
 Nível de pressão sonora 132
 Nível de volume 135
 Nível sonoro 131
 Nós 124, 159, 188, 200
 Notação exponencial 43

O

Oboé (instrumentos do tipo) 199, 201
 Oitava 64, 65, 73, 90, 139, 161, 202, 239, 246
 Oitava desafinada 68, 69
 Ondas estacionárias 124, 132, 159, 188, 212
 Ondas longitudinais 110, 117, 278
 Ondas sonoras 104, 106, 110
 Ondas transversais 111, 113, 158
 Órgão 75, 138, 141, 193, 213, 238, 255
 Orifícios (efeito deles) 203, 204
 Oscilação amortecida 108, 170
 Osciloscópio 39, 69
 Ouvido interno, *ver* Cóclea.

P

Padrão de disparo neural 230, 232
 Padrão de vibração 39, 69, 85, 90, 91, 222, 237
 Período 39, 42
 Piano (experimentos com ele) 163, 165, 172
 Pico de ressonância 200, 204
 Potência 108, 123
 Pressão (do ar) 106
 Processamento (de informação acústica) 23, 34, 145, 229, 236, 262, 273
 Progressão do som, *ver* Melodia.
 Propagação do som (*ver também* Ondas sonoras) 103
 Psicofísica 27
 Público (efeitos sobre o som) 210
 Pulso neural, *ver* Impulso neural.

Q

Qualidade do som, *ver* Timbre.
Quantum neural 52

R

Rastreamento de fundamentais (*ver também* Altura de periodicidade) 77, 81, 204
 Realimentação (em colunas de ar) 197, 200, 204
 Reconhecimento (de instrumentos musicais) 22, 187, 211, 221, 248
 Reflexão do som 121, 207

Introdução à Física e Psicofísica da Música

- Região de ressonância (na membrana basilar) 48, 62, 85, 96, 143, 145, 152, 214, 228, 236, 239, 240
- Registro agudo (de um instrumento de sopro) 202
- Registro grave (de um instrumento de sopro) 200
- Registro médio (de um instrumento de sopro) 202
- Resolução de frequência 51, 60, 259
- Resolução de timbre 219
- Ressonador 19, 170, 181
- Reverberação 113, 208, 238
- Ritmo 265
- Ruído 247
- S**
- Semitom 250, 251
- Senoidal, *ver* Movimento harmônico.
- Sensação 28
- Sensação musical 34, 265
- Senso de retorno 23, 260
- Sinapse 82
- Síntese (de sons) 177, 213
- Sistema métrico 36
- Sistema nervoso auditivo 19, 82, 144, 214, 221, 227, 266
- Sistema nervoso central, *ver também* Córtex cerebral.
- Sistemas *hi-fi* (qualidade deles) 67, 180
- Sobre-som 202, 204
- Sobretom (*ver também* Harmônicos) 161
- Som 35
- Som (unidade de volume) 136
- Soma de volumes 137, 138, 145
- Som de borda 194
- Som de diferença (*ver também* Sons resultantes) 65
- Som de palheta 196
- Som direto 207, 238
- Som puro 44
- Sons complexos 157, 179, 216, 235, 243
- Sons resultantes 64, 66
- Sons vogais 206
- Superposição de ondas (*ver também* Superposição de som) 123
- Superposição de som 53, 137, 164, 176, 235
- T**
- Temperatura absoluta 112, 190
- Tempo de reverberação 208
- Timbre 21, 165, 186, 204, 215
- Tímpano 19, 44, 106, 214
- Tom inteiro 251
- Tonalidade 23, 266, 252, 255
- Touché* (em execução pianística) 172
- Trabalho 107
- Transcorrelação (de sinais neurais) 92
- Transientes 22, 180, 211, 218, 237
- Triade 249
- Tubo de órgão, de palheta 202
- Tubo de órgão, tipo flauta 190, 193, 199, 202
- Tubo fechado, *ver* Coluna de ar fechada.
- V**
- Varição de pressão 106, 117, 131
- Velocidade (de onda) 112
- Ventres 124, 188
- Vibração 35, 38
- Vibração amortecida, *ver* Oscilação amortecida.
- Vibrato 238
- Volume 21, 44, 128, 133, 135, 143, 207
- Volume subjetivo, *ver* Volume.
- Voz humana 206
- W**
- Watt (unidade de potência) 109

<i>Título</i>	<i>Introdução à Física e Psicofísica da Música</i>
<i>Autor</i>	Juan G. Roederer
<i>Tradução</i>	Alberto Luis da Cunha
<i>Produção</i>	Ponto & Linha
<i>Projeto Gráfico</i>	Ponto & Linha
<i>Capa</i>	Andrés Sandoval
<i>Editoração Eletrônica</i>	Ponto & Linha
<i>Editoração de Texto</i>	Alice Kyoko Miyashiro
<i>Ilustrações</i>	Ponto & Linha
<i>Revisão de Texto</i>	Maria Luiza X. Souto
<i>Revisão de Provas</i>	Ricardo Koichi Miyake
	Fabrizio Waltrick
<i>Arte-final</i>	Julia Yagi
	Andrea Yanaguaita
<i>Divulgação</i>	Maria Helena Arrigucci
	Mônica Cristina G. dos Santos
	Flavia Cristina Moino
<i>Secretaria Editorial</i>	Rose Pires
	Eliane Reimberg
<i>Formato</i>	16 x 23 cm
<i>Mancha</i>	11 x 18 cm
<i>Tipologia</i>	Times Roman 10,5/15
<i>Papel</i>	Cartão Supremo 250 g/m ² (capa)
	Pólen 85 g/m ² (miolo)
<i>Número de Páginas</i>	312
<i>Tiragem</i>	1 000
<i>Laserfilm</i>	Ponto & Linha
<i>Impressão</i>	Imesp