
Faculdade de Engenharia Elétrica e de Computação
Departamento de Semicondutores Instrumentos e Fotônica

Auralização: técnicas de modelagem e simulação binaural de ambientes acústicos virtuais.

Autor: Antônio Carlos Lopes Fernandes Júnior

Orientador: Prof. Dr. Furio Damiani

Banca Examinadora:

Prof. Dr. Furio Damiani - DSIF/FEEC/UNICAMP

Prof. Dr. Peter Jürgen Tatsch - DSIF/FEEC/UNICAMP

Prof. Dr. Yaro Burian Junior - DMCSI/FEEC/UNICAMP

Prof. Dr. Jônatas Manzolli - DM/IA/UNICAMP

*Dissertação submetida à Faculdade de
Engenharia Elétrica e de Computação da
Universidade Estadual de Campinas,
para preenchimento dos pré-requisitos
parciais para obtenção do título de
Mestre em Engenharia Elétrica.*

Campinas, 28 de julho de 2005.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP**

Fernandes Júnior, Antônio Carlos Lopes

F391a Auralização: técnicas de modelagem e simulação binaural de ambientes acústicos virtuais/Antônio Carlos Lopes Fernandes Júnior.
– –Campinas, SP: [s.n.], 2005.

Orientador: Furio Damiani.

Dissertação (Mestrado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Psicoacústica. 2. Som. 3. Acústica. 4. Simulação computacional.
5. Acústica arquitetônica. 6. Arquitetura acústica. I. Damiani, Furio.
II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica
e de Computação. III. Título.

Resumo

Auralização é um termo utilizado para significar as técnicas que tornam um campo sonoro audível por modelagem física ou matemática da fonte sonora, do ambiente e do ouvinte. O presente trabalho tem como objetivo principal a análise de processos de modelagem inteiramente computacionais para a constituição de ambientes acústicos virtuais que podem ser implementados e simulados, interativamente, em um computador pessoal e experienciados, perceptualmente, via fones de ouvido. Os conceitos de resposta impulsiva e função de transferência são utilizados para modelar os espaços acústicos e as transformações impelidas pelo torso, cabeça e orelhas ao som que atinge um ouvinte. Esta dissertação dá ênfase às modelagens do ambiente e do ouvinte e ressalta os aspectos mais importantes das abordagens utilizadas para as fontes sonoras.

Abstract

Auralization is a term used to signify the techniques that render audible a sound field through the mathematical or physical modeling of the sound source, the acoustic space and the listener. The present work aims to analyse fully computed process of modeling to the constitution of virtual acoustics environments that can be implemented and interactively simulated in a personal computer and perceptually experienced through headphones. The concepts of impulse response and transfer function are used to model acoustic spaces and the transformations compelled by torso, head and ears to the sound that hits a listener. This dissertation give emphasis to the environment and listener modeling and give some light on the most important aspects of the approaches used to the sound sources.

“Eu tinha vontade de fazer como os dois homens que vi sentados na terra escovando osso. No começo achei que aqueles homens não batiam bem. Porque ficavam sentados na terra o dia inteiro escovando osso. Depois aprendi que aqueles homens eram arqueólogos. E que eles faziam o serviço de escovar osso por amor. E que eles queriam encontrar nos ossos vestígios de antigas civilizações que estariam enterrados por séculos naquele chão. Logo pensei de escovar palavras. Porque eu havia lido em algum lugar que as palavras eram conchas de clamores antigos. Eu queria ir atrás dos clamores antigos que estariam guardados dentro das palavras. Eu já sabia também que as palavras possuem no corpo muitas oralidades remontadas e muitas significâncias remontadas. Eu queria então escovar as palavras para escutar o primeiro esgar de cada uma. Para escutar os primeiros sons, mesmo que ainda bígrafos. Comecei a fazer isso sentado em minha escrivaninha. Passava horas inteiras, dias inteiros fechado no quarto, trancado, a escovar palavras. Logo a turma me perguntou: o que eu fazia o dia inteiro trancado naquele quarto? Eu respondi a eles, meio entresonhado, que eu estava escovando palavras. Eles acharam que eu não batia bem. Então eu joguei a escova fora”.

Manoel de Barros



*Dedico esta dissertação
às minhas meninas, Sara e Clara.*

Agradecimentos

À minha companheira, Sara, pela força, dedicação, carinho e por me impelir a querer ser uma pessoa melhor para estar ao seu lado.

À minha filha, Clara, por ter engrandecido minha humanidade.

Aos meus pais, Antônio Carlos e Sandra, pelos estímulos, pela confiança e sobretudo pelo amor sempre presente.

Aos meus irmãos, Daniel e Felipe, pelo companheirismo e carinho.

À minha segunda mãe, Dadai.

Ao DSIF, à FEEC e à UNICAMP pela acolhida acadêmica.

À CAPES pelo apoio financeiro prestado.

Ao Prof. Dr. Furio Damiani pela oportunidade, confiança e, acima de tudo, pela amizade.

Aos professores da Universidade Federal da Bahia, Antônio César de Castro Lima e Ana Isabela de Araújo Cunha, por me referendarem no prosseguimento de minha vida acadêmica.

Agradecimento especial ao amigo Rogério Prazeres, pelo apoio, amizade e pelas valiosas discussões em relação à dissertação.

Aos amigos e velhos companheiros de vida acadêmica: Alex Sandro Ribeiro, Eudemario Santana e Irênio Júnior, pelo apoio, amizade e por tornarem a minha estadia na FEEC mais familiar.

Aos novos amigos e companheiros: Antonio Marcelo, Hécio Wagner e Teógenes Brito, pelo apoio e por demonstrarem que as mudanças sempre nos trazem boas novidades.

Ao amigo Eduardo Allatta, pelo apoio e pela amizade.

Aos amigos da república: Carla, Fábio, Fernanda, Jaques, Maya, Mir, Nilde, Nilda, Vanessa, Pablo e Welber dentre tantos outros que me proporcionaram a construção de uma família com muitos sobrenomes e muitas origens.

À todos os companheiros e funcionários do dia a dia no DSIF e na FEEC.

Ao companheiro de engenharia de som Christian Herrera, pela valiosa prestatividade de me fornecer artigos preciosos.

Enfim, agradeço a todos que, de uma forma ou de outra, se envolveram neste saboroso processo de crescimento.

Sumário

Resumo	i
Abstract	iii
Agradecimentos	vii
Lista de Figuras	xii
Lista de Tabelas	xv
Lista de Símbolos	xvii
Lista de Abreviaturas	xix
Capítulo 1 Introdução	1
1.1 Motivação	1
1.2 Objetivo	2
1.3 Organização	3
Capítulo 2 Bases Teóricas	5
2.1 Acústica e sua Aplicação em Salas	5
2.1.1 O fenômeno de transmissão sonora	5
2.1.2 Perceptibilidade de reflexões simples	10
2.1.3 A importância da energia lateral inicial	11
2.1.4 Tempo de reverberação e seu efeito subjetivo - RT_{60}, RT ou T_{60}	11
2.2 Psicoacústica	13
2.2.1 Localização	13
2.2.2 Diferença de tempo interaural - ITD	15
2.2.3 Diferença de nível interaural - ILD	16

SUMÁRIO

2.2.4	Reversões frente/costas	17
2.2.5	Localização no interior da cabeça - <i>IHL</i>	17
2.3	Tecnologia Binaural	18
2.3.1	Funções de transferência relacionadas à cabeça - <i>HRTF</i>	18
2.3.2	Audição via fones de ouvido	21
2.3.3	Técnicas binaurais de gravação e reprodução	23
2.4	Processamento Digital de Sinais	26
2.4.1	Resposta ao impulso de um sistema	26
2.4.2	Sistemas PDS simples	28
2.4.3	Convolução e filtragem digital (filtros <i>FIR</i> e <i>IIR</i>)	29
2.4.4	Sistemas fase-mínima	31
Capítulo 3	Auralização	33
3.1	Definição – O Que é a Auralização?	33
3.2	Histórico	34
3.3	Técnicas Atuais de Auralização	35
3.3.1	Auralização inteiramente computacional	37
3.3.2	Auralização computacional com múltiplos alto-falantes	42
3.3.3	Auralização indireta com modelo acústico em escala	43
3.3.4	Auralização direta com modelo acústico em escala	46
Capítulo 4	Auralização Inteiramente Computacional	49
4.1	Estrutura da <i>RIR</i>	49
4.1.1	Decomposição da <i>RIR</i> em partes	50
4.1.2	Parâmetros acústicos extraídos da <i>RIR</i>	52
4.2	Métodos para o Cálculo da <i>RIR</i>	57
4.2.1	Métodos baseados na decomposição do campo sonoro	57
4.2.2	Métodos baseados na equação de onda	74
4.2.3	Métodos estatísticos	78
4.3	Comparação entre os métodos e hibridização	79
4.4	Influência da Absorção, Espalhamento e Difração	81
4.5	Cálculo da <i>BRIR</i>	83
4.6	Convolução para Auralização	85
4.7	Reprodução Binaural do Produto da Auralização	86
4.7.1	Transaural	86
4.7.2	Binaural	89

SUMÁRIO

Capítulo 5	Modelagem da Fonte e do Ouvinte	91
5.1	Fonte	91
5.1.1	Natureza da fonte	92
5.1.2	Geração do sinal sonoro	92
5.1.3	Direcionalidade	93
5.2	Ouvinte	94
5.2.1	Medidas de respostas impulsivas binaurais (HRIR)	94
5.2.2	HRTF generalizada	97
5.2.3	Processamento e equalização	98
5.2.4	Modelagem com filtros digitais	100
Capítulo 6	Considerações Finais	103
6.1	Conclusão	103
6.2	Trabalhos Futuros	104
	Referências Bibliográficas	106

Lista de Figuras

2.1	<i>Sistema referencial relacionado à cabeça</i>	14
2.2	<i>Diferença direcional entre evento auditivo (A_i) e sonoro (S_i) no PH</i> .	15
2.3	<i>Caminho das direções do evento auditivo em função da frequência central do sinal ruidoso de banda estreita no PM</i>	15
2.4	<i>Diferença de tempo interaural</i>	16
2.5	<i>Ambiguidade na diferença de tempo interaural</i>	16
2.6	<i>Circuito análogo à transmissão sonora na orelha.</i>	20
2.7	<i>Circuito análogo à transmissão sonora na orelha via fones de ouvido</i> .	21
2.8	<i>Medida com microfone em ponto genérico dentro da orelha externa</i> .	25
2.9	<i>SLIT</i>	27
2.10	<i>Multiplicação de uma sequência por uma constante. No exemplo acima, $g=2$.</i>	28
2.11	<i>Atraso de uma sequência por uma amostra.</i>	29
2.12	<i>Adição de duas sequências</i>	29
2.13	<i>Convolução</i>	30
2.14	<i>Filtro de resposta ao impulso finita (FIR)</i>	30
2.15	<i>Filtro de resposta ao impulso infinita (IIR)</i>	31
3.1	<i>Cadeia de Auralização</i>	37
3.2	<i>Auralização Computacional Completa</i>	38
4.1	<i>Evolução temporal de uma resposta impulsiva de ambientes</i>	50
4.2	<i>Conexão de dois sinais em transição de zero não-direcional</i>	53
4.3	<i>traçado de raios.</i>	59
4.4	<i>Deteções múltiplas</i>	60
4.5	<i>Duas condições de detecção diferenciadas devido às dimensões do receptor e a densidade dos raios</i>	61
4.6	<i>Exemplo de detecção inválida: receptor ocluso por uma parede.</i>	61

LISTA DE FIGURAS

4.7	<i>Construção de uma imagem da fonte.</i>	63
4.8	<i>Histórico de um raio sonoro a partir da construção de suas imagens da fonte.</i>	64
4.9	<i>Constelação de imagens da fonte para ambiente em formato de “caixa de sapato”.</i>	65
4.10	<i>Teste de visibilidade: traçado reverso</i>	67
4.11	<i>Reflexão irreal: não há como ocorrer uma reflexão entre as paredes 1 e 3. A imagem $F_{1,3}$ é irreal.</i>	68
4.12	<i>Estrutura em árvore de imagens da fonte possíveis para uma sala com N paredes.</i>	68
4.13	<i>Traçado de pirâmides: fonte sonora F espelhada na superfície a resultando na imagem F_a</i>	70
4.14	<i>Receptor R irradiado incorretamente pela imagem F_1. A área em cinza representa a região do espaço que está sendo irradiada incorretamente.</i>	71
4.15	<i>RA de uma esquina.</i>	72
4.16	<i>Geometria da densidade de radiação de um elemento de área dS com posição definida pelo vetor r.</i>	73
4.17	<i>Geometria da transferência de energia para o receptor.</i>	73
4.18	<i>FEM (discretização do volume do ambiente)</i>	75
4.19	<i>BEM (discretização das superfícies do ambiente)</i>	76
4.20	<i>DWM (discretização do volume do ambiente em representação bidimensional)</i>	77
4.21	<i>Transmissão do sinal binaural nos terminais do alto-falante até os canais auditivos</i>	87
5.1	<i>Diagrama de blocos da modelagem física de instrumentos de corda “puxada”.</i>	93
5.2	<i>Medida de HRTF</i>	95
5.3	<i>Diagrama de blocos: transmissão do alto-falante, nas medições, até a captação no tímpano.</i>	99

Lista de Tabelas

3.1	Critérios para os modelos físicos em escala. n é o fator de escala que relaciona o ambiente original (subscrito o) e o modelado (subscrito m).	44
4.1	Parâmetros acústicos derivados da <i>RIR</i>	53

Lista de Símbolos

A	área de superfície refletora
\bar{a}	índice médio de absorção
b	fase da <i>HRTF</i>
$B(r, t)$	densidade de radiação
C	<i>clareza</i>
D	<i>definição</i>
$e(t)$	tensão nos terminais do alto-falante
E_{af}	espectro da tensão nos terminais do alto-falante
E_{fone}	espectro da tensão nos terminais do fone de ouvido
E_{mic}	espectro da tensão nos terminais do microfone
F	fonte sonora
G	nível sonoro relativo
G_a	ganho do equalizador
$h(t)$	resposta impulsiva
$h_{d\varphi\beta}$	resposta impulsiva binaural da orelha direita
$h_{e\varphi\beta}$	resposta impulsiva binaural da orelha esquerda
I	intensidade
I_0	intensidade do raio incidente
j	$\sqrt{-1}$
M	função de transferência do microfone
$p_1(t)$	pressão sonora no centro da cabeça sem o ouvinte
$P_1(\omega)$ ou P_1	espectro da pressão sonora no centro da cabeça
$P_2(\omega)$	espectro da pressão sonora na entrada do canal auditivo bloqueado
$P_3(\omega)$	espectro da pressão sonora na entrada do canal auditivo aberto
$P_4(\omega)$	espectro da pressão sonora no tímpano
$P_5(\omega)$	espectro da pressão sonora na entrada do canal auditivo bloqueado com uso de fones de ouvido
$P_6(\omega)$	espectro da pressão sonora na entrada do canal auditivo aberto com uso de fones de ouvido
$P_7(\omega)$	espectro da pressão sonora no tímpano com uso de fones de ouvido
$Q(t)$	velocidade de volume
R	receptor
\mathbf{R}	coeficiente complexo de reflexão

LISTA DE SÍMBOLOS

Rr	distância entre elemento e receptor
rf	distância ouvinte/fonte
r_r	posição do receptor
RT_{60} ou T_{60}	tempo de reverberação
S	área total das superfícies
T_{10}	tempo de reverberação para queda de $10dB$
T_S	tempo central
\mathbf{u}_n	velocidade de partícula normal à superfície
V	volume da sala
$x(t)$	entrada de um <i>SLIT</i>
$y(t)$	saída de um <i>SLIT</i>
$Z_{\text{canal auditivo}}$	impedância acústica do canal auditivo
Z_{fone}	impedância acústica do fone de ouvido vista a partir da entrada do canal auditivo
$Z_{\text{radiação}}$	impedância externa vista a partir da entrada do canal auditivo
$Z_{\text{tímpano}}$	impedância do tímpano
α	coeficiente de absorção
β	ângulo de elevação
ΔL	diferença de nível em decibéis
δ	coeficiente de espalhamento
$\delta(t)$	<i>Delta de Dirac</i>
ε	densidade de energia
ϑ	ângulo do raio refletido
ϑ_0	ângulo do raio incidente
λ	comprimento de onda
τ_{gr}	<i>atraso</i> de grupo
φ	ângulo de azimute
ψ	fase do coeficiente complexo de reflexão
ω	frequência angular

Lista de Abreviaturas

3-D tridimensional

AR Acoustic Radiosity; Radiosidade Acústica

ASW Apparent Source Width; Largura Aparente da Fonte

BEM Boundary Element Method; Método de Elementos de Contorno

BRIR Binaural Room Impulse Response; Resposta Impulsiva Binaural do Ambiente

DIVA Digital Interactive Virtual Acoustics

DWM Digital Waveguide Mesh

FDTD Finite-Difference Time-Domain; Diferenças Finitas no Domínio do Tempo

FIR Finite Impulse Response; Resposta Impulsiva Infinita

FEC Free-Air Equivalent Coupling

FEEC Faculdade de Engenharia Elétrica e de Computação

FEM Finite Element Method; Método de Elementos Finitos

HRIR Head-Related Impulse Response; Resposta Impulsiva Relacionada à Cabeça

HRTF Head-Related Transfer Function; Função de Transferência Relacionada à Cabeça

IACC Interaural Cross-Correlation; Correlação Cruzada Interaural

IACF Interaural Cross-Correlation Function; Função de Correlação Interaural

IHL Inside the Head Locatedness; Localização no Interior da Cabeça

IID Interaural Intensity Difference; Diferença de Intensidade Interaural

IIR Infinite Impulse Response; Resposta Impulsiva Infinita

ILD Interaural Level Difference; Diferença de Nível Interaural

ISM Image Source Method; Método das Imagens da Fonte

ITD Interaural Time Difference; Diferença de Tempo Interaural

ITF Interaural Transfer Function; Função de Transferência Interaural

LEF Lateral Energy Fraction; Fração de Energia Lateral

LEV Listener Involvement; Envolvimento do Ouvinte

MTF Monaural Transfer Function; Função de Transferência Monaural

PF Plano Frontal

LISTA DE ABREVIATURAS

PH Plano Horizontal

PM Plano Médio

RIR Room Impulse Response; Resposta Impulsiva do Ambiente

RTM Ray Tracing Method; Método do Traçado de Raios

CAPÍTULO 1

Introdução

1.1 Motivação

A importância da acústica na vida do ser humano advém do fato de que esta trata, no escopo de seus assuntos, das interações do som com o ambiente no qual o homem o percebe. Devido à estreita ligação existente entre este ramo da Física e alguns dos produtos da expressão humana, tais como a fala e a música [1, 2], fica patente a necessidade de estudos relacionados a este tema. Salas de aula, estúdios de gravação, salas de audição sonora como teatros, cinemas e muitos outros ambientes, onde as características acústicas devem ser especificadas de acordo com a utilização do recinto, são alvo de projetos acústicos, ou pelo menos assim deveriam ser, para que haja qualificação da informação que se pretende transmitir. O ruído também figura entre os objetivos de estudo e projeto da acústica [3, 4]. Ambientes de trabalho, tais como os industriais, costumam ser altamente ruidosos e merecem atenção, ao serem projetados, para que a salubridade dos que nele trabalham seja garantida.

Tendo em vista toda a importância do resultado auditivo das interações do produto sonoro com o espaço que ele ocupa, surge a necessidade de predição do campo sonoro resultante em determinado ambiente e de como este será percebido pelo sistema auditivo. É do reconhecimento desta necessidade que emerge como ferramenta poderosa a *Auralização*. Esta técnica, permite que um determinado som emitido em uma sala seja escutado sem que se esteja em seu interior e sequer que este espaço exista [5]. Com este recurso, pode-se, não somente, atender toda a demanda por qualidade na confecção da ambiência desejada, como também, fortalecer diversas frentes, como informação, entretenimento e educação [5, 6]. Esta última teria na *auralização* um amparo substantivo para que as teorias a respeito do fenômeno ondulatório em questão fossem exploradas de tal forma que futuros profissionais relacionados ao áudio e à acústica, em processo de aprendizado, pudessem absorver de

1. Introdução

forma mais eficiente os conceitos estudados. Os músicos podem ser beneficiados com informações importantes já que estes são usuários primordiais dos recintos de execução e gravação sonora e, em geral, utilizam-se de sua avaliação subjetiva para obter o resultado esperado em determinada performance. Ainda na área de educação, existe a possibilidade de treinamento para pessoas com deficiência visual, já que uma correta interpretação das pistas auditivas para a localização espacial, por parte destes indivíduos, se faz necessária.

As pesquisas da psicoacústica, também, podem ser aperfeiçoadas por intermédio desta ferramenta [5, 7], pois há possibilidades de testes de como os parâmetros acústicos e dados espaciais da fonte sonora podem influenciar na percepção auditiva. No que se refere à informação e entretenimento, pode-se ampliar a qualidade de videoconferências, jogos multiusuário, e programas de realidade virtual [6]. As possibilidades na simulação de ambientes são enormes. A *Auralização computacional* [4, 5, 7] permite que o som emitido por uma ou mais fontes, seja escutado sob diversas circunstâncias de forma interativa isto é, pode-se, por exemplo, simular o posicionamento da fonte em diversos pontos do ambiente, ou mesmo simular os movimentos corpóreos do ouvinte, ou até a alteração das características físicas do próprio recinto, como, por exemplo, a absorção e formas geométricas das paredes.

Como suportes teóricos deste processo encontram-se algumas disciplinas importantes [8]. A acústica de ambientes que fornece as bases para avaliação dos fenômenos físicos que ocorrem na transmissão sonora. A psicoacústica que elucidada de que forma o sistema auditivo percebe o som. E, por fim, o processamento digital de sinais que serve como ferramenta para modelar, em ambiente computacional, as transformações advindas das teorias anteriores. Baseado no fato de ser uma ferramenta multidisciplinar que lida com a percepção humana, o interesse por este tema já é, por si só, justificado.

1.2 Objetivo

O presente trabalho tem como objetivo o estudo e a discussão sobre o processo de criação de ambientes acústicos virtuais para *simulação binaural*. Assim, os três alvos da modelagem (a fonte sonora, o ambiente e o ouvinte) são discutidos e avaliados visando a escolha das melhores abordagens para implementação de um sistema de *auralização*. As aplicações desta simulação binaural, sejam elas pesquisas na área da acústica e da psicoacústica, projetos de ambientes, ou em sistemas que se utilizem da realidade virtual, são avaliadas em relação a parâmetros como fidedignidade e

interatividade. Assim, cada processo de modelagem, principalmente o do espaço acústico, é analisado de forma a indicar a melhor composição para a escolha das técnicas mais apropriadas a determinada aplicação. Ainda, uma indicação sobre abordagem de modelagem para *auralização* que possa contemplar simulações em tempo real é realizada.

1.3 **Organização**

Esta dissertação é composta por seis capítulos. No capítulo 2, são apresentados os principais conceitos que dão suporte teórico aos demais capítulos. O capítulo 3, traz um histórico a respeito do tema e explana sobre as técnicas atuais de *auralização*. No capítulo 4, a *auralização inteiramente computacional* é discutida e seus principais parâmetros são avaliados. O foco principal deste capítulo é a *resposta impulsiva do ambiente* e as estratégias mais utilizadas para a sua determinação. O capítulo 5, refere-se às modelagens da fonte e do ouvinte. No capítulo 6, realizam-se as conclusões que indicam a possibilidade de um sistema de *auralização* interativo. Ainda neste capítulo, é delineada a direção para pesquisas futuras e implementação de tal sistema de *auralização*.

CAPÍTULO 2

Bases Teóricas

Neste capítulo as principais disciplinas que compõem as bases teóricas do processo de *auralização* são apresentadas e seus conceitos mais importantes são revisados. Figuram como alicerces teóricos a acústica de salas, a psicoacústica, a tecnologia binaural e o processamento digital de sinais.

2.1 Acústica e sua Aplicação em Salas

A Acústica pode ser definida como a geração, transmissão, e recepção de energia na forma de ondas vibracionais na matéria [3]. Quando átomos ou moléculas de um fluido ou de um sólido sofrem deslocamento, em relação à sua posição inicial, surge uma força elástica que tende a restaurar as configurações originais. É esta força restauradora que, conjuntamente com a inércia do sistema, promove o movimento de oscilação que está relacionado à geração e transmissão das ondas acústicas.

O som é uma sensação que está associada a alguns fenômenos acústicos. Quando ondas acústicas apresentam frequência de vibração na faixa entre 20 a 20.000 Hz estas são percebidas pelo sistema auditivo dos seres humanos como som. As frequências abaixo de 20 Hz são denominadas de *infrasônicas* e as frequências acima de 20 KHz são chamadas de *ultrasônicas*. É importante salientar que os outros animais também percebem vibrações mecânicas como som, porém, cada espécie com suas características próprias de faixa de frequências perceptíveis.

2.1.1 O fenômeno de transmissão sonora

Ondas acústicas geradas em fluidos com viscosidade pequena são ondas longitudinais ou seja as moléculas se movem para frente e para trás na direção de propagação da onda provocando regiões de compressão e rarefação. Este é o caso do fenômeno acústico que ocorre no ar e que figura como uma das mais importantes situações no estudo de ambientes acústicos.

2. Bases Teóricas

Das equações que governam este fenômeno acústico [3], *equação de estado*, de *continuidade* e de *força*, chega-se à *equação de onda* (2.1):

$$\nabla^2 p = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (2.1)$$

onde p é a pressão acústica em qualquer ponto, que significa a diferença entre a pressão instantânea em qualquer ponto e a pressão de equilíbrio do fluido. A grandeza $c = (331.4 + 0.6\Theta)$ é a velocidade do som no ar em função temperatura (Θ), em graus centígrados, considerando que o meio seja homogêneo e esteja em repouso. A *equação de onda* (2.1) descreve a propagação de ondas sonoras em qualquer meio no qual não haja perdas.

Se restringirmos a propagação a uma única direção, por exemplo, na direção do eixo cartesiano x , obtêm-se a equação de *onda plana* (2.2):

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \quad (2.2)$$

Em uma *onda plana*, qualquer plano perpendicular à direção de propagação, neste caso ao eixo x , apresenta fase constante para qualquer variável. A solução harmônica complexa para a pressão de tal onda, será (2.3):

$$\mathbf{p} = \mathbf{A} \cdot e^{j(\omega t - kx)} + \mathbf{B} \cdot e^{j(\omega t + kx)} \quad (2.3)$$

onde \mathbf{A} e \mathbf{B} são constantes complexas de amplitude e $k = \omega/c$ é o *número de onda*.

Da *equação de força* (2.4),

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} = -\nabla \mathbf{p} \quad (2.4)$$

chega-se à velocidade de uma partícula do ar ($\vec{\mathbf{u}}$) (2.5):

$$\mathbf{u} = \left[\frac{\mathbf{A}}{\rho_0 c} e^{j(\omega t - kx)} - \frac{\mathbf{B}}{\rho_0 c} e^{j(\omega t + kx)} \right] \quad (2.5)$$

onde ρ_0 é a *densidade de equilíbrio* do fluido.

Se considerarmos uma onda plana que se propaga no sentido positivo do eixo x as equações (2.3) e (2.5) são simplificadas (2.6), (2.7):

$$\mathbf{p}_+ = \mathbf{A} e^{j(\omega t - kx)} \quad (2.6)$$

$$\mathbf{u}_+ = \frac{\mathbf{p}_+}{\rho_0 c} \quad (2.7)$$

Da equação (2.7), nota-se que a razão entre a *pressão sonora* e a *velocidade de partícula* em uma *onda plana* independe da frequência e é real, ou seja, \mathbf{p}_+ e \mathbf{u}_+ estão em fase. Essa razão é denominada de *impedância característica* do meio e para o ar assume o valor $\rho_0 c = 414 \text{ Kg m}^{-2} \text{ s}^{-1}$. Se a onda se propaga no sentido negativo do eixo x a *impedância característica* terá um valor negativo.

Imaginemos uma superfície de área unitária perpendicular a direção de propagação. A energia sonora que flui por segundo através desta superfície é denominada de *intensidade* da onda [3,4] e é calculada pela média temporal do produto da *pressão sonora* e da *velocidade de partícula* (2.8):

$$I = \overline{\mathbf{p}\mathbf{u}} = \frac{\overline{\mathbf{p}^2}}{\rho_0 c} \quad (2.8)$$

Assim, a *densidade de energia* da onda sonora será:

$$\varepsilon = \frac{I}{c} = \frac{\overline{\mathbf{p}^2}}{\rho_0 c^2} \quad (2.9)$$

Para vibrações harmônicas, teremos:

$$I = \frac{\hat{p}^2}{2\rho_0 c} = \frac{\tilde{p}^2}{\rho_0 c} \quad (2.10)$$

$$\varepsilon = \frac{\hat{p}^2}{2\rho_0 c^2} = \frac{\tilde{p}^2}{\rho_0 c^2} \quad (2.11)$$

onde \hat{p} é a amplitude e \tilde{p} é o valor eficaz da pressão sonora.

Para além da *onda plana*, é de extrema importância o conhecimento das propriedades das *ondas esféricas*. Assim, reescrevendo a equação (2.1) em coordenadas esféricas, obtém-se:

$$\frac{\partial^2 \mathbf{p}}{\partial r^2} + \frac{2}{r} \frac{\partial \mathbf{p}}{\partial r} = \frac{1}{c^2} \frac{\partial^2 \mathbf{p}}{\partial t^2} \quad (2.12)$$

assim, as quantidades acústicas dependem da distância r em relação à origem e não mais da direção.

Se considerarmos mais uma vez a harmonicidade e avaliarmos a equação da *onda esférica* para obter a solução de uma onda que diverge a partir da origem, teremos:

2. Bases Teóricas

$$\mathbf{p} = \frac{\mathbf{C}}{r} e^{j(\omega t - kr)} \quad (2.13)$$

onde \mathbf{C} é uma constante arbitrária. Novamente calculando a *velocidade de partícula* (2.14) a partir das equações (2.4) e (2.13), teremos:

$$\mathbf{u}_r = \frac{\mathbf{p}}{\rho_0 c} \left(1 - \frac{j}{kr}\right) \quad (2.14)$$

Da equação (2.14) percebe-se que a razão entre a *pressão sonora* e a *velocidade de partícula* em uma onda esférica assume um valor complexo o que indica a diferença de fase entre as duas quantidades acústicas. Esta razão é denominada de *impedância acústica específica*. Para $kr \gg 1$, ou seja, distâncias muito maiores do que o *comprimento de onda* em questão a razão $\mathbf{p}/\tilde{\mathbf{u}}$ tende à *impedância característica* ($\rho_0 c$). Se esta condição for satisfeita, as relações de *intensidade* e *densidade de energia* para uma *onda plana* podem ser utilizadas. Assim, pode-se calcular a energia que a onda transmite em 1s:

$$P = 4\pi r^2 I = \frac{2\pi |C|^2}{\rho_0 c} \quad (2.15)$$

Por conservação de energia, percebe-se que a *potência acústica* (2.15) deve ser fornecida pela fonte sonora, na origem.

Se caso contrário $kr \ll 1$, ou seja, distâncias inferiores ao *comprimento de onda*, a *velocidade de partícula* se reduz a seguinte relação (2.16):

$$\mathbf{u}_r = \frac{\mathbf{p}}{jkr\rho_0 c} = \frac{C}{jkr^2\rho_0 c} e^{j\omega t} \quad (2.16)$$

Multiplicando esse valor limite da *velocidade de partícula* pela área da superfície esférica com ($r \ll 1/k$), obtém-se outra importante quantidade acústica, a *velocidade de volume* da fonte sonora. Esta representa o volume que é expelido da origem em 1s (2.17):

$$Q(t) = \frac{4\pi C}{jk\rho_0 c} e^{j\omega t} \quad (2.17)$$

Existe uma formulação importante da equação de onda que considera que a pressão sonora varia harmonicamente com o tempo. Se retomarmos a equação de onda (2.1) e considerarmos que $\mathbf{p} = \Upsilon \cdot e^{j\omega t}$, obteremos a seguinte equação (2.18):

$$\nabla^2 \Upsilon + k^2 \Upsilon = 0 \quad (2.18)$$

onde Υ é função da posição. Essa equação (2.18) é conhecida como equação de *Helmholtz* ou *equação de onda independente do tempo* [3].

Nível de pressão sonora.

Devido às inúmeras ordens de magnitude que separam as intensidades dos limites de percepção sonora e o limiar da dor, a unidade utilizada para representar a *pressão sonora* e a *intensidade* é o *decibel (dB)*. Assim, utilizam-se as seguintes relações logarítmicas (2.19), (2.20):

$$SPL = 20 \log(\tilde{p}/\tilde{p}_0) \quad (2.19)$$

$$IL = 10 \log(I/I_0) \quad (2.20)$$

onde *SPL* (2.19) refere-se ao *nível de pressão sonora* e é relativo à pressão de referência ($\tilde{p}_0 = 2 \cdot 10^{-5} N/m^2$) mínima para a percepção sonora de uma frequência de 1 *KHz*. O *IL* é o *nível de intensidade* em relação à *intensidade* de referência ($I_0 = 10^{-12} W/m^2$).

Coefficiente de reflexão e absorção.

Até agora, estávamos discutindo sobre a propagação sonora em um meio ilimitado. É necessário, também, discutir alguns fenômenos que ocorrem quando uma onda sonora atinge uma superfície. Em geral, quando uma *onda plana* colide com uma parede, parte da energia sonora será refletida por esta superfície com amplitude e fase diferentes da onda incidente. A quantidade acústica que representa este fenômeno é o *coeficiente de reflexão* complexo (2.21):

$$\mathbf{R} = |R|e^{j\psi} \quad (2.21)$$

Esta quantidade é uma propriedade da superfície e tanto o seu valor absoluto quanto o seu ângulo de fase dependem da frequência e do ângulo de incidência (ângulo entre a normal da superfície e a direção de propagação da onda incidente).

Após uma reflexão a intensidade da onda refletida é menor que a da onda incidente por um fator $|R|^2$. Assim, a energia perdida é dada pelo *coeficiente de absorção* (α) da superfície (2.22):

$$\alpha = 1 - |R|^2 \quad (2.22)$$

Paredes cujo $\mathbf{R} = 0$, absorvem totalmente a energia incidente. Se $\mathbf{R} = 1$, a onda refletida estará em fase com a incidente e esta parede é denominada de *parede rígida*. Para $\mathbf{R} = -1$, teremos uma reflexão defasada de 180 e a parede é dita *suave*. A partir de agora, o termo reflexão estará relacionado tanto ao fenômeno quanto à onda que resulta deste.

O *coeficiente de reflexão* descreve completamente as propriedades de uma parede. Porém, outra quantidade importante que também representa fisicamente uma parede é a sua impedância que é definida como:

$$\mathbf{Z} = \frac{\mathbf{p}}{\mathbf{u}_n} \quad (2.23)$$

onde (\mathbf{u}_n) é a *velocidade de partícula*, normal à superfície, gerada devido a presença de uma pressão (\mathbf{p}) .

2.1.2 Perceptibilidade de reflexões simples

Quando o sistema auditivo é exposto a um determinado sinal acústico vindo direto de uma fonte sonora (*som direto*) e a uma reflexão deste sinal podem ocorrer duas situações. Na primeira delas a reflexão pode ser ouvida como um sinal separado e, assim, teremos o fenômeno do *eco*. Na segunda situação esta reflexão não será percebida como tal, mas afetará o *som direto* reforçando-o, mudando a sua composição espectral, ou aumentando a *largura aparente da fonte (ASW)* [4].

O sistema auditivo dos seres humanos possui resolução temporal limitada em relação à eventos sonoros sucessivos. A percepção de uma reflexão como um *eco* depende não somente desta resolução mas, também, da intensidade relativa da reflexão, sua direção de incidência no ouvinte, do tipo de sinal e da presença de outras reflexões.

As relações entre *atraso* (diferença de tempo entre os sinais) e nível relativo entre o *som direto* e uma reflexão, são parâmetros que podem determinar a perceptibilidade ou não desta reflexão como um evento sonoro separado. Com o aumento do *atraso*, o nível necessário para que esta reflexão seja percebida diminui. Neste contexto, vale a pena ressaltar que o tipo do sinal também influi nesta relação *atraso versus* nível relativo. O sistema auditivo é menos sensível às reflexões de sinais musicais do que às reflexões de sinais de fala devido às diferenças de composição espectral destes sinais.

As direções de incidência também são determinantes para o limiar de perceptibilidade de *ecos*. Se considerarmos que o *som direto* chega de uma direção frontal ao ouvinte reflexões que venham desta mesma direção serão mascaradas muito mais facilmente do que aquelas com incidência lateral.

Essas relações de níveis e *atrasos* são de grande importância para ambientes acústicos destinados à fala e à música. Nestes locais, os *ecos* podem ser desastrosos, tanto no que tange à qualidade da informação transmitida quanto na *localização* da fonte sonora. Em relação a esta última, sabe-se, porém, que a *lei da primeira frente de onda* garante que o *som direto* é o que determina a direção da fonte sonora, mesmo em situações nas quais a reflexão apresenta níveis de intensidade 10.5 dB acima do *som direto*, desde que esteja dentro de uma faixa de *atraso* apropriada [4].

2.1.3 A importância da energia lateral inicial

Um dos efeitos subjetivos mais importantes que se apresentam em ambientes acústicos fechados é a sensação de espaço. Ela é causada devido ao fato de que os sinais acústicos atingem o ouvinte advindos de várias direções e o sistema auditivo, por não estar apto a determinar cada uma destas direções, realiza um processamento geral resultando numa sensação de espaço, denominada de *espacialidade* [4].

Ao contrário do que se pode imaginar, a espacialidade não é causada pela *reverberação*. Na verdade essa sensação de espaço pode ser conseguida por intermédio de poucas reflexões desde que estas satisfaçam às seguintes condições:

- As reflexões devem ser mutuamente incoerentes.
- Suas intensidades devem exceder um determinado limiar.
- Seus *atrasos* com respeito ao *som direto* não devem exceder 100ms (são denominadas de *reflexões iniciais*).
- As direções das reflexões devem ser laterais com relação ao ouvinte.

As reflexões laterais numa sala são mutuamente incoerentes, inclusive em relação ao *som direto*, devido aos seus tempos de chegada e faixa espectral.

2.1.4 Tempo de reverberação e seu efeito subjetivo - RT_{60} , RT ou T_{60}

O *tempo de reverberação* (*tempo de decaimento*) foi definido por *Sabine* como sendo o tempo necessário para que a pressão sonora em um ambiente sofra uma queda de 60 dB após a interrupção da emissão sonora [3,4]. Por intermédio de estudos experimentais *Sabine* chegou a conclusão de que a *reverberação* de um ambiente está relacionada ao seu volume e as características de absorção de suas superfícies (2.24):

2. Bases Teóricas

$$RT_{60} = \frac{0.161V}{S\bar{a}} \quad (2.24)$$

onde V é o volume do ambiente, S é a área total e \bar{a} é o *índice médio de absorção* que é dado por,

$$\bar{a} = \frac{1}{S} \sum_i S_i a_i$$

onde a_i é o índice de absorção de cada superfície i de área S_i . A equação do RT_{60} (2.24) assume que o campo sonoro é difuso, ou seja, que a *densidade média de energia* é a mesma para todo o volume da sala e todas as direções de propagação sonora são equiprováveis. Devido a esta consideração é preciso estar atento aos casos em que se pretende aplicar a fórmula de *Sabine*. Para que a condição de campo difuso seja atendida a sala deve possuir dimensões equilibradas para que não haja reflexões muito tardias. A absorção também deve estar bem distribuída entre as paredes do ambiente. Quando estas condições são atendidas o decaimento energético será ideal e seguirá uma lei exponencial [4].

Existem outras fórmulas para o cálculo do *tempo de reverberação* como a de *Eyring*, [3] que corrige deficiências da fórmula de *Sabine*. Porém, para o que se pretende discutir neste trabalho, os conceitos aqui apresentados são satisfatórios.

Algumas considerações subjetivas a respeito da *reverberação* permitem um entendimento melhor da existência da relação de compromisso entre RT_{60} e o tipo de informação que se pretende transmitir em um recinto [4]. Ambientes destinados principalmente à fala, como por exemplo teatros, auditórios e salas de aula, em princípio, deveriam ter $RT = 0$, para que as sílabas das palavras não se sobreponham. Porém, o *tempo de reverberação* é diretamente proporcional ao ganho que o ambiente promove no sinal acústico em comparação à sua emissão em *campo livre* (ausência de reflexões). Assim, um RT adequado pode facilitar a comunicação da informação. Para além disso, as salas ditas *secas* promovem uma sensação de desconforto para o ouvinte devido à expectativa do sistema auditivo da presença de alguma *reverberação* que indique o tamanho do ambiente. Os requerimentos de RT para espaços destinados à execução musical são diferentes dos destinados à fala. Não há necessidade de detalhamento do sinal acústico musical para que a informação seja compreendida. Na verdade, um *tempo de decaimento* adequado incrementa o *volume* (em itálico significará o *loudness*) e assegura certa continuidade à linha musical de certos estilos como por exemplo a música de câmara.

2.2 Psicoacústica

A acústica trata dos aspectos físicos do som e de suas interações com o ambiente no qual ele é gerado. Quando falamos dos efeitos subjetivos que o fenômeno físico sonoro impele ao sistema auditivo do seres humanos estamos nos referindo à psicoacústica. Este ramo da Física, trata das relações existentes entre um *evento sonoro*, o som fisicamente falando, e um *evento auditivo*, a sensação subjetiva, as imagens sonoras.

Apesar de não estarem necessariamente interconectados em todas as situações, tendo em vista o *tinnitus* que ocorre devido a algumas doenças auditivas sem que nenhum som esteja sendo gerado em termos mecânicos, *eventos sonoros* e *eventos auditivos* são avaliados conjuntamente para que sejam formadas conexões entre os atributos de um campo sonoro e as especificidades de uma “imagem sonora”, como por exemplo sua posição e extensão espacial [2]. É desta tentativa de interconexão, que surge o termo *audição espacial* que, apesar de soar redundante, significa a interpretação em relação à distância, direção e extensão que o sistema auditivo fará das características de determinado *evento sonoro*.

2.2.1 Localização

A *localização* é a lei pela qual a posição de um *evento auditivo* está relacionada aos atributos de um *evento sonoro* ou outro evento que esteja correlacionado ao *evento auditivo* [2]. Como exemplo, podemos citar a relação entre a posição de uma fonte sonora e a posição do *evento auditivo*, que não necessariamente coincidem.

O *erro de localização* (*localization blur*) é definido como a menor variação em algum atributo do *evento sonoro* que provoca uma mudança perceptiva do *evento auditivo* [2]. Um dos casos se refere às mudanças de posição do *evento sonoro*, que sejam notáveis, causadas devido a variações do sinal sonoro. Este erro indica que o *espaço auditivo* é menos detalhado que o espaço onde a fonte sonora opera, ou seja, o sistema auditivo possui resolução espacial limitada. O *erro de localização* depende do tipo de sinal, dos atributos que sofrem alterações neste sinal e, no caso de movimentos da fonte, da região de mudança da posição do emissor sonoro em relação ao ouvinte.

Para iniciar as análises de alguns aspectos destas correlações é necessário definir um sistema referencial relacionado à cabeça (fig.2.1) que será utilizado neste e nos demais capítulos. A origem do sistema está localizada no meio do segmento que conecta os limites superiores das entradas dos canais auditivos. Estes em conjunto com os limites inferiores das cavidades dos olhos definem o *plano horizontal (PH)*,

2. Bases Teóricas

onde o *azimute* (φ) pode ser definido como sendo negativo para o lado direito. O *plano frontal* (PF) intercepta estes pontos às entradas dos canais e é perpendicular ao PH . O *plano médio* (PM) é ortogonal aos anteriores, passa pela origem e divide, idealmente, a cabeça em duas partes simétricas. No PM , define-se o *ângulo de elevação* (β), com valores positivos acima do PH . O símbolo r , denota a distância da fonte à origem do sistema.

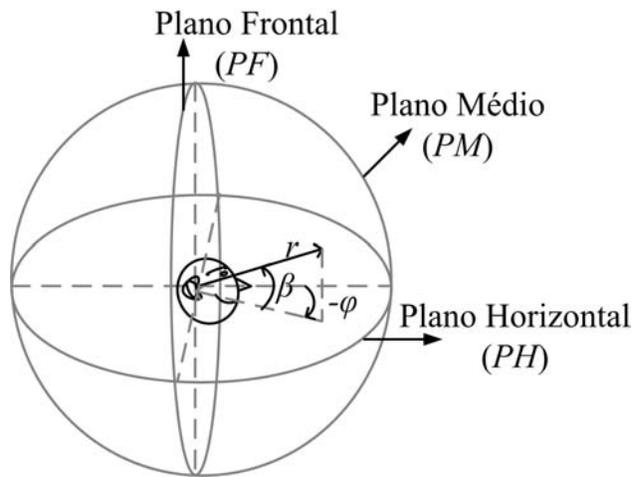


Figura 2.1: Sistema referencial relacionado à cabeça

As relações de distância (*distance hearing*) e direção (*directional hearing*) entre os eventos sonoro e auditivo, variam de acordo com diversos atributos do sinal sonoro [2]. A *localização* direcional de *eventos sonoros* no plano horizontal possui dependência em relação ao conteúdo espectral do sinal. As direções do *evento sonoro* e do *evento auditivo*, em geral, coincidem quando os sinais apresentam banda de frequência larga. O oposto ocorre para sinais de banda estreita. No PH ocorre uma anomalia frequente para tons puros e sinais de faixa estreita. O *evento auditivo* aparecerá numa direção simétrica ao do *evento sonoro* e não na mesma direção de incidência em relação ao eixo das orelhas (fig. 2.2). No PH o *erro de localização* atinge seu limite mínimo para direções frontais ao ouvinte ($\varphi = 0$). Nesta região, para sinais senoidais, a resolução máxima é de 1 *grau* [2]. Com ângulos azimutais maiores, afastamento da direção frontal, o *erro de localização* aumenta.

A *audição direcional* (*directional hearing*) no plano médio (PM) também possui especificidades que variam de acordo com o espectro do campo sonoro que atinge o ouvinte. Para sinais com largura de banda inferior a $2/3$ de oitava a *localização* do *evento auditivo* independe da direção de incidência e passa a depender da frequência

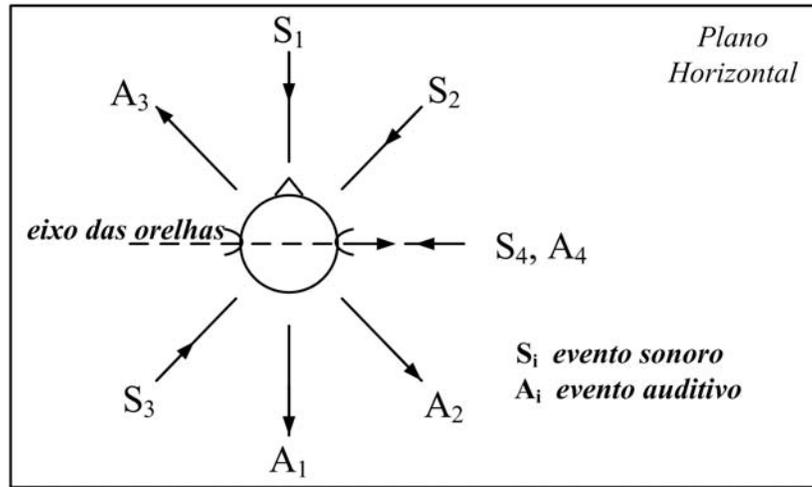


Figura 2.2: Diferença direcional entre evento auditivo (A_i) e sonoro (S_i) no PH

central do sinal [2] (fig. 2.3). A resolução espacial no *PM* é bem mais baixa que no *plano horizontal* devido ao fato de não haver diferenças no tempo de chegada nem de intensidades dos sinais no *PM* que forneçam informações sobre a fonte [2].

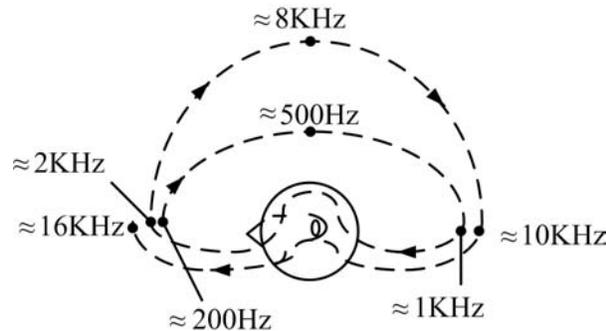


Figura 2.3: Caminho das direções do evento auditivo em função da frequência central do sinal ruidoso de banda estreita no *PM*

2.2.2 Diferença de tempo interaural - ITD

A *diferença de tempo interaural* é uma das pistas auditivas mais importantes para localizar um *evento auditivo* no *PH*. Ela é definida como sendo a diferença entre os tempos de chegada de uma frente de onda à orelha esquerda e à orelha direita. Em geral, se uma frente de onda chega à orelha esquerda antes de atingir a orelha direita, o som será percebido como estando mais próximo à orelha esquerda e, assim, haverá uma *lateralização* do *evento auditivo* para este lado. Avaliando este parâmetro

2. Bases Teóricas

interaural para tons puros, percebe-se claramente a relação entre as *lateralizações* e as diferenças de fase dos sinais às entradas dos condutos auditivos (fig. 2.4).

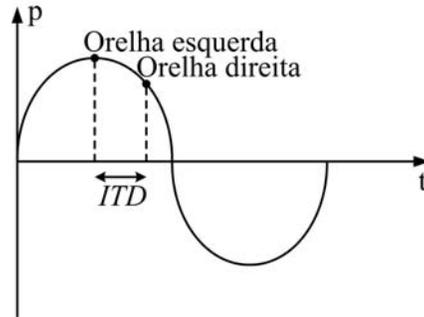


Figura 2.4: *Diferença de tempo interaural*

O sistema auditivo se utiliza quase que exclusivamente da *ITD* para localizar uma fonte sonora só até um determinado comprimento de onda, aproximadamente 1500 Hz . A partir desta frequência a informação da *ITD* se torna ambígua e não é possível determinar um ângulo azimutal único baseado na diferença de fase (fig. 2.5). Então, o sistema auditivo se utiliza de outros recursos referentes às diferenças dos sinais presentes nos canais auditivos.

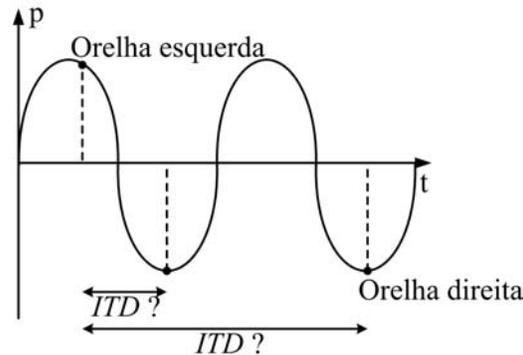


Figura 2.5: *Ambiguidade na diferença de tempo interaural*

2.2.3 *Diferença de nível interaural - ILD*

Quando o comprimento de onda se torna muito pequeno ($f > 1500\text{ Hz}$) a influência da cabeça enquanto obstáculo começa a impor uma diferença de intensidades entre os sinais acústicos que atingem as orelhas. Basicamente, a cabeça começa a fazer sombra para a orelha mais distante da fonte. A diferença entre as amplitudes dos sinais nas duas orelhas é denominada de *diferença de intensidade interaural (IID)*

ou, se considerarmos diferença de níveis de pressão sonora, de *diferença de nível interaural (ILD)*.

Assim como a *ITD* desempenha um papel importante na *lateralização* do *evento auditivo*, a *ILD* também é de grande importância para a determinação do *azimute* no *PH*. A orelha que está mais próxima da fonte sonora, receberá um sinal mais intenso que o da orelha oposta e por isto o *evento auditivo* estará deslocado para o lado de maior intensidade.

2.2.4 *Reversões frente/costas*

Das anomalias que ocorrem no *PH* e no *PM*, citadas anteriormente, percebe-se que, em algumas situações, o sistema auditivo pode localizar *eventos auditivos* em posições totalmente reversas em comparação com a posição real do *evento sonoro*. No caso específico da fonte sonora no *plano médio*, é evidente que não há diferenças interaurais e se o ouvinte não se utilizar de movimentos da cabeça para criar estas diferenças, a única informação disponível será o conteúdo espectral do sinal. Essas anomalias são denominadas de *reversões frente/costas* e são de grande importância nos sistemas de *auralização* pois exigem que haja transparência espectral nas simulações.

2.2.5 *Localização no interior da cabeça - IHL*

A *localização no interior da cabeça* está entre os fenômenos psicoacústicos relacionados à chamada *distância auditiva (distance hearing)*. Esta se refere às relações entre a distância do *evento auditivo* e os atributos dos eventos correlacionados com esta distância [2]. A distância do *evento auditivo* é tomada a partir do ponto médio do *eixo das orelhas*, a origem do sistema referencial relacionado à cabeça. A *IHL* ocorre quando esta distância é menor do que o raio do crânio e, assim, o *evento auditivo* localiza-se dentro da cabeça.

A *IHL* costuma acontecer quando se utiliza fones de ouvido. Esta é uma situação facilmente percebida por usuários de sistemas de som residenciais que experimentam a sensação de ter o som dos instrumentos no interior de suas cabeças. Isto ocorre, principalmente, devido ao fato de que ao utilizar fones de ouvido estamos eliminando os efeitos de filtragem do pavilhão auricular sobre o som incidente, que ocorreriam numa situação normal de audição. Para além disto, existem dois pré-requisitos básicos para a ocorrência de *IHL* mesmo quando se utiliza reprodução via alto-falantes:

2. Bases Teóricas

- Os dois *sinais de entrada das orelhas* devem ser similares de tal forma que ocorra um único *evento auditivo* homogêneo.
- Cada uma das fontes sonoras deve estar próxima à orelha, ou percebida como estando próxima quando estiver operando separadamente.

2.3 Tecnologia Binaural

O termo *tecnologia binaural* se refere ao conjunto de técnicas de gravação e reprodução das pressões acústicas presentes nos tímpanos para recriar um *evento auditivo* sem que haja diferenças de localização espacial e timbrais entre o que foi gravado e o que será reproduzido. Estes sinais, que são coletados nas membranas timpânicas, e que possuem todas as informações espaciais e espectrais de uma imagem sonora são denominados de *sinais binaurais*.

Os estudos psicoacústicos costumam analisar e identificar a relação entre as pistas interaurais e timbrais, em cada plano do sistema de referência, com a localização espacial do *evento auditivo*. Porém, as transformações imprimidas a um sinal de uma fonte sonora pode ser melhor representado e analisado pelo par de *respostas impulsivas binaural (HRIR)* e seus espectros de frequências denominados de *funções de transferência relacionadas à cabeça (HRTF)* [2].

2.3.1 Funções de transferência relacionadas à cabeça - HRTF

Imaginemos um alto-falante alimentado em seus terminais por um sinal de áudio $e(t)$. Esta fonte sonora promove uma pressão acústica, medida num ponto de referência, identificada por $p_1(t)$.

Se um ouvinte passa a ocupar um lugar neste espaço de tal forma que o centro de sua cabeça (origem do sistema referencial) seja o ponto onde foi tomada a pressão $p_1(t)$, teremos as seguintes funções de transferência de *campo livre* (2.25), (2.26), (2.27):

$$P_4(\omega)/P_1(\omega) = \frac{\text{pressão sonora no tímpano}}{\text{pressão sonora no centro da cabeça com o ouvinte ausente}} \quad (2.25)$$

$$P_3(\omega)/P_1(\omega) = \frac{\text{pressão sonora na estrada do canal auditivo}}{\text{pressão sonora no centro da cabeça com o ouvinte ausente}} \quad (2.26)$$

$$P_2(\omega)/P_1(\omega) = \frac{\text{pressão sonora na entrada do canal bloqueado}}{\text{pressão sonora no centro da cabeça com o ouvinte ausente}} \quad (2.27)$$

Estas funções de transferência dependem do ângulo de incidência (φ e β) e da distância até a fonte sonora (r_f). Jens Blauert [2] indica que se esta distância for maior que $3m$ a onda incidente já pode ser considerada como plana e que, assim, as *HRTF* anteriores tornam-se quase independentes da distância da fonte.

Para além das *funções de transferência de campo livre* existem dois outros tipos de *HRTF*:

- *Função de transferência monaural (MTF)* - função que relaciona a pressão sonora, em um ponto de medida no canal auditivo, de uma fonte sonora em uma posição qualquer (direção e distância) a uma pressão sonora medida no mesmo ponto mas com a fonte sonora numa posição de referência (normalmente $\varphi = 0$ e $\beta = 0$).
- *Função de transferência interaural (ITF)* - função que realaciona as pressões sonoras em pontos de medidas correspondentes nos dois canais auditivos. A pressão de referência será a da orelha que estiver voltada para a fonte sonora.

Independente do tipo de *HRTF*, a representação matemática da equação (2.28) será válida:

$$\mathbf{HRTF} = |\mathbf{HRTF}|e^{-jb} \quad (2.28)$$

onde o parâmetro (b) depende da frequência. Em geral, a magnitude ($|\mathbf{HRTF}|$) e a fase (b) são substituídos pelos parâmetros $\Delta L = 20 \log |\mathbf{HRTF}|$ (*diferença de nível*) e $\tau_{gr}(f) = db(f)/d2\pi f$ (*atraso de grupo*).

Um dos pontos capitais relacionados às *HRTF* é a relação entre direção de incidência e a variação destas funções de transferência (2.25, 2.26, 2.27) para diferentes pontos de coleta no canal auditivo. A partir da entrada do conduto, a propagação sonora passa a ser unidimensional. Isto quer dizer que independe da direção e da distância da fonte sonora [2, 9].

A equação (2.27) utiliza um ponto de medida à entrada do canal auditivo estando este bloqueado. Esta pressão pode ser imaginada como uma pressão de "circuito aberto" e ser relacionada às demais pressões $P_3(\omega)$ e $P_4(\omega)$ por intermédio de um circuito elétrico análogo (fig. 2.6):

2. Bases Teóricas

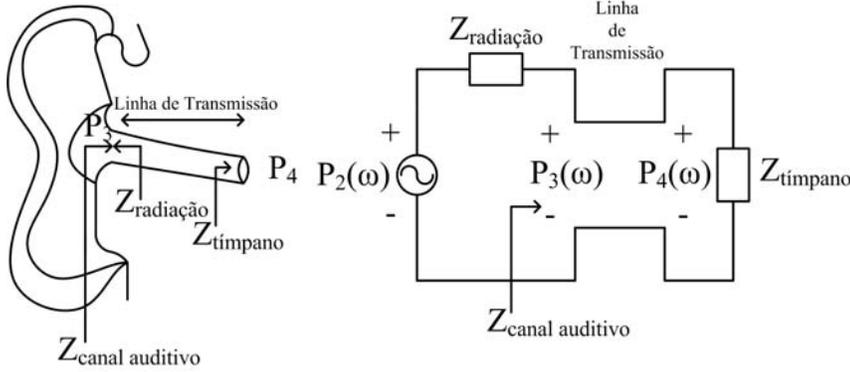


Figura 2.6: Circuito análogo à transmissão sonora na orelha.

Os argumentos dos espectros das pressões na figura (2.6) foram omitidos com o intuito de simplificar a notação. Doravante, quando necessário, variáveis em letras maiúsculas serão funções espectrais e em letras minúsculas serão funções do tempo. As impedâncias são definidas como: impedância vista a partir da entrada do canal ($Z_{\text{radiação}}$), impedância vista a partir da entrada para o interior deste ($Z_{\text{canal auditivo}}$) e impedância do tímpano ($Z_{\text{tímpano}}$). Esta analogia elétrica é válida para uma determinada faixa de frequências nas quais o comprimento de onda é muito maior do que o diâmetro do canal auditivo $d_{\text{canal}} = 8\text{mm}$, assumido que este tenha seção transversal constante. Este limite superior será dado por $f \ll \frac{340\text{m/s}}{8 \times 10^{-3}\text{m}} = 42,5\text{KHz}$. Porém, um outro limite superior, considerando um diâmetro de $(1/4)\lambda$, é utilizado e garante que o modelo comporta-se como uma linha de transmissão até cerca de 10KHz .

Se o canal auditivo está bloqueado isto significa que a *velocidade de partícula* em sua entrada é zero. Do modelo elétrico observa-se que velocidade de partícula zero na entrada do canal, analogamente, corresponde a uma corrente igual a zero através da impedância de radiação $Z_{\text{radiação}}$. Assim, a pressão $P_2(\omega)$ estará presente na entrada do canal e poderá ser medida. Para calcular, por exemplo, a pressão $P_3(\omega)$ basta utilizar o divisor de tensão correspondente (eq. 2.29):

$$\frac{P_3(\omega)}{P_2(\omega)} = \frac{Z_{\text{canal auditivo}}}{Z_{\text{canal auditivo}} + Z_{\text{radiação}}} \quad (2.29)$$

Para a realização de medidas de *HRTF*, o sinal da fonte no ponto de referência $P_1(\omega)$ não pode ser medido simultaneamente às demais pressões tomadas ao longo do conduto auditivo, já que $P_1(\omega)$ refere-se à posição do centro da cabeça. Contudo, as

funções de transferência definidas anteriormente podem ser avaliadas por intermédio da *função de transferência do alto-falante* (2.30):

$$\frac{P_1(\omega)}{E_{\text{alto-falante}}(\omega)} \quad (2.30)$$

onde $E_{\text{alto-falante}}(\omega)$ é a transformada de *Fourier* de $e(t)$, o sinal elétrico nos terminais do alto-falante. Desta forma, chegamos às *HRTF*, onde o índice i refere-se ao ponto de coleta da pressão, *ponto referencial* (2.31):

$$\frac{P_i(\omega)}{P_1(\omega)} = \frac{[P_i(\omega)/E_{\text{alto-falante}}(\omega)]}{[P_1(\omega)/E_{\text{alto-falante}}(\omega)]} \quad (2.31)$$

2.3.2 Audição via fones de ouvido

O par de *respostas impulsivas binaural (HRIR)* traduz todas as informações codificadas pelo torso, cabeça e orelhas no sinal recebido pelo ouvinte. No momento da decodificação e audição destes *sinais binaurais*, cada um deles deve ser apresentado à sua orelha de origem, ou seja, o sinal originalmente captado do lado esquerdo deverá ser apresentado, somente, deste lado, e o do lado direito só para o direito. Com estas exigências de exclusividade para as reproduções dos sinais, o uso de fones de ouvido se torna quase que imprescindível.

Para utilizar fones de ouvido é necessário, então, conhecer suas funções de transferência relacionadas ao *ponto de referência* nos quais os sinais foram captados.

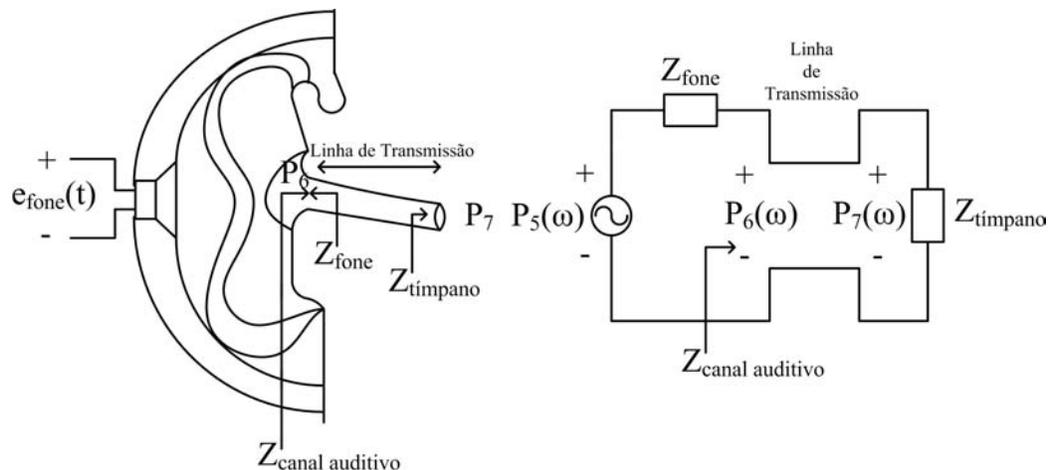


Figura 2.7: Circuito análogo à transmissão sonora na orelha via fones de ouvido

2. Bases Teóricas

Da figura (2.7), constata-se que Z_{fone} é a impedância que é vista da entrada do canal auditivo para fora. Mais uma vez, as funções de transferência são determinadas (2.32, 2.33, 2.34):

$$P_7(\omega)/E_{\text{fone}}(\omega) = \frac{\text{pressão sonora no tímpano}}{\text{tensão nos terminais do fone de ouvido}} \quad (2.32)$$

$$P_6(\omega)/E_{\text{fone}}(\omega) = \frac{\text{pressão sonora na estrada do canal auditivo}}{\text{tensão nos terminais do fone de ouvido}} \quad (2.33)$$

$$P_5(\omega)/E_{\text{fone}}(\omega) = \frac{\text{pressão sonora na entrada do canal bloqueado}}{\text{tensão nos terminais do fone de ouvido}} \quad (2.34)$$

A medida destas funções de transferência é imediata já que todas elas (2.32, 2.33, 2.34) possuem como entrada o sinal elétrico do fone. Neste caso, a representação de uma pressão medida num ponto de referência no centro da cabeça perde o sentido.

Os dois casos de captação ou gravação com ou sem fones de ouvido possuem similaridades que valem a pena ser examinadas para uma posterior *equalização* dos fones de ouvido para a reprodução. Dentro do canal auditivo a transmissão da entrada do canal até o tímpano se dá sem a interferência da posição da fonte sonora. Isto significa que $[P_7(\omega)]/[P_6(\omega)] = [P_4(\omega)]/[P_3(\omega)]$, ou seja, a propagação no canal depende somente deste e da sua impedância de carga $Z_{\text{tímpano}}$.

Realizando a mesma análise da pressão de "circuito aberto" ($P_5(\omega)$), feita anteriormente para $P_2(\omega)$, chega-se a uma relação similar a encontrada para $[P_3(\omega)]/[P_2(\omega)]$ (eq. 2.29):

$$\frac{P_6(\omega)}{P_5(\omega)} = \frac{Z_{\text{canal auditivo}}}{Z_{\text{canal auditivo}} + Z_{\text{fone}}} \quad (2.35)$$

Com o intuito de realizar a *calibração* necessária dos fones de ouvido para a reprodução fidedigna dos sinais binaurais, é necessário avaliar as duas relações para "circuito aberto":

$$\frac{[P_3(\omega)/P_2(\omega)]}{[P_6(\omega)/P_5(\omega)]} = \frac{Z_{\text{canal auditivo}} + Z_{\text{fone}}}{Z_{\text{canal auditivo}} + Z_{\text{radiação}}} \quad (2.36)$$

Se as impedâncias $Z_{\text{radiação}}$ e Z_{fone} são iguais, ou seja, as impedâncias vistas a partir da entrada do canal para fora, com e sem fone, se equivalem, e se $Z_{\text{fone}} \ll$

$Z_{\text{canal auditivo}}$, então, este fone possui propriedades *FEC* (*free-air equivalent coupling*), pois sua presença não altera a impedância de radiação que existiria sem o seu uso. Desta forma, para fones de ouvido *FEC* a relação (2.36) torna-se $(P_3/P_2) \approx (P_6/P_5)$.

2.3.3 Técnicas binaurais de gravação e reprodução

Como pudemos ver até agora, existem três possibilidades de escolha para os *pontos de referência* para se executar medidas de *HRTF* ou gravações de sinais binaurais. Como o objetivo de qualquer um destes procedimentos é a reprodução, principalmente via fones de ouvido, devido a sua flexibilidade de uso em relação aos ambientes e à destinação dos sinais pertinentes à cada orelha, se faz necessária uma equalização referente à cadeia de dispositivos elétricos percorrida pelo áudio binaural.

A *calibração* de um fone de ouvido, é o nome que se dá ao levantamento de sua função de transferência [9]. Para cada *ponto de referência* em que se captam os *sinais binaurais* deve-se proceder a uma *calibração* para que as pressões, na reprodução, estejam equalizadas corretamente.

No primeiro caso, o sinal é reproduzido com fones de ouvido e captado nos tímpanos. Considerando que o microfone utilizado para captação possui uma função de transferência M , a função de transferência total entre o campo sonoro $p_1(t)$ e a pressão nos tímpanos na reprodução pelos fones $p_7(t)$, será:

$$\frac{P_7}{P_1} = \frac{P_7}{E_{\text{fone}}} \cdot \frac{E_{\text{fone}}}{E_{\text{mic}}} \cdot M \cdot \frac{P_4}{P_1} \quad (2.37)$$

O objetivo é que esta função de transferência seja igual a $[P_4/P_1]$, para que haja audição do mesmo evento ocorrido no momento da captação. Desta forma, chega-se a seguinte equação equalizadora:

$$Ga_1 = \frac{E_{\text{fone}}}{E_{\text{mic}}} = \frac{1}{M \cdot (P_7/E_{\text{fone}})} \quad (2.38)$$

Ga_1 é a função de transferência que deve ser aplicada para compensar as respostas em frequência dos microfones e dos fones de ouvido.

No segundo caso, *ponto de referência* na entrada do canal, realizando a mesma análise para compensar os circuitos elétricos transdutores, teremos a seguinte função de transferência entre $[P_6/P_1]$:

$$\frac{P_6}{P_1} = \frac{P_6}{E_{\text{fone}}} \cdot \frac{E_{\text{fone}}}{E_{\text{mic}}} \cdot M \cdot \frac{P_3}{P_1} \quad (2.39)$$

2. Bases Teóricas

O resultado desta transferência deve ser igual a (P_3/P_1) , logo:

$$Ga_2 = \frac{E_{\text{fone}}}{E_{\text{mic}}} = \frac{1}{M \cdot (P_6/E_{\text{fone}})} \quad (2.40)$$

Assim, Ga_2 representa, neste caso, a equalização para este *ponto de referência*.

No terceiro caso, medida de pressão à entrada do canal bloqueado, a função de transferência, será:

$$\frac{P_5}{P_1} = \frac{P_5}{E_{\text{fone}}} \cdot \frac{E_{\text{fone}}}{E_{\text{mic}}} \cdot M \cdot \frac{P_2}{P_1} \quad (2.41)$$

Igualando-a a condição de fidedignidade da reprodução (P_2/P_1) a equalização será:

$$Ga_3 = \frac{E_{\text{fone}}}{E_{\text{mic}}} = \frac{Z_{\text{canal auditivo}} + Z_{\text{fone}}}{Z_{\text{canal auditivo}} + Z_{\text{radiação}}} \cdot \frac{1}{M \cdot (P_5/E_{\text{fone}})} \quad (2.42)$$

Para um fone de ouvido *FEC* a relação $\frac{Z_{\text{canal auditivo}} + Z_{\text{fone}}}{Z_{\text{canal auditivo}} + Z_{\text{radiação}}}$ é unitária, e a equalização é simplificada:

$$Ga_3 = \frac{E_{\text{fone}}}{E_{\text{mic}}} = \frac{1}{M \cdot (P_5/E_{\text{fone}})} \quad (2.43)$$

Da análise da equação (2.42) percebe-se que o ganho deve compensar tanto o microfone e os fones de ouvido quanto as diferenças dos acoplamentos acústicos causados pelas impedâncias diferentes. Se um fone *FEC* ideal for utilizado essa compensação para impedâncias não será necessária.

É preferível que a *calibração* dos fones de ouvido seja feita com o mesmo microfone utilizado para a captação dos *sinais binaurais*, pois desta forma compensações para diferenças de respostas entre microfones distintos se tornam desnecessárias.

É importante ressaltar que medidas em cabeças reais realizadas na entrada do canal bloqueado não perturbam o campo sonoro medido desde que o microfone possua dimensões apropriadas e esteja embutido no plug que veda o canal [9, 10]. Se o microfone estiver na entrada ou for introduzido no canal auditivo aberto, haverá interferência do instrumento mensurador nas medidas. Assim, um estudo sobre a influência destas perturbações na função de transferência total se faz necessário.

Utilizemos um ponto de medida genérico, onde a pressão é $p_x(t)$ e seu espectro com notação simplificada P_x (fig. 2.8), entre os *pontos de referência* das pressões P_3

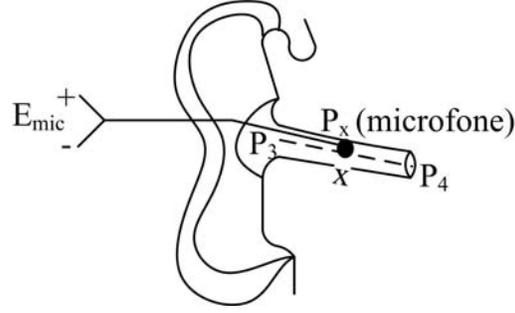


Figura 2.8: Medida com microfone em ponto genérico dentro da orelha externa

e P_4 . Desta forma, a função de transferência entre a pressão de *campo livre* (P_1) e a tensão nos terminais do microfone será:

$$\frac{E_{mic}}{P_1} = \frac{E_{mic}}{P_x} \cdot \left(\frac{P_x}{P_1}\right)_{pert}; \quad \frac{E_{mic}}{P_x} = M \text{ função de transferência do microfone} \quad (2.44)$$

o fator $(P_x/P_1)_{pert}$ dá conta das perturbações introduzidas pelo microfone. Se a calibração do fone de ouvido é realizada com o mesmo microfone, nas mesma posição no canal:

$$G = \frac{1}{M \cdot (P_x/E_{fone})_{pert}} \quad (2.45)$$

A função de transferência total até os tímpanos, será:

$$(P_7/P_1) = (P_7/E_{fone}) \cdot G \cdot M \cdot (P_x/P_1)_{pert} = \frac{(P_x/P_1)_{pert} \cdot (P_7/E_{fone})}{(P_x/E_{fone})_{pert}} \quad (2.46)$$

Deseja-se que $(P_7/P_1) = (P_4/P_1)$. Então, pode-se avaliar a perturbação imposta pelo microfone pelo erro $(P_7/P_1)/(P_4/P_1)$, que será:

$$erro = \frac{(P_x/P_1)_{pert} \cdot (P_7/E_{fone})}{(P_4/P_1) \cdot (P_x/E_{fone})_{pert}} \quad (2.47)$$

Lembrando que dentro do canal auditivo com ou sem fone a seguinte relação é válida, $(P_x/P_6)_{pert} = (P_x/P_3)_{pert}$ e simplificou o resultado da equação (2.47). Utilizando as equações (2.29), (2.7) e $(P_7/P_6) = (P_4/P_3)$, a equação do *erro* (2.47) pode ser simplificada:

$$erro = \frac{(P_3/P_2)_{pert} \cdot (P_6/P_5)}{(P_3/P_2) \cdot (P_6/P_5)_{pert}} = \frac{Z_{canal\ auditivo} + Z_{radiação}}{Z_{canal\ auditivo\ pert} + Z_{radiação}} \cdot \frac{Z_{canal\ auditivo\ pert} + Z_{fone}}{Z_{canal\ auditivo} + Z_{fone}} \quad (2.48)$$

Da equação do *erro* em função das impedâncias (2.48), percebe-se que este está relacionado ao descasamento destas impedâncias acústicas. Se um par de fones *FEC* forem utilizados, para os quais $Z_{fone} = Z_{radiação}$, então o *erro* introduzido será nulo já que a equação 2.48 terá um valor unitário para este caso.

2.4 Processamento Digital de Sinais

Com o crescente avanço da tecnologia digital e da capacidade de processamento dos computadores, a manipulação dos sinais de áudio, por intermédio do processamento digital, se tornou corriqueira e até mesmo necessária às demandas atuais. Quando se pretende implementar um sistema de *auralização*, que possa ser utilizado num computador pessoal, as teorias e técnicas sobre sinais e sistemas e sobre o processamento de sinais discretos no tempo passam a orientar os resultados da acústica e da psicoacústica em direção a realização virtual de ambientes e *eventos auditivos*.

2.4.1 Resposta ao impulso de um sistema

Para utilizar todas as possibilidades de implementação dos sistemas *PDS* é preciso compreender algumas propriedades básicas dos *sistemas lineares invariantes no tempo* (*SLIT*) [11] vigentes neste trabalho:

- Causalidade - um sistema físico realizável não pode produzir saída sem que haja uma entrada prévia.
- Estabilidade - as saídas de um sistema estável serão finitas se as entradas forem sinais finitos.
- Invariabilidade temporal - os parâmetros do sistema não variam ao longo do tempo.
- Linearidade - o sistema possui a propriedade de *superposição* (*adição e homogeneidade*). Na primeira, se cada saída está relacionada a uma determinada entrada então, a soma de todas as entradas produzirá a soma de todas estas saídas. Na segunda propriedade, um sistema homogêneo produzirá uma saída multiplicada por um fator, se este estiver multiplicado pela entrada que a gerou.

Estas propriedades formam a base de toda a discussão subsequente deste texto. Um sistema *linear e invariante no tempo* pode ser descrito completamente por sua *resposta impulsiva* $h(t)$ (fig. 2.9):

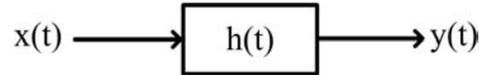


Figura 2.9: *SLIT*

A *resposta impulsiva* relaciona a entrada $x(t)$ e a saída $y(t)$ por intermédio da *convolução* (eq. 2.49):

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(\tau) \cdot h(t - \tau) d\tau \quad (2.49)$$

A função $h(t)$ recebe o nome de resposta ao impulso, pois, se o *SLIT* for excitado pela função *Delta de Dirac* $\delta(t)$ apresentará em sua saída a própria $h(t)$. Se passarmos para o domínio da frequência via *transformada de Fourier* é obtida a seguinte relação (2.50):

$$Y(\omega) = X(\omega) \cdot H(\omega) \Rightarrow H(\omega) = \frac{Y(\omega)}{X(\omega)} \quad (2.50)$$

onde $X(\omega)$, $Y(\omega)$ e $H(\omega)$ são os espectros de frequências das funções contínuas no tempo $x(t)$, $y(t)$ e $h(t)$, respectivamente.

Até aqui, todas estas definições consideram que os sinais são contínuos no tempo. Para proceder às manipulações destes sinais no formato digital, se faz necessário redefinir algumas propriedades anteriormente citadas, para *sinais discretos no tempo* [11].

A função *impulso*, $\delta(t)$, terá sua versão discreta definida por (eq. 2.51):

$$\begin{aligned} \delta[n] &= 1, n = 0 \\ \delta[n] &= 0, n \neq 0 \end{aligned} \quad (2.51)$$

Mais uma vez, a resposta de um sistema linear discreto e invariante no tempo, ao *impulso* discreto, relaciona a sequência de entrada $x[n]$ e a sequência de saída $y[n]$ via *convolução linear discreta* (eq. 2.52):

2. Bases Teóricas

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n - k] \quad (2.52)$$

Se aplicarmos a *transformada discreta de Fourier* (*DFT*) obteremos os conteúdos espectrais das sequências acima: $x[n] \rightleftharpoons X(z)$, $y[n] \rightleftharpoons Y(z)$ e $h[n] \rightleftharpoons H(z)$ (*pares transformados*).

2.4.2 Sistemas PDS simples

Os algoritmos de *PDS* utilizados para processamento de sinais de áudio geralmente são combinações de sistemas simples que executam operações básicas definidas. Dentre estas, as mais importantes modificam uma determinada sequência de entrada $x[n]$ por intermédio da multiplicação por um fator, da introdução de atraso, ou adicionando-a a outro sinal [11].

A multiplicação de uma sequência por uma constante é feita pelos *multiplicadores* (fig. 2.10).

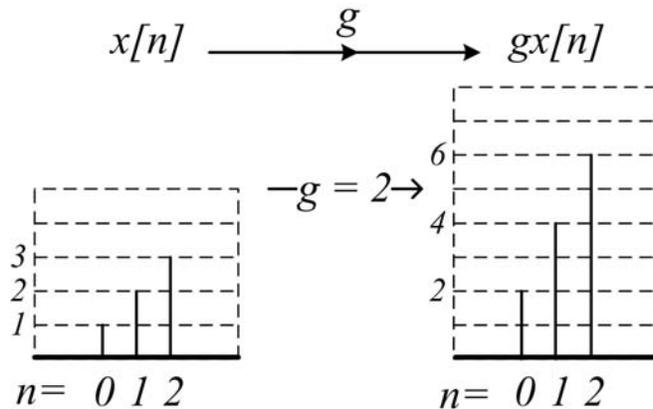


Figura 2.10: *Multiplicação de uma sequência por uma constante. No exemplo acima, $g=2$.*

O *atraso* de uma sequência é dado em amostras. Assim um dispositivo deste tipo atrasa a transferência da sequência de entrada $x[n]$ para a saída por um tempo igual a um determinado número de amostras. Tal sistema está representado na figura (2.11).

A adição de sequências é um procedimento trivial utilizado no áudio digital e está representada na figura (2.12).

No caso da adição, é interessante observar que as aplicações reais necessitam de um limite em relação ao resultado da soma das sequências para que o sinal de saída

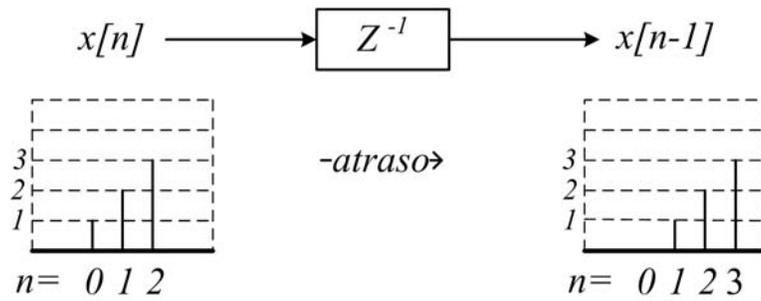


Figura 2.11: Atraso de uma sequência por uma amostra.

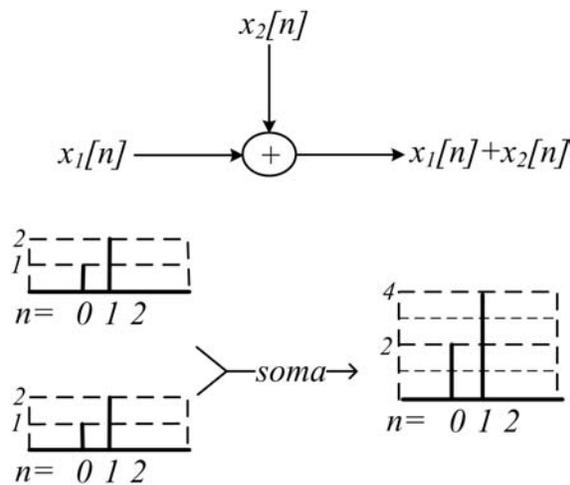


Figura 2.12: Adição de duas sequências

não ultrapasse o valor máximo de quantização. Por isso, antes da soma, as sequências podem ser multiplicadas por um determinado *ganho* para limitar o valor final.

2.4.3 Convolução e filtragem digital (filtros FIR e IIR)

Os conceitos de *filtragem* e *convolução* estão profundamente ligados. Quando um sinal $x[n]$ passa através de um filtro $h[n]$, matematicamente, os seus espectros serão multiplicados, ou seja, a saída será $X[z] \cdot H[z]$. No domínio do tempo, o que se observa é o sinal $x[t] * h[t]$.

A convolução (eq. 2.52) é uma operação que envolve multiplicações e somas de sequências. A avaliação da convolução no domínio do tempo das duas sequências da figura (2.13) evidencia esta relação filtro e convolução.

O sinal $x[n]$ será filtrado pelo filtro $h[n]$, o primeiro passo da figura (2.13). No segundo passo, avaliando a equação (2.52), a primeira parcela do somatório será $x[n] \cdot h[0]$, ou seja, o primeiro coeficiente de $h[n]$ multiplicado pelo sinal a ser filtrado,

2. Bases Teóricas

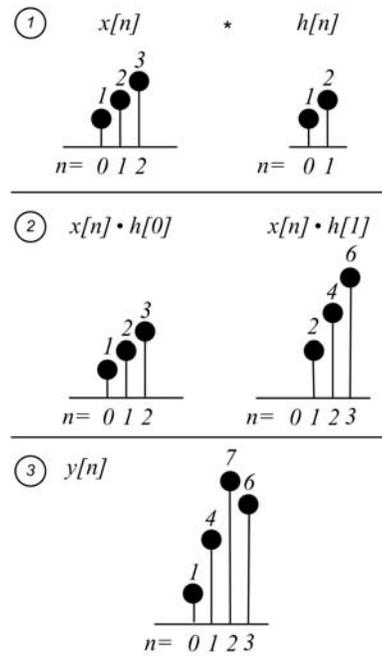


Figura 2.13: *Convolução*

e a segunda parcela será $x[n] \cdot h[1]$, o segundo coeficiente da resposta impulsiva do filtro multiplicado pela sequência de entrada. No terceiro passo, as sequências resultantes dos passos anteriores são adicionadas para resultar no sinal $y[n]$, filtrado.

Uma análise mais atenta das parcelas que compõem o somatório que resulta em $y[n]$, mostra que $h[0]$ promove um *atraso* no sinal original de $n = 0$ amostras e um *ganho* dado pelo coeficiente 1. Já $h[1]$, produz um *atraso* em $x[n]$ de $n = 1$ amostras e um *ganho* igual a 2. Se utilizarmos os sistemas básicos de *PDS* discutidos anteriormente para representar estas modificações, teremos o diagrama da figura 2.14.

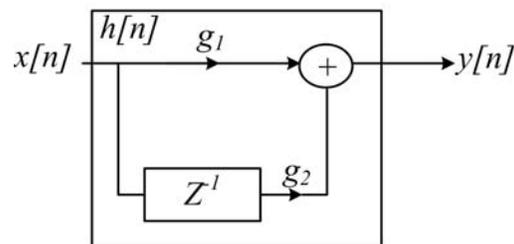


Figura 2.14: *Filtro de resposta ao impulso finita (FIR)*

Se um sistema promove uma operação que consiste em gerar em sua saída uma soma da sua entrada atual e de versões atrasadas ele é denominado de *filtro FIR* (*resposta ao impulso finita*).

Se ao invés de alimentar a saída com versões do sinal de entrada realimentar-mos a entrada com versões *atrasadas* da saída, como na figura 2.15, obteremos o chamado *filtro IIR (resposta ao impulso infinita)*. Observa-se que deve haver uma limitação em relação aos valores de g_2 para que as amplitudes na saída do sistema não assumam valores infinitos.

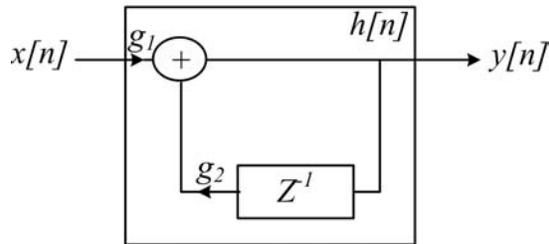


Figura 2.15: *Filtro de resposta ao impulso infinita (IIR)*

2.4.4 *Sistemas fase-mínima*

Para um determinado *SLIT* com função de transferência $H(z)$, o sistema inverso correspondente é definido como sendo aquele que possui função de transferência $H_i(z)$ de tal forma que se eles forem dispostos numa configuração em cascata a função de transferência equivalente será unitária [11], ou seja (eq. 2.53):

$$G(z) = H(z) \cdot H_i(z) = 1 \quad (2.53)$$

Isto implica que (eq. 2.54),

$$H_i(z) = \frac{1}{H(z)} \quad (2.54)$$

E a equivalência no domínio do tempo é (eq. 2.55),

$$g[n] = h[n] * h_i[n] = \delta[n] \quad (2.55)$$

A resposta em frequência de um *SLIT* é a transformada de *Fourier* de sua resposta impulsiva. Em geral, o conhecimento sobre a magnitude não provê nenhuma informação sobre a fase, e vice-versa. Contudo, para sistemas descritos por equações a diferenças lineares com coeficientes constantes há ligação entre a magnitude e a fase. Em particular, se a magnitude da resposta em frequência e o número de pólos e zeros for conhecido, então só haverá um número finito de escolhas para a fase associada. De forma similar, por intermédio de um fator de escala, haverá um número

2. *Bases Teóricas*

finito de possibilidades para a magnitude se o número de pólos, número de zeros e a fase forem conhecidos. Se o sistema for estável e causal, os pólos devem estar dentro do círculo unitário [11], mas a estabilidade e a causalidade não restringem o *locus* dos zeros. Se o sistema inverso também for estável e causal, aí sim, os zeros do sistema original estarão dentro do círculo unitário. Sistemas que obedecem tais restrições são denominados de *sistemas fase-mínima* e, assim, a partir da magnitude de sua resposta em frequência pode-se especificar exatamente a sua fase, e a fase especifica a magnitude por intermédio de um fator de escala.

CAPÍTULO 3

Auralização

Neste capítulo definimos o conceito de *auralização* e apresentamos um breve histórico a respeito do tema nos centros de pesquisa e universidades. As técnicas atuais de auralização são apresentadas e, do confronto de suas vantagens e desvantagens, escolhe-se a auralização inteiramente computacional como objeto de pesquisa.

3.1 Definição – O Que é a Auralização?

Nos últimos anos a história da acústica de ambientes incorporou um de seus mais fascinantes capítulos escritos até agora: a auralização [4]. Este termo foi cunhado para significar as técnicas que permitem tornar audível um campo sonoro de uma fonte, num determinado espaço, por modelagem física ou matemática, de tal forma que se possa simular a experiência de audição binaural numa posição específica do ambiente modelado [5].

Primordialmente, o objetivo desta tecnologia não é recriar a sensação da fala, da música ou de qualquer outro sinal de interesse relevante, e sim recriar a impressão aural das características acústicas de um espaço, seja ele fechado ou aberto, que estão codificadas na informação emitida. A auralização pode ser feita utilizando modelos acústicos em escala ou modelagem computacional para obtenção da *BRIR* ou funções de transferência. Estas últimas são usadas para filtrar o sinal anecóico da fonte determinando-lhe, assim, a assinatura acústica do ambiente usando processamento digital de sinais. As *respostas impulsivas binaurais* podem, também, ser medidas diretamente num ambiente existente com a aplicação de técnicas binaurais, como por exemplo o uso de *cabeças artificiais* [9].

As técnicas de auralização, ou simulação binaural de ambientes oferecem a possibilidade do uso dos ouvidos para escutar a acústica de um ambiente ainda durante a fase de projeto [7]. Muitos problemas num espaço podem ser facilmente detectados

3. Auralização

perceptualmente, enquanto que, apenas com o uso de parâmetros acústicos objetivos, a dificuldade de percebê-los seria demasiadamente complexa.

3.2 Histórico

Os primeiros experimentos de auralização foram realizados por Spandöck e seu grupo de pesquisa em Munique em 1934. Eles utilizavam modelos físicos em escala do ambiente sonoro em estudo [12] e um método denominado de técnica do gramofone. Primeiramente, um sinal anecóico, sinal este desprovido das distorções impostas pela acústica de um ambiente, era gravado num cilindro de cera. Este, então, era reproduzido no modelo, por intermédio de um alto-falante, com a velocidade de reprodução ajustada para um valor 5 vezes maior que a de gravação, pois a escala de redução era de 1:5. O som, no modelo, era captado com a mesma velocidade de rotação que foi inserido e era registrado num outro cilindro. Usando fones de ouvido Spandöck podia ouvir o som do espaço modelado com o cilindro de cera girando à sua velocidade normal de reprodução, ou seja, novamente, 5 vezes mais lenta para obter a impressão do evento auditivo.

Os predecessores dos sistemas de auralização computacional foram os algoritmos de reverberação artificial. Pioneiramente, Schroeder [13, 14] os implementou, permitindo que, mais tarde, reverberadores digitais fossem desenvolvidos para a indústria da música e do áudio. As similaridades entre os objetivos dos sistemas de reverberação e de auralização são justificadas, por ambos gerarem a sensação de ambiência pelo decaimento do campo sonoro [8].

No final da década de 1970 Allen e Berkley realizaram um algoritmo de simulação para levantamento da *resposta impulsiva* de um ambiente puramente digital [15]. Este método de simulação foi concebido para ambientes pequenos e era baseado em fontes virtuais sendo a geometria das salas modeladas retangular. Esta escolha foi feita devido às seguintes considerações: interesse em ambientes de escritório, que são normalmente retangulares; a implementação de um programa computacional deste modelo de sala é mais simples; o método das *fontes virtuais* converge, rapidamente, para uma solução exata da equação de onda quando as paredes são *rígidas*.

Na França, no final da década de 1980, no Centre Scientifique et Technique du Bâtiment, foi realizada uma das primeiras tentativas de modelagem acústica de ambientes para simulação binaural de salas de concerto, ou seja, auralização [16]. O

programa computacional *Epidaure*, que calculava a *RIR* (*resposta impulsiva do ambiente*), foi incrementado com a capacidade de gerar *BRIR* (*resposta impulsiva binaural do ambiente*).

A partir do início da década de 1990 muitas pesquisas referentes à modelagem de reflexões e métodos de auralização foram conduzidas na *Chalmers University of Technology*, Suécia [17]. O programa de modelagem acústica de ambientes *CATT Acoustic* [18], que inclui auralização, foi desenvolvido pelo Dr. Bengt-Inge Dalenbäck em cooperação com cientistas da universidade.

Na *Technical University of Denmark*, durante a última década, as pesquisas se concentraram na modelagem e auralização de salas de concerto [7] que culminaram na implementação de um programa de modelagem acústica: o *Odeon* [19, 20].

A partir de meados da última década, na *Università di Parma*, Itália, os pesquisadores voltaram suas atenções para a acústica automotiva, sistemas de som para automóveis e acústica de salas de concerto [21, 22]. Destes estudos surgiu um programa de modelagem acústica de ambientes chamado *Ramset* [22].

Estudos relacionados à computação gráfica, métodos de modelagem de ambientes acústicos e algoritmos de renderização sonora em tempo real estão em desenvolvimento desde o ano de 1997 na Universidade de Princeton e nos laboratórios Bell, EUA, [6, 23].

Os programas *EASE*, simulação acústica, e posteriormente o *EARS*, auralização, foram desenvolvidos [24] para simular e auralizar, respectivamente, um ambiente durante a década de 1980 e no início dos anos 90.

Mais recentemente, na *Helsinki University of Technology*, Finlândia, um grupo de pesquisadores desenvolveu o *DIVA*, *Digital Interactive Virtual Acoustics*, que oferece auralização em tempo real para o usuário interferindo interativamente na posição do ouvinte no ambiente virtual [8, 25].

Cabe aqui ressaltar que apesar da idéia ser razoavelmente antiga, o termo *auralização* só foi introduzido por Mendel Kleiner no ano de 1989 na convecção anual da *sociedade de engenharia de áudio*, *AES*. Segundo o próprio autor [5] esta palavra é análoga à visualização no que se refere a capacidade de "antevisão" da acústica de um ambiente sem que se esteja nele ou sequer que ele exista.

3.3 Técnicas Atuais de Auralização

O antigo sonho dos pesquisadores da acústica de obter uma impressão auditiva relacionada a um determinado espaço a priori de sua construção já pode ser realizado.

3. Auralização

A simulação binaural de um ambiente para a predição das propriedades acústicas e para *auralização* pode ser executada com modelos computacionais ou físicos. Atualmente existem quatro técnicas básicas de auralização [5]:

- *Auralização inteiramente computacional* - Toda a transmissão sonora e as propriedades acústicas do local em estudo são modeladas em ambiente digital. Desta forma calcula-se a *RIR* do ambiente. Levando-se em conta as características de transmissão da *HRTF* é determinada a *BRIR* do espaço para uma determinada posição e orientação do ouvinte. A ambiência pode ser percebida por intermédio da convolução no computador, em um programa, desta resposta impulsiva e do sinal anecóico. A apresentação deste resultado é feita utilizando sistemas binaurais, via fones de ouvido ou transaurais (reprodução via alto-falantes com processamento adequado para cancelar o sinal de um emissor na orelha oposta).
- *Auralização computacional com múltiplos alto-falantes* - Aqui, a predição da *RIR* é feita da mesma forma que na *auralização inteiramente computacional*. Porém, a convolução é multi-canal permitindo, assim, que um conjunto de alto-falantes projete cada sinal numa direção específica em relação ao ouvinte criando uma direcionalidade mais natural do campo sonoro. Desta forma, não há modelagem do ouvinte, ou seja, a *HRTF* não é utilizada para a construção e uso da *BRIR*.
- *Auralização indireta por modelo acústico em escala* - Esta técnica utiliza modelos acústicos em escala e técnicas ultrasônicas para medir a *BRIR* de um determinado ambiente. Assim a convolução desta resposta impulsiva com o sinal de interesse pode ser feita como na *auralização inteiramente computacional*.
- *Auralização direta por modelo acústico em escala* - Este método, menos flexível que o indireto, difere deste apenas pelo fato de que a convolução é realizada em “tempo real”. O sinal desejado, devidamente escalonado, é executado no modelo, captado, convertido novamente para a escala normal, e escutado sem possibilidade de medição e armazenagem da *BRIR*.

Apesar da diversidade dos sistemas, a tendência atual é a utilização da *auralização inteiramente computacional* ao invés dos modelos em escala e outros métodos. Isto se deve ao fato de que o sistema baseado puramente no ambiente computacional é muito mais flexível, ou seja, passível de mudanças no que se refere às características da fonte, do receptor e principalmente na geometria e nos parâmetros do ambiente [7].

3.3.1 Auralização inteiramente computacional

Após décadas de pesquisa e desenvolvimento os ambientes acústicos virtuais amadureceram [7]. O uso do computador na recriação das características aurais de espaços tem gerado frutos significativos [18, 19, 25]. A auralização baseada em ambiente puramente digital é, hoje em dia, a ferramenta mais poderosa no auxílio à construção de eventos auditivos realmente convincentes, relacionados a uma fonte num determinado local.

Em todo sistema de auralização computacional completa existe um encadeamento básico denominado cadeia de auralização [25, 26], conforme a figura (3.1).

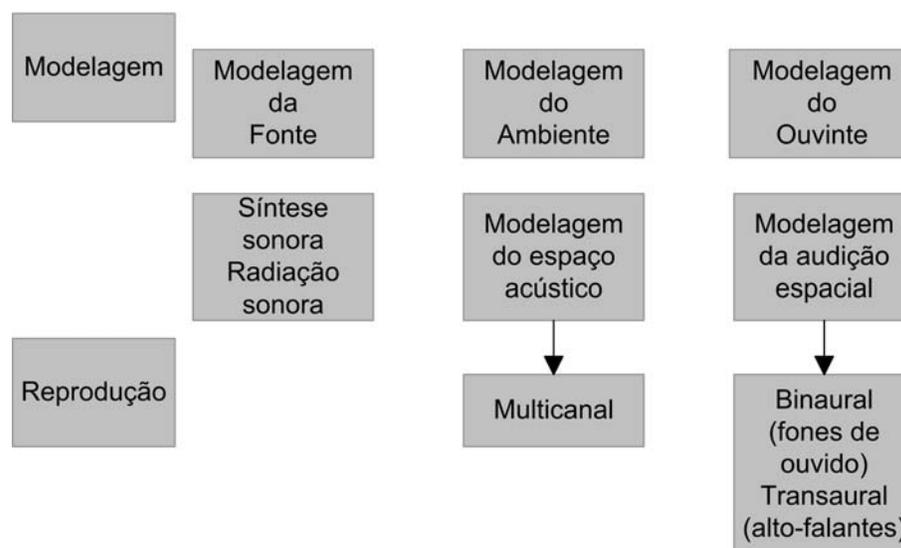


Figura 3.1: Cadeia de Auralização

Estes três componentes, que formam uma típica cadeia fonte-meio-receptor de um sistema de comunicação [27], devem ser modelados para possibilitar a constituição de um ambiente acústico virtual. As propriedades acústicas das superfícies e suas coordenadas são inseridas no computador, assim como os dados referentes à fonte sonora e o ouvinte. Com as devidas modelagens a *RIR* é calculada e em seguida a *BRIR* para que se possa efetuar a convolução com o sinal anecóico desejado. O áudio filtrado pela *resposta impulsiva binaural do ambiente* pode então ser ouvido. Esse sistema de auralização encontra-se representado na figura 3.2 [5].

Fonte sonora

Na primeira etapa da cadeia de auralização, ou seja, na modelagem da fonte sonora, há uma distinção básica entre a natureza dos transmissores. Alto-falantes

3. Auralização

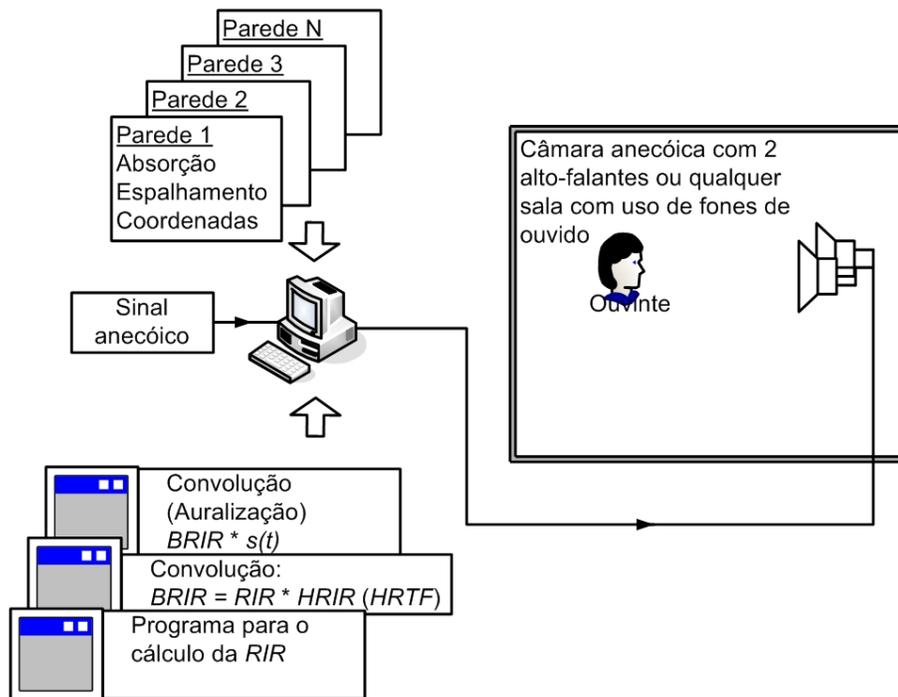


Figura 3.2: *Auralização Computacional Completa*

e *fontes naturais*, tais como instrumentos musicais e seres humanos, possuem propriedades bem distintas e necessitam, então, de abordagens diferentes para um modelamento adequado. O primeiro grupo possui suas características de diretividade e resposta em frequência bem determinadas e são, por isso, não variantes com respeito ao tempo. Já o segundo, é um dos maiores problemas no processo de auralização de auditórios. A descrição física das particularidades das *fontes sonoras naturais* é complicada, não somente devido à complexidade acústica destas, mas também devido a sua variação temporal e diferenças entre o número de transmissores em questão [5]. Por exemplo, imaginemos a riqueza de detalhes em relação aos fenômenos ondulatórios que emergem se as nossas *fontes naturais* constituem uma orquestra. A *difração* e o *espalhamento* introduzidos pelos músicos não podem ser negligenciados e ainda por cima variam de acordo com a performance.

Tendo em vista as distinções entre as naturezas das fontes, o próximo passo é imaginar como o áudio emitido por estas será introduzido no ambiente virtual. As abordagens mais simples são as gravações anecóicas dos sinais desejados ou a síntese sonora [26]. Dentre as diversas técnicas que são utilizadas para esta geração podemos citar como exemplo de síntese a modelagem física de instrumentos musicais com algoritmos computacionais [28], que tenta imitar o processo físico de

construção sonora. Mais adiante, no capítulo 5, discutiremos a respeito deste e outros métodos comumente utilizados.

Uma das características das fontes sônicas que deve ser levada em consideração, para fins de auralização, e que está intrinsecamente ligada às suas dimensões e geometria é a diretividade. A abordagem mais simples é considerar um emissor sonoro como sendo um ponto com emissão de ondas omnidirecional. Tipicamente esse não é o caso, já que as fontes geralmente possuem diretividade dependente da frequência de operação o que indica que esta propriedade deve ser modelada para que se possa alcançar resultados reais. Por exemplo, um instrumento musical como um violão além de possuir características de radiação bastante complexas, devido ao fato de sua energia acústica derivar dos seus modos de ressonância, ainda sofre influência do corpo do músico que introduz reflexões e mascaramentos contribuindo, assim, para características de diretividade do transmissor. Em ambiente computacional os filtros são utilizados para modelar o comportamento direcional desejado [25].

Meio de propagação sonora

A modelagem do meio, segundo elo da cadeia, ou do espaço acústico propriamente dito, leva em conta a propagação do som no ar e as suas interações com as superfícies tais quais paredes, teto, piso, e objetos que tenham dimensões relevantes em comparação com os comprimentos de onda do campo sonoro. Podemos pensar numa situação geral de audição em um ambiente como sendo um sistema linear invariante no tempo (*SLIT*) se a temperatura e a umidade relativa do ar se mantiverem constantes, se não houver correntes de ar e assumindo que as fontes, o receptor e outros objetos não estão em movimento no momento da observação [12]. Desta forma, a clássica abordagem de caracterização de um *SLIT* por intermédio de sua resposta ao impulso pode ser utilizada.

Com algoritmos que permitem a determinação de todos os *caminhos de reverberação* (histórico de uma reflexão) do transmissor até o ouvinte uma resposta ao impulso, entre estes, pode ser determinada. Matematicamente a propagação sonora é descrita pela equação de onda ou equação de *Helmholtz* [3]. Teoricamente uma resposta impulsiva entre uma fonte e um receptor pode ser obtida por intermédio da solução da equação de onda mas analiticamente isto é muito custoso. Assim, outros métodos de modelagem são usados para evitar este problema. Atualmente os mais utilizados são:

- métodos geométricos

3. Auralização

- métodos baseados na equação de onda
- métodos estatísticos

Os métodos geométricos, que são baseados na acústica geométrica, são o método do *Traçado de Raios*, o método das *Imagens Virtuais da Fonte* e o método do *Traçado de Pirâmides*. No traçado de raios uma fonte pontual emite um grande número de partículas em várias direções. Essas partículas, em seu movimento ao longo de raios sonoros, vão perdendo energia por absorção do meio e reflexão com as superfícies. Aquelas que atingem o receptor serão computadas em relação à sua energia, tempo de chegada e direção. Estas informações podem por exemplo contribuir para a formação de um *ecograma* que a depender da quantidade de partículas emitidas e da resolução temporal pode ser aproximado pela resposta impulsiva da sala [29].

A diferença substantiva entre o método anterior e o das Imagens da Fonte fica evidente já no modo como este define os caminhos das reflexões dos raios acústicos que chegam ao receptor. Esta técnica avalia os trajetos refletidos especularmente considerando os caminhos diretos de raios emitidos por *fontes virtuais* que são geradas pelo espelhamento da posição da fonte original sobre cada superfície poligonal do ambiente [15]. Desta forma a constelação de fontes espelhadas que é gerada passa a representar as interações da onda sonora com as paredes da sala determinando todos os trajetos até o ouvinte. Excitando, assim, todos os transmissores, real e virtuais, com o impulso de Dirac, pode-se obter a resposta impulsiva do espaço em estudo em um ponto determinado.

No método do traçado de Pirâmides os raios são substituídos por feixes ou seja um conjunto de raios. Assim os caminhos de propagação entre a fonte sonora e o receptor são encontrados com a transmissão do feixe piramidal. As superfícies de interseção são poligonais e sempre que um feixe as atinge ele é recortado removendo-se assim a região coberta pelo feixe na outra face do polígono. Dessa forma são construídos um feixe de transmissão sobre a região removida e um feixe de reflexão pelo espelhamento do feixe de transmissão sobre a superfície poligonal de incidência [6]. No vértice de cada pirâmide originada por reflexão encontra-se uma fonte virtual que corresponde, exatamente, à região do espaço coberta pelo feixe. Desta forma tem-se a certeza de que aquela fonte é "visível" para todos os pontos englobados pelo traçado piramidal correspondente.

Os métodos baseados na equação de onda são mais precisos quando usados para prever o comportamento de um campo sonoro do que os geométricos. Isto é justificável pois essas soluções numéricas da equação de onda levam em conta os fenômenos ondulatórios tais como a difração e a reflexão difusa. Comumente, são utilizados métodos de elementos e métodos de diferenças finitas [5, 25]. Em acústica, o *Método de Elementos Finitos (FEM)* e o *Método de Elementos de Contorno (BEM)* trazem como idéia central a subdivisão, do espaço em questão, em elementos e o cálculo da equação de onda para cada um destes. A diferença básica entre os dois é justamente a discretização do ambiente. O *BEM* discretiza apenas as superfícies do local enquanto no *FEM* todo o espaço é discretizado.

Nos métodos das *Diferenças Finitas no Domínio do Tempo (FDTD)*, outra técnica para simulação de ambientes acústicos é utilizada. A idéia central é a substituição das derivadas na equação de onda por diferenças finitas [25]. O método mais utilizado, dentre estes, é o *Digital Waveguide Mesh*. O espaço é dividido em partes denominadas junções que estarão ligadas por meio de guias de ondas bidirecionais [30]. Estes guias são modelados, em ambiente computacional, por intermédio de *linhas de atraso digitais (Digital Delay Lines)*.

O terceiro tipo de abordagem para modelamento do espaço acústico são os Métodos Estatísticos. Na verdade, eles não são utilizados isoladamente e sim sempre combinados com um dos outros apresentados para conferir características específicas à simulação ou usados para predizer etapas da evolução temporal de um campo sonoro que possuem um comportamento aleatório.

Geralmente as técnicas estatísticas irão fornecer partes da *RIR* que representam o trecho em que o campo sonoro passa a independer dos atributos geométricos do ambiente ou seja onde, temporalmente, o campo se torna difuso [4]. Conhecida como *cauda reverberante*, esta partição da resposta impulsiva de um ambiente não precisa ser determinada de forma precisa. Assim, somente a estrutura macroscópica temporal e espectral é estatisticamente avaliada resultando no decaimento de energia compatível com espaços reais.

Alguns métodos implementam fenômenos ondulatórios como reflexões difusas por intermédio de métodos estatísticos [31]. No capítulo 4 trataremos das aplicações desta e das outras técnicas de modelagem de ambientes, suas vantagens e desvantagens, características específicas com relação à faixa de frequência, dimensões e formas geométricas do ambiente, esforço computacional, considerações sobre os fenômenos

3. Auralização

ondulatórios, limitações no que se refere à interatividade em tempo real de um sistema de auralização que balise seus algoritmos em uma delas e da conjugação de algumas destas em metodologias híbridas, capazes de representar diversos casos de simulação de campos sonoros em salas.

Receptor e apresentação do produto da auralização

Para fins de auralização, a modelagem do receptor é baseada em propriedades tais como a diretividade, impostas pelo conjunto tímpano, canal auditivo, orelha externa, cabeça e torso, aos sinais que alcançam o ouvinte [2]. Uma função de transferência é utilizada para representar as distorções lineares introduzidas pelos elementos acima. A *HRTF* (Função de Transferência Relacionada à Cabeça) representa, então, o filtro que codifica o sinal acústico do *evento sonoro* gerado pela fonte no ponto de audição para que este seja compatível com um *evento auditivo* real. Existem diversos métodos para aquisição da *HRTF* e estes serão explorados no capítulo 5.

Em ambientes computacionais as propriedades da *HRTF* são modeladas com filtros digitais. Desta forma, estes levam em conta em seus processos de transformações lineares a direção de incidência, distância e a frequência dos sinais originados pela fonte sonora. Para cada caminho de propagação das ondas sonoras que atingem o ouvinte o filtro *HRTF* deve ser aplicado.

Uma vez que a *função de transferência da cabeça* e a *resposta impulsiva do ambiente* foram levantadas, elas serão utilizadas para a confecção da *BRIR*, que nada mais é do que a resposta ao impulso de todo o sistema modelado. Fazendo a convolução desta *resposta impulsiva binaural do ambiente (BRIR)* com o sinal “seco”, ou anecóico, obtém-se os dois canais binaurais de saída deste sistema. A partir destes sinais o resultado da auralização pode ser avaliado por intermédio de fones de ouvido, técnica binaural. Pode-se, ainda, apresentar os canais num sistema transaural que consiste numa reprodução em câmara anecóica com um par de alto-falantes contendo, em sua entrada, filtros de *cancelamento de crosstalk* (supressão do sinal do alto-falante esquerdo para a orelha direita e vice-versa).

3.3.2 Auralização computacional com múltiplos alto-falantes

Todos os princípios de modelagem da fonte e do meio citados no item anterior são utilizados nesta técnica. A diferença básica está no fato de que não há a necessidade de modelamento da *HRTF* para a representação das características do ouvinte. Ao invés disso, um conjunto de alto-falantes irá projetar, individualmente, os sinais acústicos correspondentes a cada direção de incidência. Nota-se, com isso, que a

convolução deve ser realizada para cada canal do sistema de reprodução. Desta forma o campo sonoro apresentado ao ouvinte possui características direcionais mais naturais do que no método computacional completo. Os sons que devem atingir o receptor por trás, realmente virão desta direção evitando, assim, problemas como *IHL* (*localização no interior da cabeça*) e *reversões frente/costas* [2].

Apesar da aparente vantagem deste sistema ele é mais sofisticado, tanto em aparelhagem quanto em custo, do que o anterior. O conjunto de alto-falantes deve estar numa câmara anecóica para que não haja influência da reverberação do ambiente no momento da audição. O aparato dos falantes é outro problema, pois, além de numerosos, estes precisam ser pequenos, para evitar contaminação do campo simulado por fenômenos como a *difração* e o *espalhamento*, e, além disto, os transdutores devem possuir resposta em frequência excelente [5, 32, 33].

No presente trabalho elegemos o aprofundamento da *auralização inteiramente computacional* utilizando técnicas binaurais, pois, esta demonstra maior flexibilidade no que se refere não só a investimentos financeiros mas, também, em relação ao seu uso fora dos laboratórios o que está de acordo com as motivações deste trabalho.

3.3.3 *Auralização indireta com modelo acústico em escala*

Como já citado anteriormente, desde o início do século passado, os estudiosos da acústica se utilizam de modelos acústicos em escala na tentativa de obter impressões auditivas reais de um ambiente determinado. A idéia central que rege este sistema é a confecção de um modelo em escala reduzida e a aplicação de técnicas ultra-sônicas, ou seja, aumento do valor das frequências e implementação de fontes e receptores compatíveis, para simular a transmissão sonora do ambiente real. Dois pré-requisitos devem ser observados para que uma correta modelagem seja efetuada: a similaridade geométrica e a acústica em relação à sala original. Para que isto fosse garantido Spandöck [12] formulou algumas regras que devem ser observadas para a construção do modelo reduzido. As leis do escalonamento foram derivadas, originalmente, da redução do comprimento de onda do campo sonoro do modelo assumindo que a velocidade do som fosse constante tanto para o espaço original quanto para o de dimensões inferiores. Na tabela (3.1) encontram-se listadas as principais regras supracitadas.

Normalmente o fator de escala assume valores entre 5 e 40 [12]. Obviamente, quanto maior o fator de escala menor as dimensões lineares e conseqüentemente maior o cuidado com os detalhes geométricos na confecção do espaço. Há ainda um

3. Auralização

Tabela 3.1: Critérios para os modelos físicos em escala. n é o fator de escala que relaciona o ambiente original (subscrito o) e o modelado (subscrito m).

parâmetro	relação
comprimento de onda	$\lambda_0/\lambda_m = n$
dimensão linear	$l_0/l_m = n$
frequência	$f_m/f_0 = n$
tempo	$T_0/T_m = n$
coeficiente de absorção das superfícies	$\alpha_m(f_m) = \alpha_0(f_0)$
coeficiente de absorção do meio	$\xi_m(f_m) = n \cdot \xi_0(f_0)$

outro ponto que limita o fator de escala que é a implementação da instrumentação necessária, pois, com as altas frequências que devem ser alcançadas para a utilização adequada do modelo, as fontes e os receptores acabam por se tornar extremamente complexos tecnicamente.

A obtenção da similaridade geométrica é conseguida já no estágio de construção do modelo reduzido baseada no fator de escala selecionado de acordo com o objetivo da simulação. É necessário, então, notar que há um compromisso entre o grau de precisão do experimento e os seus custos. Já a similaridade acústica deve ser encarada com maiores cuidados, pois esta irá determinar as características do meio de propagação e das superfícies do espaço em questão. Normalmente o modelo rudimentar é construído em compensado e o seu interior é envernizado para simular paredes de concreto e tijolos. Desta forma, as características de absorção e reflexão podem ser inseridas adicionando materiais acústicos apropriados às superfícies internas.

Meio de propagação sonora

Baseado nas regras do escalonamento, a faixa de frequências em que se pretende trabalhar deve ser transformada. No caso de performances musicais esta faixa deve iniciar em algumas dezenas de Hertz e se estender até, no mínimo, 16 KHz. Se for considerado um fator de escala de 10, como, por exemplo, em Xiang e Blauert [12], teremos um limite superior de 160 KHz que deve ser garantido. Porém, a absorção do ar para as frequências ultra-sônicas é extremamente elevada em comparação com a sofrida pelas frequências audíveis. Para remediar esta situação dois métodos podem ser utilizados: ar seco (baixa humidade relativa) e substituição do ar por outro meio. Em [12] os autores utilizaram nitrogênio para garantir uma absorção consistente com a que seria percebida no ambiente real.

Correções matemáticas podem ser implementadas em alguns casos para uma correta absorção por parte do meio. Porém, para fins de auralização, a substituição é o método mais confiável, pois a relação sinal/ruído da *BRIR* medida permite uma medição fidedigna desta [12].

Superfícies e auditório

As características de absorção de um ambiente são dadas pelas superfícies compostas por paredes, teto, piso e pelo auditório. No modelo em escala estas propriedades devem ser reproduzidas na faixa de frequência escalonada. Vários materiais são utilizados em acústica arquitetônica e seus respectivos coeficientes de absorção são fornecidos pelos seus fabricantes após medidas realizadas em câmaras de reverberação. Um procedimento semelhante é utilizado para a medição dos coeficientes nas frequências do modelo reduzido, mas as medidas são feitas em câmaras de reverberação consistentes com os ambientes modelados, ou seja, câmaras com dimensões lineares reduzidas. Aplica-se nestas, também, o mesmo método de compensação para a correção da absorção do meio.

Em uma sala de execução musical, normalmente, o piso encontra-se coberto pelo auditório. Logo as características absorvedoras do público devem ser levadas em conta na modelagem. Assim, um modelo de ouvinte que encerre não somente características de absorção mas também de difração de uma pessoa real deve ser implementado. Simplificações na forma geométrica desses modelos podem ser efetuadas. N. Xiang e J. Blauert [12] confeccionaram modelos de pessoas em isopor cujos ajustes do coeficiente de absorção foram realizados em uma câmara de reverberação em miniatura.

Receptores

A HRTF deve ser levada em conta na medida das respostas impulsivas para que haja uma auralização correta do evento auditivo. Assim os receptores sonoros devem possuir as características de transferência sonora do conjunto torso, cabeça e pinna. Então uma *cabeça artificial* [12] adequada e em miniatura deve ser utilizada para simular estas propriedades na faixa de frequência escalonada.

Fontes sonoras

As fontes utilizadas nos modelos em escala podem ser divididas em dois tipos: fontes contínuas e impulsivas. Estas últimas, tipicamente, promovem uma descarga elétrica que gera o conteúdo sonoro com o espectro de frequências rico. Apesar de

3. Auralização

poderem ser construídas em dimensões a ponto de serem consideradas fontes pontuais, elas produzem uma energia sonora inadequada de tal forma que uma grande quantidade de médias deve ser efetuada para obtenção da relação sinal/ruído desejada.

Ao contrário das impulsivas, as fontes sonoras contínuas são geralmente transdutores eletroacústicos de diversos tipos. Uma de suas vantagens é a possibilidade de radiar sinais sonoros adequados à aplicação do experimento. Essas fontes devem ser desenvolvidas para atender as seguintes exigências:

- Eficiência alta
- Resposta em frequência plana dentro da faixa em questão
- Linearidade adequada
- Características direcionais desejadas, tais como, omnidirecionalidade.

As dimensões do transdutor devem ser menores que os comprimentos de onda produzidos. Já eficiências altas requerem transdutores maiores. Isto limita as exigências listadas acima a uma relação de compromisso do incremento de umas em detrimento das outras. Para solucionar estes problemas foram desenvolvidos transdutores ultra-sônicos como por exemplo em [12]. Usando filmes piezoelétricos de *PVDF* (polifluoreto de vinilideno), que possuem entre tantas vantajosas propriedades a flexibilidade mecânica, alta constante piezoelétrica, peso reduzido, além do baixo custo, foi confeccionado um transdutor que atende a maior parte da faixa de frequências altas necessárias ao modelo acústico em escala com a característica de direcionalidade desejada. Alto-falantes dinâmicos podem ser utilizados para cobrir as frequências baixas.

3.3.4 *Auralização direta com modelo acústico em escala*

Esta abordagem para auralização, que foi inaugurada por Spandöck [12], também se fundamenta nos modelos físicos em escala e todas as suas regras de confecção anteriormente citadas. Porém, ela se utiliza do sinal escalonado diretamente, ou seja, a convolução ocorre em "tempo real". Isto quer dizer que esta técnica é menos sofisticada do que o método indireto, pois, não é seu objetivo medir a *BRIR* e sim ouvir o sinal diretamente logo após a conversão para as frequências normais. Para além disso, ela é menos flexível. No caso indireto a relação sinal/ruído pode ser melhorada se médias forem efetuadas, distorções lineares e não-lineares podem ser

reduzidas e, além da medição da *BRIR*, a possibilidade desta ser armazenada torna-se factível o que acaba sendo muito mais conveniente que armazenagem de modelos físicos.

CAPÍTULO 4

Auralização Inteiramente Computacional

A *auralização inteiramente computacional* define sistemas que se utilizam de modelagens puramente digitais da fonte, do ambiente e do ouvinte. Estas duas últimas são baseadas no conceito de resposta impulsiva. A forma de apresentação das pressões sonoras para o ouvinte, resultantes da simulação *binaural*, pressupõe fones de ouvido ou um sistema estéreo com processamento adequado (*transaural*). A seguir, discutiremos as principais características das *respostas impulsivas acústicas de ambientes*, parâmetros objetivos que podem ser calculados a partir destas, métodos utilizados para determinar estas respostas, e a sua utilização para conferir a assinatura acústica dos espaços aos sinais que se deseja simular.

4.1 Estrutura da *RIR*

Como já visto anteriormente, todo ambiente acústico cujas condições atmosféricas, temperatura e pressão de equilíbrio, se mantém constantes possui características próprias de propagação sonora que, desta forma, não variam ao longo do tempo. Todo sistema linear invariante no tempo pode ser caracterizado por intermédio de sua resposta impulsiva.

Por definição, uma resposta ao impulso num ambiente acústico é calculada para um par fonte-receptor [4]. Num ambiente acústico a resposta entre uma fonte simples e uma determinada posição de recepção descreve completamente o sistema acústico entre estes dois pontos. Deve-se deixar claro que tal resultado não pode ser generalizado para representar todo o sistema acústico do espaço. Quando, no interior de um recinto, um indivíduo bate palmas uma única vez produzindo um som de curta duração ele escuta, aproximadamente, a "resposta ao impulso" deste lugar. Esta audição já possibilita uma avaliação perceptual em relação à característica reverberante da sala.

4. Auralização Inteiramente Computacional

O método de caracterização de um espaço pela RIR é uma ferramenta muito versátil e poderosa para a modelagem e simulação de ambientes acústicos. Ele permite que parâmetros acústicos objetivos sejam extraídos da resposta impulsiva e que a *auralização* seja executada.

4.1.1 Decomposição da RIR em partes

A evolução temporal da resposta impulsiva de um ambiente é marcada, basicamente, por três fases (fig. 4.1). A primeira é composta pelo *som direto* que, como o próprio nome sugere, é o primeiro sinal emitido pela fonte que atinge o receptor sem que haja reflexão em qualquer superfície. A segunda fase é composta pelas chamadas *reflexões iniciais*, que são os sons refletidos pelas paredes que chegam no ponto de recepção e essas chegadas ainda podem ser distinguidas temporalmente, uma a uma, pelo sistema auditivo [34]. Essa etapa da RIR fornece ao ouvinte a maior parte das informações espaciais a respeito do ambiente devido à capacidade de discriminação das direcionalidades, de distinção dos tempos de chegada e da intensidade relativamente alta dessas contribuições [34]. Vale ressaltar que apesar dessa fase ser dominada por *reflexões especulares*, há componentes de *reflexão difusa* e *difração*. A última, nem por isso menos importante, é a *cauda reverberante* ou *reverberação* propriamente dita que representa a porção temporal na qual o campo sonoro, após inúmeras reflexões, se torna bastante difuso e denso, inviabilizando a distinção dos sinais que chegam ao ouvido.

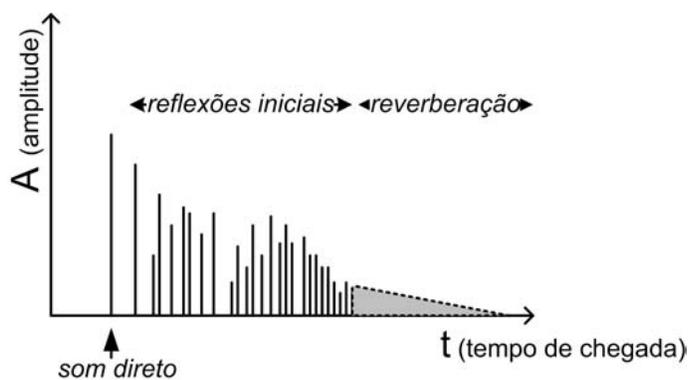


Figura 4.1: *Evolução temporal de uma resposta impulsiva de ambientes*

Esta abordagem de estudo da composição temporal de uma RIR se torna muito interessante quando se observa que cada parte desta resposta é interpretada perceptualmente pelo sistema auditivo de forma diferenciada. A estrutura temporal e

direcional do *som direto* e das *reflexões iniciais* é de extrema importância para a sensação de *volume*, *clareza*, e *espacialidade*. É também determinante para as variações das condições de audição em partes diferentes de um recinto [4]. Já a *cauda reverberante*, que possui grande densidade de reflexões, está associada a impressão auditiva muito mais pelas suas características temporal e espectral gerais do que pela sua estrutura fina [29]. Desta forma, do ponto de vista perceptual não há necessidade do mesmo detalhamento no cálculo de todas as partes da RIR para a construção de um evento auditivo convincente. Assim podemos dizer que há duas abordagens básicas que podem ser utilizadas para o cálculo das diferentes fases de evolução no tempo de um campo sonoro numa sala: as determinísticas e as estatísticas.

O *som direto* e as *reflexões iniciais* são computados com métodos determinísticos, já que a estrutura fina destas etapas deve estar bem especificada para uma correta *auralização*. Fazem parte destes, os *métodos geométricos* e os *baseados em solução numérica da equação de onda*.

As abordagens estatísticas são utilizadas para modelar a partição difusa da resposta impulsiva que fornece informações gerais sobre o decaimento do campo sonoro num ambiente.

Como consequência direta desta forma de entender e trabalhar a resposta impulsiva está o fato de que os esforços referentes aos cálculos computacionais se tornam muito menores àqueles que teríamos se, por exemplo, toda a RIR fosse calculada por intermédio de métodos determinísticos. Quando tratarmos das técnicas de modelagem de ambientes acústicos, alguns dos parâmetros que permitem comparações entre estas técnicas, como o esforço computacional, esclarecerão este fato.

Obviamente se forem utilizados métodos diferenciados para o cálculo das partes que compõem a RIR surge a necessidade de decisão em relação ao momento de transição entre uma abordagem e outra, ou seja, em que ponto da resposta impulsiva comutaremos de um método determinístico para um estocástico. Essa decisão dependerá da evolução temporal do campo sonoro no ambiente em questão. Quando este tornar-se difuso, não mais haverá a necessidade de detalhamento da estrutura deste sinal. A estatística será, a partir deste ponto escolhido, suficiente para modelar o comportamento da resposta ao impulso. Esta transição pode ser balizada em alguns critérios de truncamento da parte determinística da *RIR*. Estes podem ser baseados no comprimento da resposta impulsiva do ambiente, na energia, na ordem máxima de reflexões a ser considerada ou até numa combinação destes.

4. Auralização Inteiramente Computacional

Apesar de serem critérios importantes a ordem máxima de reflexões e a energia não são satisfatórios. A energia individual dos raios decresce com o passar do tempo mas a densidade temporal das reflexões cresce. Assim um critério de truncamento absolutamente baseado num destes é difícil de se estabelecer. Um dos efeitos que podem acontecer refere-se a uma determinada reflexão que venha de uma direção perceptualmente relevante ou possua características espectrais particularmente audíveis podendo ser percebida individualmente, muito embora sua energia total seja relativamente baixa.

O comprimento máximo da parte determinística da *RIR* é um critério mais apropriado para o truncamento. Ele pode ser estimado por intermédio do tempo de reverberação da sala ou por outra evidência perceptual relevante. Kuttruff [29] sugere uma faixa entre 100 - 150ms a partir da qual, para qualquer sala, o campo sonoro pode ser considerado como difuso e um tempo característico pode ser definido para separar as *reflexões iniciais* da reverberação dentro desta faixa.

Após a obtenção das partições da resposta ao impulso acústico será preciso juntar os sinais de modo a formar a *RIR* completa, que descreve toda a evolução temporal do campo sonoro para um par de pontos num espaço, sem que haja distúrbios audíveis. Algumas condições devem ser obedecidas para garantir que esta conexão seja bem sucedida [31]. A primeira delas, é garantir que as duas partes sejam conectadas suavemente uma à outra. Isso pode ser realizado fazendo uma junção numa transição de zero não-direcional (fig. 4.2). Na segunda, deve-se observar se há grandes diferenças espectrais entre os sinais no ponto de conexão, pois o ouvido humano é bastante sensível a tais variações. A melhor maneira de evitar tais problemas é utilizar métodos semelhantes para a síntese das diferentes secções da RI ou fazer uma região de sobreposição ou faixa de assimilação entre as partes preditas.

4.1.2 Parâmetros acústicos extraídos da RIR

Para além de permitir que o comportamento acústico de um ambiente em face de uma determinada excitação sonora seja escutado, a resposta impulsiva de um determinado espaço nos fornece critérios acústicos que são de importantes para a previsão de como um sinal emitido será afetado [4]. A maioria dos parâmetros padronizados, que caracterizam acusticamente um local, podem ser derivados da *RIR*. Abaixo, encontra-se uma tabela que lista os principais critérios resultantes da resposta impulsiva (tab. 4.1).

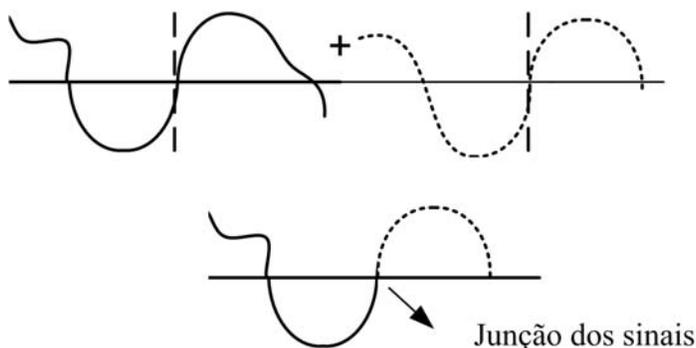


Figura 4.2: Conexão de dois sinais em transição de zero não-direcional

Tabela 4.1: Parâmetros acústicos derivados da *RIR*

parâmetro	símbolo	aspecto subjetivo	fonte
tempo de reverberação (s)	RT_{60}	reverberação	ISO 3382
tempo de queda inicial (s)	T_{10}	reverberação, clareza	ISO 3382
tempo central (s)	T_S	clareza	ISO 3382
clareza (dB)	C	clareza	ISO 3382
definição (dB)	D	definição	ISO 3382
nível sonoro relativo (dB)	G	nível relativo	ISO 3382
fração de energia lateral (%)	LEF, L_F	impressão espacial	ISO 3382
coeficiente de correlção cruzada interaural	$IACC$	impressão espacial	ISO 3382

Tempo de reverberação - RT_{60}

Definido como o tempo necessário para que a energia num recinto decresça 60 dB do seu valor inicial após o súbito desligamento da emissão de uma fonte, o tempo de reverberação é um dos parâmetros acústicos mais importantes derivados da *RIR*. O RT_{60} é obtido das *curvas de decaimento* resultantes da *integração de Schroeder* [4]: a resposta impulsiva $h(t)$ é elevada ao quadrado $h^2(t)$, resultando em energia acústica, passa por um processo de integração reversa que origina as *curvas de decaimento*. Uma regressão linear é aplicada para que seja determinada a melhor reta que se adapta às curvas e a partir desta, já com os níveis convertidos para decibéis, é determinado o tempo de decaimento. A equação 4.1 define formalmente a relação entre a *RIR* e as *curvas de decaimento* $y(t)$:

4. Auralização Inteiramente Computacional

$$\langle y^2(t) \rangle_{conjunto} = \int_t^\infty h^2(x)dx = \int_0^\infty h^2(x)dx - \int_0^t h^2(x)dx \quad (4.1)$$

Tempo de queda inicial - EDT

O *EDT* (*Early Decay Time*) é definido como o intervalo de tempo requerido, por um processo de decaimento do nível de energia sonora, para atingir uma queda de 10 dB após o término do sinal de excitação [4]. Muito utilizado nos dias atuais, pode-se notar que este parâmetro caracteriza a taxa de queda sonora na porção inicial do processo. Decaimentos de energia acústica encontrados em situações reais não seguem uma lei exponencial e sim possuem taxas diferentes para vários trechos, contrariando a definição do RT_{60} . Mesmo assim, eles podem ser caracterizados pela sua seção inicial que é a mais significativa para a reverberação percebida subjetivamente [4].

Para uma comparação entre o RT_{60} e o *EDT* basta que este seja multiplicado por um fator igual a 6. Num processo ideal de queda, ou seja, com decaimento exponencial, o valor esperado do *EDT* vezes o fator é igual ao RT_{60} . O seu cálculo é similar ao do *Tempo de Reverberação* porém com os níveis para ajuste da “melhor reta” fixados entre 0 e 10dB.

É importante saber que o *Tempo de Reverberação* é um parâmetro que não varia de acordo com a geometria do ambiente em questão nem com a posição dentro deste. Isto é facilmente observado da sua própria definição. Já o *EDT*, é extremamente sensível tanto à geometria quanto à posição, pois, a porção inicial da curva de queda energética, sobre a qual ele é avaliado, é constituída, principalmente, por algumas reflexões isoladas e intensas.

Clareza - C_{50} e C_{80}

O índice de *Clareza* (*Clarity*) [4] é mais um parâmetro acústico objetivo derivado da resposta impulsiva. Ele mede a transparência da informação sonora por intermédio da razão logarítmica entre a energia do som inicial (*som direto e reflexões iniciais*) e a energia da *reverberação*.

$$C_{50} = 10 \log \frac{\int_0^{50ms} h^2(t)dt}{\int_{50ms}^\infty h^2(t)dt} \quad (4.2)$$

$$C_{80} = 10 \log \frac{\int_0^{80ms} h^2(t)dt}{\int_{80ms}^\infty h^2(t)dt} \quad (4.3)$$

Os limites superiores das integrais no numerador correspondem aos lapsos temporais para a chegada do *som direto* e das *reflexões iniciais*. Para a fala este intervalo é de 50 *ms* e para a música é de 80 *ms*. Valores altos deste critério indicam uma grande quantidade de energia inicial da *RIR* o que corresponde a sensação subjetiva de clareza da informação. Valores baixos significam campo sonoro com reverberação excessiva.

Definição - D

A *Definição (Definition)* [4] é um parâmetro acústico objetivo que expressa a clareza da fala e da música por intermédio da razão logarítmica entre a energia de uma *RIR* antes de um determinado tempo, onde ocorrem as *reflexões iniciais*, e a energia total desta resposta impulsiva.

$$D_{50} = 10 \log \frac{\int_0^{50ms} h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \quad (4.4)$$

$$D_{80} = 10 \log \frac{\int_0^{80ms} h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \quad (4.5)$$

Assim como na *Clareza* o limite de 50 *ms* se refere à fala e o de 80 *ms* à música.

Tempo central - T_S

O *Tempo Central (Center Time)* [4] pode ser entendido como o centro de gravidade temporal de $h(t)$ e é definido matematicamente com a seguinte equação (4.6):

$$T_S = \frac{\int_0^{\infty} t \cdot h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \quad (4.6)$$

Usualmente dado em milisegundos, um valor baixo de T_S indica que a maior parte da energia chega no início de $h(t)$, sugerindo uma sensação de clareza, enquanto que valores altos sinalizam muita energia reverberante. Este critério raramente fornece alguma informação adicional ao *EDT*.

Nível sonoro relativo - G

O *nível sonoro relativo (strenght)* expressa a influência na percepção do nível sonoro que um ambiente acústico proporciona [4]. Ele é representado pela razão logarítmica entre a energia total da *RIR* e a energia do som direto captado a uma distância de 10 metros (eq. 4.7).

4. Auralização Inteiramente Computacional

$$G = 10 \log \frac{\int_0^{\infty} h^2(t) dt}{\int_0^{\Delta t} h_0^2(t) dt} \quad (4.7)$$

Onde $h_0(t)$ é a resposta ao impulso à distância de referência (10 metros) e Δt é a duração da excitação, que garante que somente o som direto seja computado. A variação dos valores de G em relação à distância num ambiente indica o grau de difusão de um campo sonoro em seu interior.

Fração da energia lateral - LEF

A *Fração de Energia Lateral* (*lateral energy fraction*), diferentemente dos anteriores, é um parâmetro binaural [4]. Ela se refere à razão entre a *energia inicial* que incide lateralmente ao ouvinte e a energia do *som direto* e das *reflexões iniciais* que chegam omnidirecionalmente (eq. 4.8).

$$LEF = \frac{\int_0^{80ms} h^2(t) \cos \Psi dt}{\int_0^{80ms} h^2(t) dt} \quad (4.8)$$

O ângulo de incidência Ψ é definido entre o eixo que transpassa as orelhas e a direção do som incidente, enquanto que o limite de 80 *ms* se refere a energia inicial. Investigações a respeito da *Espacialidade* [4], determinaram que a contribuição de uma reflexão à sensação de espaço independe do tempo de atraso desta na faixa entre 0 e 80 *ms* e que a causa primordial da sensação de espaço pode ser obtida por poucas reflexões desde que estas sejam incoerentes, relativamente intensas, seus atrasos não excedam 100 *ms* e que elas sejam reflexões laterais.

Correlação cruzada interaural - IACC

Uma boa medida da *Espacialidade* de um campo sonoro dentro de um ambiente são as similaridades e dissimilaridades entre os sinais que chegam aos ouvidos. Por intermédio de uma análise de correlação, e considerando as pressões acústicas dentro dos canais auditivos, teremos para as respostas impulsivas binaurais $h_e(t)$ (orelha esquerda) e $h_d(t)$ (orelha direita) a seguinte função de correlação cruzada [4] (eq. 4.9):

$$IACF(\tau) = \frac{\int_{t_1}^{t_2} h_e(t) h_d(t + \tau) dt}{\sqrt{\int_{t_1}^{t_2} h_e^2(t) dt \int_{t_1}^{t_2} h_d^2(t) dt}} \quad (4.9)$$

Obviamente de funções binaurais, a *HRTF* deve ser utilizada para a determinação da *Resposta Impulsiva Binaural* a partir da *RIR*. A restrição para a porção inicial da resposta impulsiva deve ser feita especificando o limite superior de integração para o

intervalo desejado. A variável τ significa a diferença de tempo de chegada dos sinais para as orelhas e, assim, assume valores entre $-1ms$ e $+1ms$, que corresponde ao atraso máximo entre os sons devido à cabeça.

Com o intuito de obter um número que mesure a similaridade máxima das ondas sonoras que atingem as orelhas o parâmetro denominado *IACC* (*Correlação Cruzada Interaural*) é definido como (eq. 4.10):

$$IACC = \max |IACF(\tau)|; \quad |\tau| < 1ms \quad (4.10)$$

Originalmente, o *IACC* (*Inter-aural cross correlation*) era calculado para toda a extensão da resposta impulsiva. Mais recentemente este critério passou a ser calculado para a porção inicial e a final da *RIR* em separado. Assim, existem vários critérios *IACC* com limites de integração diferenciados dentre os quais podemos destacar o $IACC_{early}$ ($t_1 = 0 ms$ e $t_2 = 80 ms$) que se refere a partição inicial da resposta impulsiva e mede a *ASW* (*Largura Aparente da Fonte*) e o $IACC_{late}$ ($t_1 = 80 ms$ e $t_2 = 750 ms$) que está relacionado à seção final e é uma medida para o *LEV* (*Envolvimento do Ouvinte*) [34].

É importante ressaltar que todos os critérios acústicos discutidos acima, assim como o *EDT*, são extremamente sensíveis à geometria da sala em questão e à posição que é tomada como referência para o seu cálculo. Logo, enquanto o RT_{60} é um parâmetro que descreve de uma forma geral um determinado comportamento acústico de um sistema, os demais critérios representam, apenas, um ponto específico do espaço no interior de tal ambiente.

4.2 Métodos para o Cálculo da *RIR*

Calcular a *resposta impulsiva de um ambiente* é proceder à modelagem acústica de uma determinada sala, ou seja, é inferir como as ondas sonoras, emitidas por uma fonte, irão interagir com o meio, com os objetos e as superfícies que compõem este recinto e quais as mudanças que este processo impelirá a este campo sonoro em um dado ponto do espaço em estudo. Existem várias abordagens que servem a este propósito. A seguir trataremos das técnicas mais usuais que são utilizadas para o levantamento da *RIR*.

4.2.1 Métodos baseados na decomposição do campo sonoro

Os métodos que são balizados na decomposição do campo sonoro consideram este, e conseqüentemente a *RIR*, como sendo formado por componentes elementares que

4. Auralização Inteiramente Computacional

são representados por fontes virtuais ou raios [35]. Estas técnicas são pautadas na acústica geométrica [4] cujo postulado principal implica na substituição do conceito de ondas sonoras por raios sonoros.

Quando o comprimento de onda de um sinal sonoro for muito menor que as dimensões das paredes e teto de um ambiente, e também inferior às distâncias percorridas dentro deste, pode-se assumir a representação por intermédio de raios. Estes, são definidos como uma pequena porção de uma onda esférica com abertura infinitesimal que se origina num ponto determinado e se desloca segundo uma trajetória retilínea. Essa condição para representação sonora é aceita, em ambientes acústicos, a partir de 1KHz , o que corresponde a cerca de 34cm de comprimento de onda.

De acordo com esta definição, se o meio não introduzir perdas, a energia total contida no raio se mante-rá constante ao longo do trajeto. Contudo, deve-se observar que a intensidade de um feixe de raios divergentes cai de acordo com a lei do inverso do quadrado da distância ($1/r^2$), como numa onda esférica.

Apesar da praticidade, é importante ressaltar o fato de que propagações em linhas retas desprezam o fenômeno ondulatório da *difração*. Da mesma forma, desconsideram a *interferência* pois componentes superpostos de um campo sonoro não irão interagir gerando cancelamentos e reforços baseados em suas relações de fase. Ao invés disso, suas densidades de energia ou intensidades são somadas. Essa simplificação é permitida quando componentes diferenciados são mutuamente *incoerentes* o que, em espaços acústicos, é verificado, já que os sinais possuem um vasto espectro de frequências [4].

No que tange às reflexões, ao atingir uma superfície, um raio sonoro seguirá a *lei de Snell*, ou seja, ângulo de incidência igual ao ângulo de reflexão. Essa *reflexão especular*, também é válida para feixes de raios sonoros em superfícies curvas [4]. Se a superfície refletora possuir irregularidades de ordem de grandeza comparáveis ao comprimento de onda a *lei de Snell* perde-rá a validade. Assim, a energia incidente sofrerá um *espalhamento* segundo um determinado ângulo sólido, o que se refere à *reflexões difusas*. Uma superfície pode ser perfeitamente ou parcialmente difusora. No primeiro caso o ângulo de incidência não influenciará em nada o comportamento da energia refletida e esta será espalhada de acordo com a *lei cossenoidal de Lambert* (4.11):

$$I(r) = I_0 dS \frac{\cos \vartheta \cos \vartheta_0}{\pi r^2} \quad (4.11)$$

onde supõe-se dS um elemento de área *iluminado* por um feixe de raios paralelos, ou aproximadamente paralelos, que formam um ângulo ϑ_0 com a normal da superfície e cuja intensidade seja I_0 . Dessa forma, a intensidade do raio espalhado numa direção ϑ a uma distância r pode ser determinada pela equação (4.11) [4]. Porém, na acústica de salas, as superfícies parcialmente difusoras são a regra geral para o comportamento das paredes. Apesar disto, em muitos casos uma aproximação por reflexões perfeitamente *difusas* é mais adequada do que assumir um comportamento meramente *especular* para o fenômeno da reflexão.

O uso da acústica geométrica na modelagem de ambientes tem como principal resultado a determinação dos caminhos de propagação mais significativos entre uma determinada fonte e um receptor. A seguir, serão discutidos os principais métodos que são utilizados para a solução deste problema.

Traçado de raios (RTM)

O método do *traçado de raios* [29,31], assim como todo algoritmo geométrico, tem como principal objetivo determinar os caminhos de propagação mais significativos entre a fonte e o ouvinte dentro de um ambiente. Para que isto seja feito, a fonte sonora emite certos sinais em todas as direções do espaço a ser modelado [36]. Estes, podem ser raios, partículas, feixes cônicos ou piramidais. Tais sinais, se propagam em linhas retas, já que o ar dentro da sala é suposto um meio homogêneo e isotrópico, e são refletidos segundo a *lei de Snell*; suas características físicas, como energia, tempo de chegada e direção serão armazenadas no detector/receptor para posterior confecção da *RIR*.

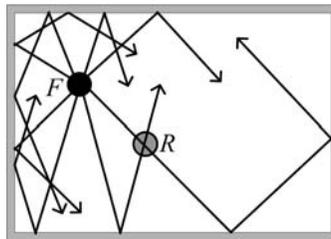


Figura 4.3: *traçado de raios*.

O *RTM* funciona bem para modelar reflexões especulares e difusas. Sempre que um determinado raio atingir uma superfície parte de sua energia será refletida segundo um ângulo especular ou será espalhada o que depende do *coeficiente de espalhamento* (δ) da superfície. No caso de ocorrer uma *reflexão difusa*, o algoritmo gera

4. Auralização Inteiramente Computacional

dois ângulos sendo um azimutal (φ) e um polar (β) que representam a direção do raio refletido. O ângulo β segue a *lei de Lambert* nas formulações mais simplificadas. Numa situação real, os *coeficientes de espalhamento* são dependentes da frequência. Porém, para efeito de modelagem, introduzir esta dependência para cada banda de frequência simulada, em cada ponto de colisão, gera um procedimento computacionalmente pesado [35]. Dessa forma, um único valor de δ é geralmente utilizado para todas as bandas.

No que se refere ao formato geométrico das paredes, o traçado de raios é utilizado, também, em superfícies curvas, fato que gera boa flexibilidade quando este método é usado conjuntamente com o *imagens da fonte* [19, 20], que só lida com paredes planas.

Se o sinal emitido pela fonte pode ser uma partícula ou um raio ideal surge uma incoerência: como atingir um ponto de recepção infinitesimal? Para solucionar este impasse o detector deve assumir dimensões finitas para registrar um número de raios adequado à representação da resposta impulsiva de uma posição no ambiente. É preferível, que este alvo possua um formato geométrico que não favoreça certas direções de incidência em detrimento de outras. Portanto, a esfera apresenta a melhor característica de isotropia espacial além de ser de fácil implementação [4, 36, 37]. Quando o detector assume um tamanho finito, surgem problemas relacionados à sua função primordial de detecção [37]. O primeiro, é o erro na quantidade de caminhos detectados e é conhecido como *deteções múltiplas*. A figura (4.4) ilustra um exemplo deste tipo.

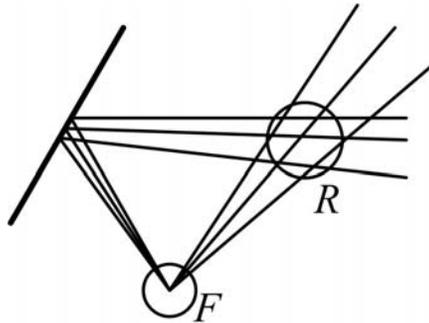


Figura 4.4: *Deteções múltiplas*

O fenômeno de detecção de mais de um caminho sonoro é natural quando se trata de ondas esféricas. Assim, para a contagem energética, essas detecções representam a lei de propagação $1/r$. Contudo, já que o *RTM* é utilizado para determinação de

caminhos de reverberação entre fonte e receptor, a melhor opção é descartar essas detecções múltiplas e implementar numericamente a *lei do inverso da distância* correspondendo ao comprimento do caminho sonoro [36]. Este procedimento é justificado, por exemplo, pela dependência da energia coletada em relação a posição relativa do receptor para com os raios (fig. 4.5).

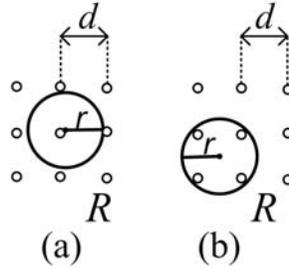


Figura 4.5: *Dois condições de detecção diferenciadas devido às dimensões do receptor e a densidade dos raios*

Claramente, percebe-se a variação por um fator 4 quando as dimensões lineares do detector são da ordem da distância entre raios adjacentes [37].

Outro erro de detecção está relacionado a caminhos sonoros inválidos. A depender do tamanho do detector algumas inconsistências podem ocorrer como por exemplo raios detectados por receptores que estão oclusos por uma superfície (fig. 4.6).

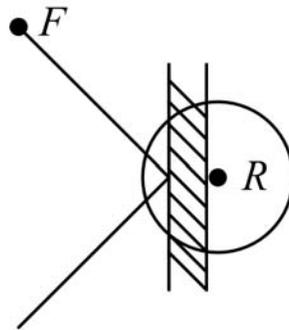


Figura 4.6: *Exemplo de detecção inválida: receptor ocluso por uma parede.*

É de extrema importância a análise da relação entre o tamanho do receptor e o número de raios que ele será capaz de detectar. Isto deve ser feito, pois é objetivo central de um processo de *traçado de raios* determinar caminhos de reverberação válidos e não detectar os inválidos. Para os detectores esféricos, o tamanho mínimo

4. Auralização Inteiramente Computacional

necessário é determinado seguindo o princípio de que nenhum caminho válido deva ser desperdiçado [36], assim, obtêm-se (eq. 4.12):

$$r_e = l_{max} \left(\frac{2\pi}{N} \right)^{1/2} \quad (4.12)$$

onde r_e é o raio da esfera, l_{max} o comprimento do raio detectado e N o número de raios. Desta fórmula (4.12), nota-se a dependência do tamanho do receptor em relação ao comprimento dos raios. Isto sugere um receptor de tamanho variável que pode ser usado para corrigir erros relacionados à densidade de raios detectada.

No *traçado de raios* existe uma relação estreita entre a sua resolução espacial, número de raios a serem emitidos, e a geometria da sala a ser modelada [36]. A probabilidade de uma determinada reflexão acontecer é função da área da parede refletora (A) e do ângulo de incidência (ϑ_0). Esta probabilidade é alta se após um tempo de viagem t a área da frente de onda por raio não for maior que $A/2$. Destas considerações [7, 36], resulta o número mínimo de raios em função de t e A e de c (velocidade do som no ar)(eq. 4.13).

$$N \geq \frac{8\pi c^2}{A} t^2 \quad (4.13)$$

Da equação (4.13) em função da área e do tempo, percebe-se que a utilização do *RTM* para calcular as secções mais tardias da resposta impulsiva da sala, a *cauda reverberante*, torna necessário um número computacionalmente grande de raios. Já o cálculo das *reflexões iniciais*, e de algumas reflexões intermediárias, é factível.

O resultado de um processo com *traçado de raios* não é a *RIR* propriamente dita, e sim uma distribuição temporal de energia [29]. Assim, é preciso converter os *ecogramas* em uma resposta ao impulso equivalente de pressão que possua a mesma composição temporal e espectral, mas com uma estrutura fina totalmente nova. Para que esta conversão seja executada é preciso subdividir a faixa de frequência que se pretende modelar em bandas. Para cada banda, é rodado um *traçado de raios* com seus respectivos coeficientes de absorção e atenuação do ar. A resolução temporal para os *ecogramas* é um ponto crítico do processo. Ela nem pode ser muito longa, já que isto provoca perdas de detalhes importantes para a *RIR*, nem muito curta, sob pena de flutuações indesejadas. Considerando a resolução temporal do sistema auditivo, é suficiente que ela esteja entre 5 e 10ms.

O uso do *RTM* não está relacionado somente à coleta de partículas que possuam dados físicos conhecidos tais quais intensidade, tempo de chegada e direção de incidência no ponto de recepção. Isto ocorre quando o *traçado de raios* é utilizado como os chamados *testes de visibilidade e oclusão* para validar fontes virtuais em *métodos híbridos* usando, também, o *método das imagens*.

Imagens da fonte ou fontes virtuais (ISM)

A reflexão de um raio sonoro, proveniente de uma fonte, numa superfície, pode ser representada por intermédio do espelhamento desta fonte sobre o plano refletor [4,5,7,35]. Isto, gera uma *imagem da fonte* ou *fonte virtual* que substitui-rá o efeito causado, no raio, pela superfície. O trajeto do raio refletido do ponto *A* até o ponto *B* é construído, após o rebatimento, ligando a imagem da fonte (A_1) até o ponto de recepção *B* e, posteriormente, fazendo a ligação entre *A* e o ponto de intersecção do segmento de reta $\overline{A_1B}$ com o plano (4.7).

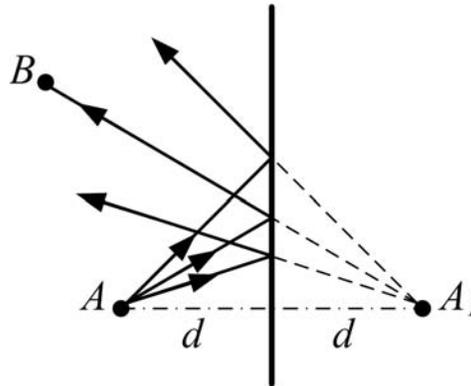


Figura 4.7: *Construção de uma imagem da fonte.*

Uma vez que a imagem foi determinada, a superfície pode ser descartada, pois, seu comportamento estará representado pela *fonte virtual*.

Se ao invés de um único plano infinito, considerarmos que a fonte está dentro de uma sala com paredes planas, podemos avaliar o comportamento de um raio sonoro construindo as imagens da fonte referentes a esta trajetória (fig. 4.8):

Onde cada reflexão adicional estará associada a uma fonte de determinada ordem. Assim A_1 se refere à primeira reflexão (de primeira ordem) e A_{12} se refere à segunda reflexão (de segunda ordem). Para um dado ambiente com N paredes planas, serão geradas N imagens de primeira ordem. Estas, por sua vez, irão produzir $(N - 1)$ imagens de segunda ordem, cada uma, ou seja, teremos um total de $N(N - 1)$

4. Auralização Inteiramente Computacional

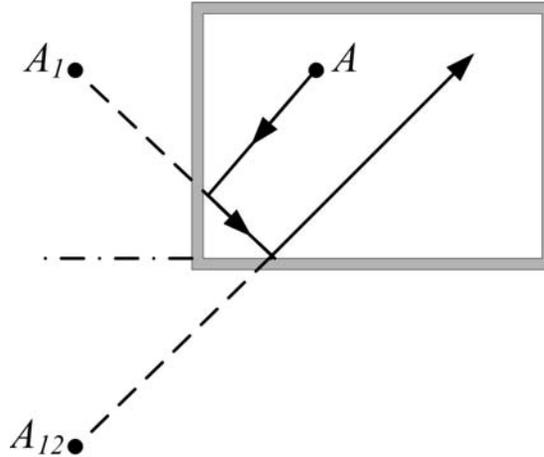


Figura 4.8: *Histórico de um raio sonoro a partir da construção de suas imagens da fonte.*

imagens. Desta forma, recursivamente, pode-se obter para uma ordem i_o a seguinte fórmula que expressa o número total (n_{imag}) de imagens obtidas (eq. 4.14):

$$n_{imag}(i_o) = N \frac{(N-1)^{i_o} - 1}{N-2} \quad (4.14)$$

A esta altura, fica claro que este procedimento gera uma constelação de *fontes virtuais* que passarão a representar as paredes de uma sala. No caso de um ambiente com formato geométrico de "caixa de sapato", haverá um padrão bem definido para esta constelação, como podemos conferir na figura (4.9).

Obviamente este padrão se mantém para a terceira dimensão da figura.

Uma vez que as imagens virtuais tenham sido encontradas para uma determinada posição da fonte, se excitarmos estas e a fonte original com a função de *Dirac* obteremos, para uma posição dada de recepção, a resposta ao impulso do par fonte/receptor, especificados. Aos sinais de cada imagem, devem ser aplicados a absoção das superfícies $(1 - \alpha)^{n_{cruz}}$, onde o expoente significa o número de cruzamentos com as superfícies refletoras, ou seja, o número de reflexões. Deve-se, ainda, multiplicar a energia de cada sinal pelo fator $1/r^2$ referente à *lei do inverso do quadrado da distância*. Desta forma, no ponto de recepção, para cada sinal individual, teremos informação sobre a intensidade e o tempo de chegada resultando assim num *ecograma* (fig. 4.1).

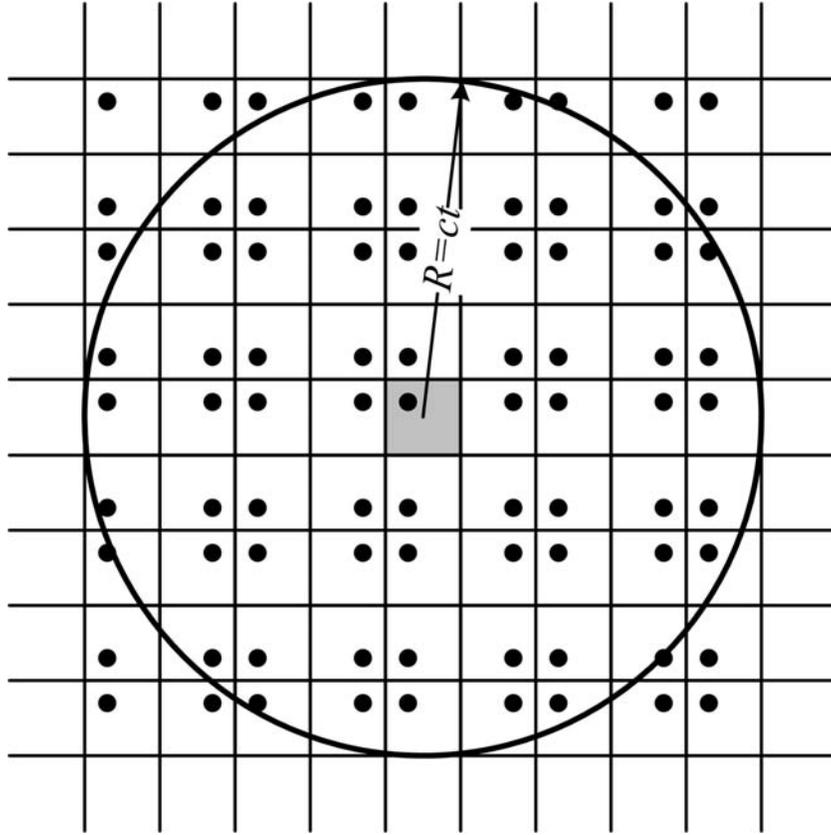


Figura 4.9: Constelação de imagens da fonte para ambiente em formato de “caixa de sapato”.

Apesar de ser um método bastante eficaz para salas retangulares, devido ao padrão da “nuvem” de imagens, o *ISM* torna-se muito custoso para computar reflexões de ordem muito alta. Da equação (4.14) pode-se perceber que o número de fontes virtuais cresce exponencialmente com a ordem das reflexões [4, 7]. Porém, é preciso observar que há duas qualidades de imagens da fonte que são geradas no modelo: as *inválidas* e as *invisíveis*. Quando se trata do último tipo, estamos falando de fontes virtuais que não são visíveis a partir de um determinado ponto de recepção. Logo, estas devem ser desconsideradas pois não contribuem para a *RIR*.

Vamos avaliar quantitativamente a relação entre o número total de imagens produzidas e o número de imagens efetivamente visíveis. Consideremos uma sala retangular para a qual foi originada uma constelação de fontes para uma determinada posição da original (fig. 4.9):

Se considerarmos que até um tempo t , todos os sinais das imagens atingem o centro da sala, pode-se imaginar uma esfera de raio $R = ct$ (fig. 4.9) que conterà

4. Auralização Inteiramente Computacional

todas as imagens. Assim do volume desta esfera ($V_{esf} = \frac{4\pi(ct)^3}{3}$), pode-se inferir que o volume de uma imagem da sala estará contido na esfera $n_{refl} = \frac{4\pi c^3}{3V} t^3$ vezes. Esta quantidade, também representa o número de fontes virtuais contidas na esfera e, conseqüentemente, o número de reflexões até um tempo t após a emissão sonora [4,7]. Estatisticamente, esta equação é válida para qualquer formato de sala.

Se compararmos o valor de n_{refl} (número de reflexões até um tempo t) com o de n_{imag} (eq. 4.14)(número de fontes até a ordem i_0) para uma sala de $15.000m^3$ com 30 superfícies, *caminho livre médio* de $16m$ (grandeza que representa o valor médio da função densidade de probabilidade que governa a ocorrência de um determinado caminho de comprimento l entre duas reflexões subsequentes de uma partícula sonora com superfícies do ambiente), teremos que, para o cálculo de uma *RIR* de $600ms$ serão necessárias fontes até a 13 ordem. Assim, teremos $n_{imag} \approx 10^{19}$ e $n_{refl} \approx 2.450$ [7]. Nota-se, então, que apesar do crescimento exponencial do número de fontes virtuais com a ordem das reflexões, apenas uma pequena parte das imagens geradas efetivamente contribuirão para uma determinada posição de recepção. Desse fato, surge a necessidade dos chamados *testes de visibilidade* [38] para garantir que só as fontes visíveis sejam levadas em consideração. Este tipo de teste, é um dos pontos críticos do *ISM* em relação ao tempo e esforço computacionais.

Dentre os *testes de visibilidade* podemos destacar o *traçado reverso* (*back tracing*) e o teste que se utiliza de um processo de *RTM*. O princípio do *traçado reverso* é executar um traçado do receptor até a fonte sonora verificando todas as superfícies envolvidas no trajeto. Para um determinado *caminho de reverberação* teremos uma cadeia de fontes até a i -ésima ordem:

$$F \rightarrow F_{n_1} \rightarrow F_{n_1 n_2} \rightarrow \dots \rightarrow F_{n_1 n_2 \dots n_{i-1}} \rightarrow F_{n_1 n_2 \dots n_i} \quad (4.15)$$

onde $n_k \neq n_{k+1}$ pode assumir valores de 1 até o número total de paredes do ambiente. A figura (4.10) ilustra um *teste de visibilidade* para uma imagem de segunda ordem. O ponto de recepção (R) está ligado à imagem (F_{12}), que está sob teste, por um segmento de reta ($\overline{RF_{12}}$). O último índice da imagem indica que a superfície 2 foi a última reflexão da trajetória. Se o ponto de intersecção entre o segmento de reta ($\overline{RF_{12}}$) e o plano que define a parede 2 está situado no polígono que define a parede real a imagem é considerada visível a partir do ponto de recepção. Neste caso este ponto de intersecção assume a função de receptor e o mesmo expediente é executado para a *fonte virtual* (F_1) e este ponto. Isto se repete até que a fonte sonora original

seja alcançada. Se caso contrário um dos pontos de intersecção não estiver situado no polígono que define a parede a fonte sob teste estará invisível.

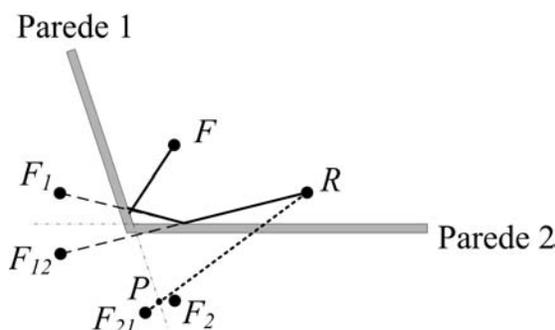


Figura 4.10: *Teste de visibilidade: traçado reverso*

O *RTM* também pode ser utilizado para testar visibilidades de *imagens*. Sempre que uma partícula atingir uma superfície o índice relacionado a esta parede será armazenado formando uma lista de superfícies atingidas para determinada partícula detectada. Cada partícula representa um caminho possível entre a fonte e o receptor e, assim, define uma imagem visível. Obviamente, mais de uma partícula pode representar a mesma imagem por isto é preciso saber se determinada *fonte virtual* já foi detectada. Desta forma, a lista de índices referente a cada detecção será comparada com as imagens encontradas.

Como citado anteriormente, algumas *imagens da fonte* geradas a partir de um ambiente devem ser avaliadas quanto a sua validade física. Este teste independe da posição do receptor podendo ser feito a priori deste. Ainda em relação à visibilidade, muitas fontes podem não ser visíveis de ponto algum dentro do ambiente e por isso são inválidas não contribuindo assim para o processo de conservação de energia. Algumas, porém, são irreais do ponto de vista físico e devem ser descartadas em prol da redução do tempo de processamento [39]. Elas podem ser reflexões que acontecem diretamente entre superfícies que não estão visíveis uma para a outra (fig. 4.11).

Avaliando a construção de uma constelação de fontes virtuais por intermédio de uma estrutura em árvore [39] (fig. 4.12), na qual os nós sejam as imagens geradas a partir da equação n_{imag} (4.14), ou seja, sem conhecimento de formas geométricas, poderemos perceber que quando existir um nó referente a uma *imagem inválida*, este, e todas as fontes descendentes, poderão ser eliminados a priori do conhecimento da posição de recepção (fig. 4.12). Para uma ordem de reflexão (i_0) o número de *imagens da fonte* será de $N(N - 1)^{i_0 - 1}$, onde N é o número de paredes.

4. Auralização Inteiramente Computacional

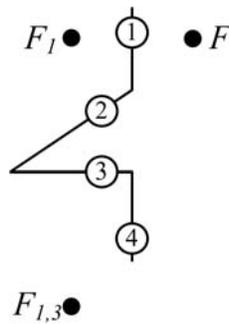


Figura 4.11: *Reflexão irreal: não há como ocorrer uma reflexão entre as paredes 1 e 3. A imagem $F_{1,3}$ é irreal.*

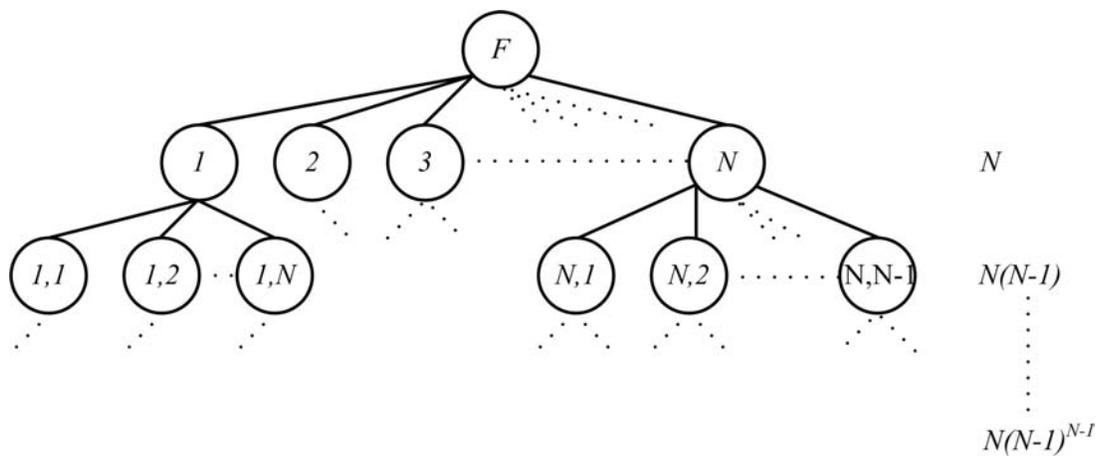


Figura 4.12: *Estrutura em árvore de imagens da fonte possíveis para uma sala com N paredes.*

Do ponto de vista computacional, as eliminações de ramos da árvore viabilizam a utilização do método. Porém, as fontes invisíveis não permitem que seus descendentes sejam eliminados pois estes, sim, podem estar visíveis em reflexões de ordem mais elevada.

O método das imagens da fonte é bastante robusto para o cálculo exato da RIR, mas possui limitações que devem ser compreendidas e avaliadas de acordo com a aplicação que se pretende implementar. Uma das mais óbvias é o fato desta técnica se limitar a superfícies planas [4,5,7,35]. Em alguns casos, a depender da geometria do ambiente que se deseja modelar, pode ser feita uma aproximação pela sala retangular ou aproximações de superfícies curvas por planos, mas isto, para fins de auralização, compromete a fidedignidade do ambiente virtual [35].

Outra desvantagem relacionada à geometria das salas é o fato dele não lidar corretamente com o tamanho finito das paredes que compõem o ambiente, como visto anteriormente. Para esquinas *rígidas* ou formadas por superfícies com baixo coeficiente de absorção, isto pode ser amenizado pela introdução de fontes secundárias denominadas de *edge sources* [35, 40]. Esta é uma forma de introduzir o fenômeno ondulatório da difração, que originalmente não é levado em conta pelo *ISM* [15], para fins de auralização. O princípio desta idéia é espelhar a quina (linha de encontro de duas paredes) nas superfícies do ambiente, substituir as imagens geradas por fontes pontuais, e fazer com que esta fonte gere uma resposta que siga o modelo matemático proposto em [40] que é baseado em respostas impulsivas de quinas.

A posição de cada imagem não se altera com o deslocamento do receptor, por isso não há necessidade de proceder à modelagem novamente. Uma vez estabelecido um conjunto válido de *imagens*, a flexibilidade para simulações é enorme já que mudanças de diretividade da fonte, do receptor, e de impedâncias das paredes não alteram nem as posições das *fontes virtuais* nem o *teste de visibilidade*. Estes detalhes são vantajosos para fins de processos interativos, porém, deve-se observar que mudanças de posição do ouvinte implicam em novos testes de visibilidade.

A abordagem por *imagens* não opera com superfícies difusoras o que acaba sendo uma limitação já que em paredes reais há uma mistura entre energia sendo refletida especularmente e sendo espalhada segundo um ângulo sólido. Para incluir este fenômeno métodos pseudo-estatísticos [31] e até mesmo o *RTM* são utilizados [19].

O *ISM* é um método determinístico capaz de encontrar todos os caminhos de reverberação da fonte até o ouvinte. Para fins de auralização o seu uso se torna vantajoso quando ele é utilizado para o cálculo das *reflexões iniciais*, pois, assim, serão necessárias imagens de ordens baixas. Baseado no fato de que ele não lida com o fenômeno ondulatório do *espalhamento* e, sabendo da importância perceptual dessa energia espalhada para a formação de um evento auditivo que corresponda à situação real, se faz necessário o uso das *imagens da fonte* conjuntamente com outros métodos não só para incluir as *reflexões difusas*, mas, também, para proceder ao cálculo da *cauda reverberante*.

Traçado de pirâmides

O método do *traçado de pirâmides* [6, 41] é essencialmente uma derivação do *traçado de raios*. Aqui, ao invés de raios, feixes piramidais são traçados recursivamente através do ambiente (fig. 4.13) [6, 23]. Ao interceptar um polígono no interior

4. Auralização Inteiramente Computacional

do espaço, a zona de sobreamento é retirada e será contruído um feixe de transmissão associado a esta. A reflexão, será o espelhamento deste feixe transmitido sobre o polígono (superfície a) alvo. A esta reflexão estará associada uma fonte F_a .

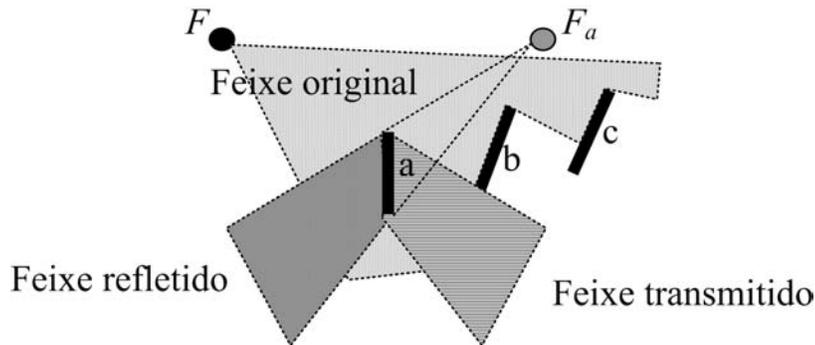


Figura 4.13: Traçado de pirâmides: fonte sonora F espelhada na superfície a resultando na imagem F_a

A justificativa para o uso de tal formato para o feixe, é o fato de que este cobre, exatamente, todo o espaço de direções a partir da fonte sonora. A escolha deste tipo de feixe possui a vantagem de representar infinitas intersecções entre raios e uma superfície, evitando assim os problemas de amostragem espacial inerentes ao *RTM*. Consequentemente, todos os caminhos potenciais de propagação podem ser enumerados.

No *RTM* a fonte sonora emite raios centralizados nas áreas dos triângulos esféricos que são obtidos a partir da subdivisão das faces de um icosaedro em triângulos, e posterior projeção destes na esfera que circunscreve o icosaedro. No traçado piramidal estes triângulos são as próprias faces das pirâmides. Assim, a exigência de um tamanho finito para o receptor perde o sentido e este pode assumir dimensões infinitesimais, ou seja, receptor pontual. Para que um feixe seja detectado no ponto de recepção, basta que o detector pontual esteja entre os planos que definem a pirâmide.

Apesar de não sofrer com problemas relacionados à amostragem espacial, o *traçado de pirâmides* também possui suas desvantagens. A primeira delas é que o número finito de feixes acaba limitando a determinação do número correto de *imagens da fonte*. Outro problema é que as pirâmides não se dividem no momento de uma reflexão quando transpassam duas ou mais paredes. Na verdade as reflexões de um feixe piramidal são regidas pelo raio central deste e as superfícies que este raio atinge [41] (fig. 4.14).

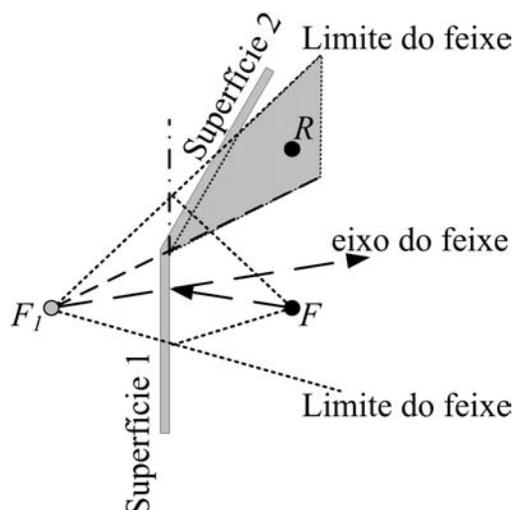


Figura 4.14: Receptor R irradiado incorretamente pela imagem F_1 . A área em cinza representa a região do espaço que está sendo irradiada incorretamente.

Das situações geradas por esta indivisibilidade do feixe no momento da reflexão podemos citar as imagens que irradiam receptores que não estão visíveis (fig. 4.14) e as imagens que não iluminam receptores visíveis.

O *traçado de pirâmides* possui uma vantagem sobre o *ISM* em relação ao número de fontes virtuais a ser considerado. Como cada região do espaço coberta por um feixe representa todos os pontos para os quais uma imagem da fonte está visível, fontes virtuais de ordens mais elevadas só serão consideradas para os polígonos que são interceptados pelo feixe.

Este método, observando as particularidades discutidas anteriormente, é eficaz no cômputo das reflexões especulares e por isso é utilizado para o cálculo das *reflexões iniciais*. Em relação à modelagem da difusão, ele não possibilita soluções baseadas puramente em seus algoritmos [41].

Transição de energia ou radiosidade acústica (AR)

O método da *transição de energia*, apesar de ser baseado na acústica geométrica, é bem diferente dos anteriores. O *AR* discretiza as superfícies do espaço modelado e simula a troca de energia entre pares da cadeia de transmissão sonora, ou seja, da fonte para elementos das paredes, de elementos para elementos, e destes para o receptor [35, 42] (fig. 4.15).

O princípio fundamental da *transição de energia* é que todas as superfícies refletem energia espalhada. Baseado neste fundamento, o seu uso é aconselhado para

4. Auralização Inteiramente Computacional

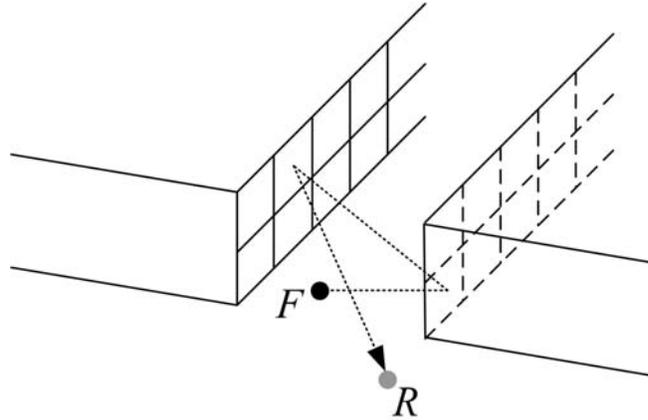


Figura 4.15: RA de uma esquina.

a determinação de *reflexões difusas* e da *cauda reverberante*, já que na partição final da resposta impulsiva de uma sala a energia predominante é resultado de reflexões que causam *espalhamento* [4, 42]. Assim, se assumirmos que as paredes refletoras seguem a *lei de Lambert* (eq. 4.16):

$$I(\theta, R) = I(0, R) \cos \theta \quad (4.16)$$

onde $I(\theta, R)$ é a intensidade sonora a uma distância R espalhada em uma direção θ ($0 \leq \theta \leq \pi/2$) em relação à normal da superfície e $I(0, R)$ é a intensidade da onda incidente a uma distância R que independe do ângulo de incidência, estaremos assumindo que cada elemento nos quais a parede está discretizada obedece a esta lei. Isto simplifica bastante as fórmulas que descrevem as trocas energéticas no modelo.

As equações analíticas que compõem o AR partem da definição de uma *densidade de radiação*, $B(r, t)$, onde r é a posição de um elemento infinitesimal de área (fig. 4.16). Essa densidade representa a taxa na qual a energia deixa uma superfície de área unitária [42]. De acordo com [42], obtém-se (eq. 4.17):

$$B(r, t) = \frac{\rho(r)}{\pi} \int_S B(r', t - R/c) e^{(-mR)} \frac{\cos \theta \cos \theta'}{R^2} dS' + B_d(r, t) \quad (4.17)$$

Essa equação avalia a contribuição de cada elemento dS' para a *densidade de radiação* de um elemento dS mais a parcela de energia direta $B_d(r, t)$ proveniente da fonte sonora. A integração é executada sobre toda a superfície do recinto (S). O termo $e^{(-mR)}$ se refere a absorção do ar, R é a distância entre o elemento dS e o

dS' , e os ângulos θ e θ' são aqueles formados entre a distância R e as normais dos elementos dS e dS' respectivamente (fig. 4.16).

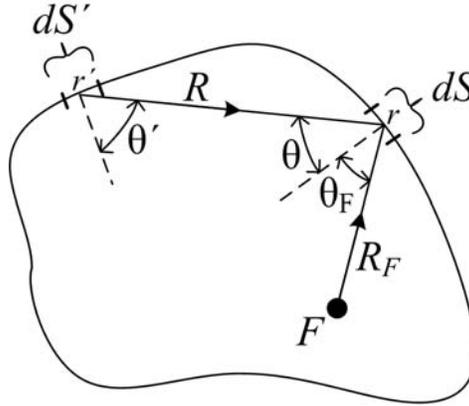


Figura 4.16: Geometria da *densidade de radiação* de um elemento de área dS com posição definida pelo vetor r .

Uma vez que as *densidades de radiação* de cada elemento de área foram avaliadas por intermédio da equação anterior (eq. 4.17) a intensidade no receptor é calculada por (eq. 4.18),

$$I(r_r, t) = \frac{1}{\pi} \int_S \frac{B(r, t - R_r/c) \cos \theta_r}{R_r^2} e^{(-mR_r)} dS + I_d(r_r, t) \quad (4.18)$$

onde R_r é a distância entre elemento e receptor (fig. 4.17). A função sob integração, para toda a superfície do recinto, resulta na soma das intensidades de cada elemento discretizador no ponto de recepção (r_r) levando em consideração as respectivas atenuações do meio. A parcela $I_d(r_r, t)$ inclui a contribuição direta da fonte, facilmente calculada com o conhecimento da potência acústica da fonte sonora [3, 4, 42].

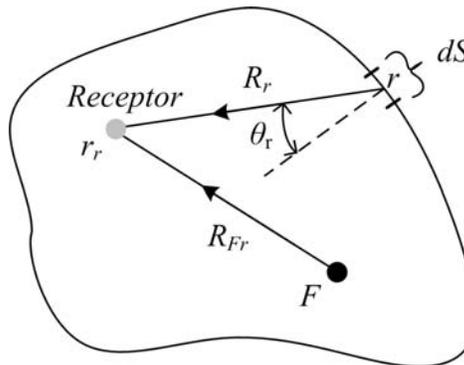


Figura 4.17: Geometria da transferência de energia para o receptor.

4. Auralização Inteiramente Computacional

Das equações (4.17 e 4.18), é perceptível uma das maiores vantagens do *AR*. As *densidades de radiação* de cada elemento que discretiza as superfícies da sala, $B(r, t)$, independem da posição do receptor. Na verdade $B(r, t)$ depende apenas da composição geométrica do ambiente, da discretização espacial e dos parâmetros da fonte, tais como potência e posição. Desta forma, para posições novas de recepção, não há necessidade de proceder a todo o cálculo novamente.

Em geral, estas equações (4.17 e 4.18) analíticas devem ser resolvidas numericamente. Nestas soluções numéricas o interior da sala é discretizado em elementos planos finitos S_i . Desta forma, teremos como soluções numéricas para as equações de densidade e intensidade:

$$B_i(t) = \rho_i \sum_{j=1}^N B_j(t - R_{ij}/c) e^{(-mR_{ij})} F_{ij} + B_{di}(t) \quad (4.19)$$

$$I(r_r, t) = \frac{1}{\pi} \sum_{i=1}^N \int_{S_i} \frac{B_i(t - R_{ri}/c) \cos \theta_r}{R_r^2} e^{(-mR_{ri})} dS + I_d(r_r, t) \quad (4.20)$$

onde N é o número de elementos discretizadores, ρ_i é o coeficiente de reflexão do elemento i (constante dentro do elemento) e R_{ij} distância entre pontos centrais dos elementos S_i e S_j . Os F_{ij} são conhecidos como *fator de forma* entre o elemento i e o j . Estes podem ser entendidos fisicamente como a fração de energia que deixa o elemento i e que incide no elemento j [35, 42].

O método da *transição de energia* permite uma grande flexibilidade em relação à precisão e a velocidade que se deseja atingir por intermédio da manipulação do número de elementos que discretizam as superfícies do ambiente. Para além disso, o tamanho desses elementos não precisa ser menor que o comprimento de onda do sinal sonoro em questão [35], fato esse, computacionalmente relevante. Porém, o *AR* é limitado no uso com reflexões especulares e não lida com difrações.

4.2.2 Métodos baseados na equação de onda

Ao contrário dos *métodos geométricos* que não lidam, em sua formulação básica, com fenômenos ondulatórios como a difração, aqueles *baseados em soluções numéricas da equação de onda* produzem os resultados mais precisos para as modelagens dos ambientes acústicos [5, 43].

O princípio básico destes métodos é a discretização do volume ou das superfícies de um espaço e a solução da equação de onda para cada um destes elementos.

Obviamente esta subdivisão espacial deve ser de tal forma que os comprimentos de onda ainda possam ser avaliados com precisão dentro de um elemento discretizador. Assim, modelos baseados nestas técnicas possuem um limite superior de operação em relação à frequência.

Apesar de lidarem com ambientes de formatos geométricos arbitrários, estes métodos possuem, também, limitações em relação às dimensões do espaço. Para estes métodos as *condições de fronteira* devem ser bem especificadas, ou seja, é necessário que haja um conhecimento a respeito das impedâncias complexas das superfícies em questão [25] o que na prática ainda não está bem determinado para ambientes reais.

Elementos finitos (FEM)

O *método de elementos finitos* é baseado na solução da *equação de Helmholtz* para elementos que discretizam todo o volume de um ambiente [25] (fig.4.18). Assim, a pressão nos pontos de geração e recepção são calculados e a função de transferência entre estas posições é obtida.

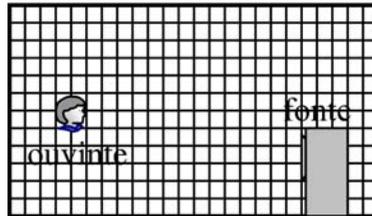


Figura 4.18: *FEM (discretização do volume do ambiente)*

Em situações gerais de audição em um recinto o meio é considerado homogêneo, com isto os cálculos são simplificados. O número de elementos aumenta em função da frequência e do tamanho do ambiente o que resulta num esforço computacional maior devido ao número de cálculos a serem executados.

A grande vantagem deste método é que para locais que necessitem de um detalhamento maior devido a produção de fenômenos acústicos complexos, tal como uma esquina, a estrutura dos elementos discretizadores pode ser refinada permitindo a modelagem das interações das ondas com estes pontos críticos.

Como o *FEM* soluciona numericamente a equação de onda, os fenômenos ondulatórios como difração, espalhamento e refração são incluídos em sistemas que se utilizam desta técnica de modelagem. O resultado da solução, via *elementos finitos*, é a função de transferência no domínio da frequência. Para utilizar a *convolução* é necessário, então, aplicar a *transformada inversa de Fourier* nesta função

4. Auralização Inteiramente Computacional

de transmissão e obter a *reposta impulsiva*. Para além disto, no domínio do tempo as informações a respeito do comportamento do ambiente são mais acessíveis.

O *FEM*, para a *simulação binaural* com uso de *HRTF*, não é apropriado devido ao fato de não lidar com a informação das direções de incidência necessárias para o uso das *HRTF*. Assim, abordagens simplificadas de geração dos *sinais binaurais* devem ser implementadas como por exemplo o cálculo de pressões em pontos que possuam relações interaurais (distância entre orelhas) ou pressões sobre uma esfera [43].

No que se refere à faixa de frequências e às dimensões dos recintos, o *FEM* possui boa aplicabilidade em relação às baixas frequências e à ambientes pequenos.

Elementos de contorno (BEM)

O *método de elementos de contorno* também soluciona a equação de onda, porém, o faz somente para os elementos que discretizam as superfícies do recinto (fig. 4.19).

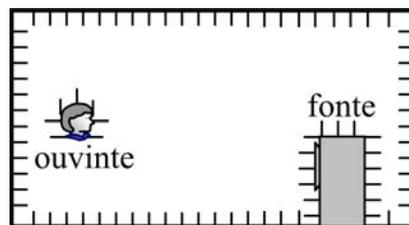


Figura 4.19: *BEM (discretização das superfícies do ambiente)*

Para esta abordagem, a equação de onda é expressa como equação de *Green* [35], sua formulação integral, e é solucionada por intermédio da subdivisão das superfícies do ambiente considerando que a pressão seja combinação linear de um número finito de funções destes elementos. Desta forma, o campo sonoro, no interior de um ambiente, pode ser construído a partir das contribuições da pressão sonora e da *velocidade de partícula normal* do contorno das superfícies do espaço modelado.

A grande vantagem do *BEM* em relação ao *FEM* é que o número de elementos utilizados para discretizar o ambiente é muito menor. Porém, as matrizes resultantes no *BEM* não são tão esparsas quanto às do *FEM*. Em geral elas são menores porém mais densas e não possuem simetrias que costumam facilitar as soluções.

Assim como no *FEM*, o *BEM* torna-se crítico para ambientes muito grandes e para frequências elevadas. Tipicamente, os elementos discretizadores devem ter $(1/8)\lambda$. Desta forma, este método é mais apropriado para a modelagem de ambientes pequenos e para frequências baixas.

Diferenças finitas no domínio do tempo (FDTD)

O método de diferenças finitas no domínio do tempo aproxima as derivadas com respeito ao tempo e ao espaço da equação de onda por diferenças finitas. Assim como no *FEM*, todo o volume do ambiente é discretizado para que sejam efetuados os cálculos das equações [35, 44].

Este método possui formulações capazes de lidar com formatos de *mesh* (estrutura de elementos discretizadores) cartesianos ou não-cartesianos, cuja utilização depende das superfícies do ambiente. No caso de um formato cúbico para elementos do *mesh*, a solução é muito mais simples do que para formas mais complexas [44].

Dentre as variações de *FDTD* mais utilizadas atualmente figura o *DWM* (*digital waveguide mesh*) [25, 30]. Neste método, os elementos nos quais o ambiente foi discretizado são junções que estão interconectadas por *guias de onda*. O conceito de *guia de onda* se refere a qualquer meio no qual a propagação da onda é unidimensional. Considerando que não haja perdas, as soluções podem ser expressas como ondas que chegam e partem num *guia* bidirecional (fig. 4.20).

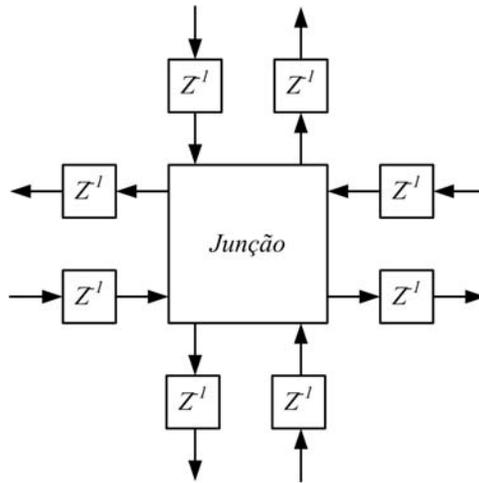


Figura 4.20: *DWM* (discretização do volume do ambiente em representação bidimensional)

Estes *guias* são modelados em ambiente digital como *linhas de atraso*. Para o caso bidimensional com *junções* quadradas (fig. 4.20) existem $N = 4$ guias interconectando os elementos amostrados. Se considerarmos que em uma linha de atraso de um guia de índice i flui um sinal p_i^+ e na outra linha uma pressão p_i^- , a pressão sonora num determinado ponto será:

4. Auralização Inteiramente Computacional

$$p_i = p_i^+ + p_i^- \quad (4.21)$$

Se as *junções* não introduzem perdas as condições (eq. 4.22) e (eq. 4.23) [3] são válidas para estes elementos discretizadores:

$$\sum_{i=1}^N v_i^+ = \sum_{i=1}^N v_i^- \quad (4.22)$$

$$p_1 = p_2 = \dots = p_i = \dots = p_N \quad (4.23)$$

Destas condições (eq. 4.22 e eq. 4.23) a pressão numa *junção* (p_J) pode ser calculada por (4.24):

$$p_J = [2 \sum_{i=1}^N (p_i^+ / Z_i)] / [\sum_{i=1}^N (1 / Z_i)] \quad (4.24)$$

Avaliando o resultado da modelagem representado pela equação (4.24), um sinal sonoro introduzido num guia, se propagará até atingir uma *junção* e deste ponto irá se espalhar de acordo com as impedâncias de cada guia que estão atrelados às superfícies dos quadrados, para o exemplo bi-dimensional da figura (4.20). No caso de ambientes acústicos, em geral, o meio é homogêneo, assim todas as impedâncias dos guias do *grid* serão iguais. Porém, as superfícies que limitam o ambiente devem assumir valores de impedância diferenciados para que as devidas reflexões ocorram.

4.2.3 Métodos estatísticos

Existe uma relação estreita entre as teorias sobre ambientes acústicos e as teorias estatísticas. A definição do *tempo de reverberação* é toda baseada em considerações de comportamento aleatório das ondas sonoras após um determinado tempo de emissão [4]. Os *métodos estatísticos* geralmente são utilizados para o cálculo da *cauda reverberante* da *RIR* pois, como visto anteriormente, esta partição da evolução temporal da energia de um ambiente possui um comportamento mais aleatório e consequentemente energia distribuída de maneira difusa [4].

Estas técnicas costumam ser utilizadas em conjunto com as anteriores, principalmente com o *ISM* [16]. Neste caso, dois sinais de *ruído branco gaussiano*, independentes, são gerados e manipulados com base em parâmetros obtidos das reflexões

de ordem mais baixa para serem posteriormente apresentados às orelhas. Essas manipulações são balisadas em três expedientes: imposição de correlacionamento dos sinais ruidosos baseado no *IACC*, filtragem baseada nas características absorptivas das superfícies e janelamento temporal.

Utilizando, ainda, o fato de não ser necessário o conhecimento a respeito da estrutura fina da *cauda reverberante*, esta parte final na *RIR* pode ser sintetizada por intermédio de um processo *Poissoniano* de filtragem [31]. As características deste filtro, que devem estar relacionadas ao decaimento energético, são obtidas por um *RTM* de baixa resolução temporal (períodos de coleta de energia ampliados) para algumas bandas de frequência. São consideradas, neste processo, energia especular e espalhada. Desta forma, para cada intervalo de coleta energética, é obtido um espectro de potência que deverá ser suavizado e interpolado para que se obtenha uma função contínua em cada intervalo. Em seguida, tira-se a raiz quadrada para transformar as respostas energéticas em respostas de pressão resultando, assim, na magnitude do filtro. A fase é atribuída segundo uma distribuição de *Poisson* utilizada para selecionar pontos no eixo do tempo [29]. Este expediente de atribuição aleatória de fase é justificado pelo fato das reflexões na *cauda* estarem superpostas aleatoriamente.

4.3 Comparação entre os métodos e hibridização

Cada método de modelagem acústica possui características próprias em relação a faixa de frequência de operação, formato geométrico das superfícies do ambiente, dimensões do espaço e capacidade de modelagem de fenômenos ondulatórios. Para além disso, as aplicações em *simulações binaurais* tornam necessários o conhecimento dos requerimentos computacionais do sistema tal como precisão, velocidade e interatividade para a escolha da técnica adequada.

Como visto anteriormente os métodos baseados na acústica geométrica assumem que o comprimento de onda é muito inferior às dimensões dos obstáculos dentro do ambiente e das superfícies que o delimitam. Em contraste, os métodos baseados em soluções da equação de onda, na prática, se limitam à comprimentos de onda maiores devido a necessidade de elementos discretizadores entre $(1/6)\lambda$ e $(1/10)\lambda$ para uma modelagem correta [35]. Como estamos interessados em sistemas que proporcionem *auralização* na faixa de frequências referentes à voz e à música ($20Hz - 20KHz$) a hibridização é um recurso que pode ser utilizado para ampliar a *banda* dos sinais modelados.

4. Auralização Inteiramente Computacional

Com relação às dimensões, o confronto dos *métodos geométricos* e de *soluções da equação de onda* sugere uma relação de compromisso entre frequência do som e volume do recinto. Para a faixa de comprimentos de onda desejada os métodos geométricos devem ser utilizados com cuidado já que ambientes muito pequenos podem invalidar a pressuposição de propagação segundo raios ao invés de ondas. Já o *FEM*, o *BEM* e o *FDTD* são praticamente inviáveis computacionalmente para ambientes de volumes grandes.

Os fenômenos ondulatórios como a *difração* e o *espalhamento* se tornam indispensáveis para ambientes cujas dimensões dos obstáculos em seus interiores são comparáveis a λ . Numa sala de audição musical, por exemplo, a platéia origina difrações da informação sonora para um determinado ouvinte. É interessante que este e outros fenômenos como o espalhamento devido a rugosidade de superfícies estejam representados na modelagem para que a *simulação binaural* represente corretamente o campo sonoro real. Os métodos baseados em soluções da equação de onda, inerentemente, já lidam com os fenômenos ondulatórios. Em suas formulações originais, os métodos geométricos não modelam a *difração*. O *ISM* possui uma representação baseada em fontes secundárias denominada de *edge source* [40] que modela a difração baseado no conceito de resposta ao impulso de uma quina (linha que delimita o encontro de duas superfícies). No que se refere ao *espalhamento*, o *RTM* modela este fenômeno, porém, o *ISM*, não.

A modelagem de superfícies curvas, côncavas ou convexas, é executada pelos *métodos baseados na equação de onda* e pelos *geométricos*, com exceção do *ISM*. Para utilizá-lo nestes casos seria necessário aproximar tais superfícies por várias superfícies planas menores, mas tal expediente é muito custoso computacionalmente.

Com essas considerações, torna-se evidente a necessidade de concepção de modelagens calcadas em mais de uma técnica para simular largas faixas de frequência e ambientes com dimensões e formatos variados. A hibridização também é necessária para a obtenção de tempos computacionais adequados a cada situação. Os *métodos baseados em soluções numéricas da equação de onda* demandam esforço e tempo computacionais que impedem o seu uso para processamentos em tempo real e conseqüentemente para o uso em sistemas interativos nos quais uma mudança de posição do ouvinte ou da fonte acarretam novos cálculos. Porém, pré-processamentos utilizando estas técnicas podem ser vantajosos para predeterminar a *reverberação* que é uma característica geral do ambiente e independe do ponto de audição.

Na maioria dos sistemas de *auralização* atuais [16, 18, 19] o *método das fontes virtuais* e o de *traçado de raios* são utilizados para computar as *reflexões iniciais*. Para a *cauda*, um método estatístico é geralmente executado. O *DIVA* [8, 25], utiliza o *ISM* para computar as *reflexões iniciais* e uma hibridização composta pelo *DWM*, frequências baixas, e pelo *RTM*, frequências altas, para sintetizar a *cauda reverberante*.

4.4 Influência da Absorção, Espalhamento e Difração

Os fenômenos sob os quais um sinal sonoro está sujeito, em um ambiente, são complexos e dependem das características reais das superfícies do espaço. Paredes difusoras, de dimensões limitadas, curvas e até mesmo estruturas ressonadoras são encontradas frequentemente em ambientes acústicos.

Prever o comportamento de um sinal após sua colisão com uma superfície, para os *métodos geométricos*, é modelar um sistema linear. A função de transferência que representa este sistema, razão entre os espectros do sinal acústico refletido e do sinal incidente, nada mais é do que o *coeficiente de reflexão* da superfície (\mathbf{R}). Como vimos anteriormente, ele é um valor complexo que depende da frequência e do ângulo de incidência. Na prática, porém, este coeficiente não é utilizado devido a escassez e as dificuldades inerentes às suas medições.

Para relacionar os sinais de entrada e saída deste sistema, utiliza-se o *coeficiente de absorção* (α) que costuma ter um valor médio e a sua dependência em relação à frequência é mensurada em *faixa de oitava*. Esta dependência de α por faixas, induziu duas abordagens de simulações [36], uma de *faixa larga* e outra de *faixa estreita*. Esta última é executada em *faixas de oitava* e cada α é utilizado para sua *faixa* correspondente. Já para a simulação em *faixa larga* a resposta impulsiva da superfície é convoluída com os sinais que a atingem, provenientes de raios ou *imagens*. Esta resposta ao impulso é obtida a partir do *coeficiente de absorção* que permite calcular a magnitude do *coeficiente de reflexão* para algumas frequências ($\alpha = 1 - |\mathbf{R}|^2$). Em seguida promove-se uma interpolação linear para a obtenção da magnitude de R para toda a faixa de frequências. Considerando que este sistema é de *mínima fase* [11] obtêm-se a fase da função de transferência e com a transformação inversa de Fourier, chega-se a resposta impulsiva.

É importante observar que ambientes cujos *coeficientes de absorção* das superfícies não estejam distribuídos uniformemente enfraquecem a aplicação do conceito

4. Auralização Inteiramente Computacional

clássico de RT_{60} . As fórmulas, que são função do volume e da absorção do ambiente, assumem que o campo sonoro é difuso, no decaimento energético [4]. Desta forma, no que se refere às aproximações que são utilizadas para o cálculo da *cauda reverberante*, baseadas em requerimentos perceptuais, é preciso ter cuidado para que decaimentos perfeitamente exponenciais não sejam aplicados à *cauda* em ambientes cuja distribuição de energia, nesta etapa da *RIR*, não seja uniforme. Para contornar esta situação é que são utilizados, por exemplo, métodos que avaliam as características espectrais e temporais gerais da evolução temporal da energia tardia no ambiente, como o *RTM* de baixa resolução citado anteriormente.

O *espalhamento* causado por *superfícies parcialmente difusoras*, também depende da frequência e do ângulo de incidência [4]. Este fenômeno ainda é tratado de forma bastante aproximada devido a escassez de informações a respeito de superfícies difusoras. Recentemente, em abril de 2004, um padrão para execução de medidas de *coeficientes de espalhamento* de superfícies foi publicado na *ISO 17497-1*. O *coeficiente de espalhamento* (δ) é definido como sendo a razão entre a energia difusa refletida e a energia total refletida (difusa e especular).

Existem abordagens diferenciadas, para simular *espalhamento*, que variam desde os métodos de modelagem acústica utilizados, como por exemplo o *RTM* e o *AR*, até se o δ será dependente para faixas de frequência ou não. Todas estas abordagens assumem que a energia da *RIR* que deriva do *espalhamento*, devido a superfícies rugosas, não precisa ser determinada com a mesma precisão que é necessária para as *reflexões especulares*. Desta forma, é suficiente que o sinal sintetizado para representar a energia difusa tenha um comportamento espectral e temporal geral.

Heinz [31] executa um *RTM* de baixa resolução temporal que só estima a energia *difusa*. Assim, quando uma partícula atinge uma superfície parte da sua energia será reduzida por um fator $(1 - \alpha)$, sendo α o *coeficiente de absorção* para uma determinada faixa de frequência. Uma comparação entre um número aleatório e o *coeficiente de espalhamento* da superfície decide se a reflexão é *especular* ou seguirá a *lei de Lambert* (4.11). Obviamente, se δ variar para cada faixa de frequência o *RTM* deverá ser executado para todas estas faixas o que onera ainda mais a simulação em relação a esforço e tempo computacionais. O receptor só irá detectar uma partícula se a sua última reflexão for não-especular. Com este procedimento é construído o *ecograma* que representa a evolução temporal da energia difusa.

O *espalhamento* é perceptualmente substantivo para a *auralização*. Sistemas de simulação que não levam em conta esta qualidade de energia acústica não soam naturais. Por isto, é interessante que os algoritmos possam levar em conta este fenômeno. Porém, implementações objetivando a variação deste parâmetro com respeito à frequência podem comprometer sistemas que se destinam a interatividade em tempo real.

Outro importante fenômeno ondulatório para a *auralização*, principalmente, de salas de concerto e outros ambientes que possuam platéia ou obstáculos, é a *difração*. Imaginemos a seguinte situação: uma orquestra executando uma peça musical em uma sala de concertos lotada de ouvintes. Certamente, numa posição central da platéia, o sinal será composto de inúmeras *difrações*, principalmente as que ocorrem devido à presença dos músicos e aquelas originadas nos ouvintes próximos desta posição. Portanto, é necessário que a *difração* seja modelada nestas situações para que haja *simulações binaurais* que correspondam à realidade. Porém, simular *difrações* e até mesmo *espalhamentos* que ocorrem nas proximidades da fonte e do receptor, por intermédio de *métodos geométricos*, é extremamente complicado.

Recentemente, os pesquisadores passaram a estudar com mais profundidade quão audíveis são os efeitos da *difração* e do *espalhamento* em *auralizações* de ambientes acústicos [45]. Essas pesquisas pretendem elucidar quais os parâmetros mais importantes para estes dois fenômenos, como por exemplo, qual seria a faixa de frequências do sinal na qual a dependência do *coeficiente de espalhamento* seria perceptualmente mais audível.

O que se observa [17, 35, 45] é que estes fenômenos ondulatórios precisam ser computados de maneira precisa para a partição da *RIR* onde ocorrem as *reflexões iniciais*. É nesta etapa da resposta do ambiente que suas contribuições são perceptualmente mais importantes.

4.5 Cálculo da *BRIR*

Uma vez que a *resposta impulsiva do ambiente* tenha sido determinada torna-se necessário o cálculo da *resposta impulsiva binaural do ambiente* para que a *auralização*, ou seja, convolução da *BRIR* com o sinal anecóico e audição do resultado, seja executada. Para este cálculo é necessário que o ouvinte seja representado por suas funções de transferência (*HRTF*), ou alguma simplificação de modelamento que permita a obtenção de dois sinais que possuam, ao menos, *diferenças interaurais*.

4. Auralização Inteiramente Computacional

Para cada *caminho de reverberação*, identificado por um índice i , estarão associados os parâmetros

$$r_i, \varphi_i, \beta_i, h_{ci}(t)$$

onde r_i é a distância percorrida pela onda sonora, φ_i e β_i os ângulos de incidência em relação ao ouvinte e $h_{ci}(t)$ é a resposta impulsiva de um determinado *caminho de reverberação* que incorpora atenuações referentes à distância e ao ar, filtragens das superfícies e diretividade da fonte sonora.

O índice i possui um conjunto de valores finitos e assume valor igual a zero, para o caminho direto da fonte ao receptor, e um valor máximo igual a N . Vale salientar que este número máximo de *caminhos de reverberação*, para aplicações em que é necessária a redução do esforço computacional, é limitado pelo número de *reflexões iniciais* às quais os direcionamentos possuem extrema relevância para a formação do *evento auditivo*.

Utilizando o par de *respostas impulsivas binaural (HRIR)*,

$$h_{e\varphi_i\beta_i}(t), \text{ orelha esquerda}$$

$$h_{d\varphi_i\beta_i}(t), \text{ orelha direita}$$

convoluindo a resposta impulsiva de cada caminho, $h_{ci}(t)$, com seu respectivo par de *HRIR*,

$$h_{ci}(t) * h_{e\varphi_i\beta_i}(t)$$

$$h_{ci}(t) * h_{d\varphi_i\beta_i}(t)$$

e executando o somatório destas convoluções

$$BRIR_e = \sum_{i=0}^N h_{ci}(t) * h_{e\varphi\beta}(t)$$

$$BRIR_d = \sum_{i=0}^N h_{ci}(t) * h_{d\varphi\beta}(t)$$

obtemos a *resposta impulsiva binaural do ambiente* considerando N *caminhos de reverberação*.

Para proceder a convolução das *HRIR* com as respostas de cada *caminho de reverberação* é necessário relacionar as direções de incidência (φ_i e β_i) com o *sistema*

de coordenadas relacionadas à cabeça que foi utilizado para as medidas das *HRIR*. Os grupos direcionais estarão alocados em setores que dividem este sistema. Felizmente, o sistema auditivo possui resolução espacial limitada [2] (*erro de localização*) e por isso o número de setores aos quais estarão relacionados um par de *HRTF* é finito. O número de setores utilizados varia de acordo com a partição da *RIR* em questão. Para convoluir as *HRIR* com o *som direto* é necessária uma estrutura de setorização mais refinada que para as outras partes. Diferentemente, para as *reflexões iniciais* e para a *cauda reverberante*, esta setorização é muito mais geral, englobando um número pequeno de setores [31,43]. Kuttruff [29] sugere para as três fases da evolução temporal da *RIR* o seguinte número de grupos direcionais: 40 para o *som direto*, 5 para as *reflexões iniciais* e para a *cauda reverberante*. Este expediente de atribuir números diferenciados de setores às partições da *RIR* torna o processamento mais eficiente.

A transformação dos sinais que atingem o ouvinte originários de cada *caminho de reverberação* em *sinais binaurais* pode ser feita de várias formas [5]. A mais simples delas é atribuir ao sinal monofônico de uma determinada direção uma representação estereofônica garantindo pontos de recepção com *diferenças interaurais*. Outra maneira é calcular a distribuição de um campo sonoro na superfície de uma esfera. Porém, para um sistema de *auralização* capaz de apresentar ao usuário um campo sonoro virtual perceptualmente equivalente a um campo real são necessárias as medidas das *HRTF* em *cabeças artificiais* ou no próprio ouvinte. No capítulo 5, os métodos de medidas e processamento destas funções de transferência serão explorados.

4.6 Convolução para Auralização

A *resposta impulsiva binaural do ambiente* (*BRIR*) incorpora as distorções lineares impostas por toda a *cadeia de auralização* ao sinal que chega aos tímpanos de um ouvinte. O sinal original, que será filtrado, está desprovido de informações do contexto do ambiente, sinal anecóico.

A última etapa do processo de *auralização* é justamente a convolução do par de *respostas impulsivas binaurais do ambiente* com o sinal “seco” da fonte sonora (sem informações de ambiente), a tensão $s(t)$, (eq. (4.25)):

4. Auralização Inteiramente Computacional

$$\begin{aligned}e_e(t) &= BRIR_e * s(t) \\ e_d(t) &= BRIR_d * s(t)\end{aligned}\tag{4.25}$$

onde $e_e(t)$ e $e_d(t)$ são os *sinais binaurais* de tensão que carregam todas as informações da fonte, do ambiente e do ouvinte.

As *BRIR* também podem ser medidas diretamente em um ambiente, com o uso de *cabecas artificiais* ou de um ouvinte, por intermédio das técnicas binaurais de gravação discutidas no capítulo 2. Estas medições permitem que as *BRIR* sintetizadas sejam comparadas com estas respostas medidas e que comparações entre os seus parâmetros acústicos sejam efetuadas. Pode-se realizar, também, confrontações perceptuais dos dois resultados, medidos e preditos, por intermédio de *auralização* destas duas respostas impulsivas. Este tipo de expediente torna-se bastante útil para avaliações a respeito da validade de sistemas de *auralização* [46].

4.7 Reprodução Binaural do Produto da Auralização

Como visto anteriormente, os *sinais binaurais* resultantes da convolução da *BRIR* com o sinal anecóico devem ser apresentados cada um ao seu respectivo tímpano. Existem duas formas de realizar tal tarefa. Na primeira delas, técnica *transaural*, um par de alto-falantes é utilizado nos quais um processamento adequado permite que o sinal de uma orelha não seja captado pela outra. Este sistema pressupõe audições em câmaras anecóicas. Na segunda forma de audição são utilizados fones de ouvido. A seguir estas técnicas serão discutidas. Porém, vale a pena ressaltar que estamos interessados em sistemas que podem ser utilizados fora dos laboratórios o que corresponde ao uso, óbvio, de fones de ouvido.

4.7.1 *Transaural*

A reprodução das *pressões binaurais* pode ser feita por intermédio de um sistema estéreo em uma câmara anecóica, evitando, assim, a contaminação dos sinais pelo ambiente. Obviamente, para que cada canal, esquerdo e direito, alcance somente sua respectiva orelha, é preciso utilizar uma técnica conhecida como cancelamento de *crosstalk* (*crosstalk cancelation*) [5, 9, 47].

4. Auralização Inteiramente Computacional

Na figura (4.21) pode-se avaliar os princípios desta técnica. Utilizando, para simplificar a análise, as funções espectrais das pressões e dos sinais de tensões, teremos o seguinte conjunto de sinais com notação simplificada:

- S_e , sinal binaural da orelha esquerda que resultou da convolução da $BRIR_e$ com o sinal anecóico.
- S_d , sinal binaural da orelha direita que resultou da convolução da $BRIR_d$ com o sinal anecóico.
- $E_{af,e}$ = sinal nos terminais do alto-falante esquerdo.
- $E_{af,d}$ = sinal nos terminais do alto-falante direito.
- $P_{r,e}$ = pressão reproduzida para a orelha esquerda.
- $P_{r,d}$ = pressão reproduzida para a orelha direita.
- $H_{lado,lado}$ = funções de transferência de um alto-falante (esquerdo ou direito) para uma determinada orelha (esquerda ou direita).
- $P_{1,lado}$ = pressão na posição referente ao centro da cabeça sem o ouvinte fornecida pelo falante de um dos lados.

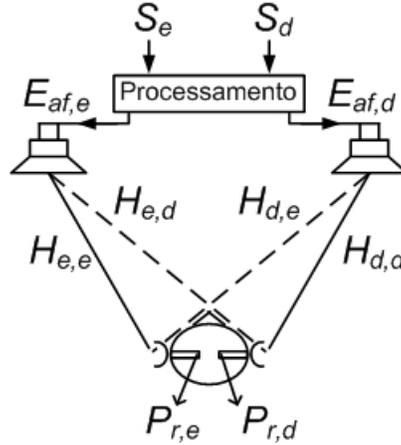


Figura 4.21: *Transmissão do sinal binaural nos terminais do alto-falante até os canais auditivos*

Os sinais reproduzidos para as duas orelhas podem ser expressos pelas seguintes equações (4.26):

4. Auralização Inteiramente Computacional

$$P_{r,e} = H_{e,e} \cdot E_{af,e} + H_{d,e} \cdot E_{af,d} \quad (4.26)$$

$$P_{r,d} = H_{e,d} \cdot E_{af,e} + H_{d,d} \cdot E_{af,d}$$

Para que a reprodução seja equivalente ao uso de fones de ouvido, os sinais reproduzidos devem ser equivalentes aos *sinais binaurais* resultantes da convolução (eq. 4.27):

$$P_{r,e} = P_{g,e} \quad (4.27)$$

$$P_{r,d} = P_{g,d}$$

Combinando as equações (4.26) e (4.27) e isolando os termos referentes aos sinais nos terminais dos alto-falantes, temos (eq. 4.28):

$$E_{af,e} = \frac{H_{d,d} \cdot P_{g,e} - H_{d,e} \cdot P_{g,d}}{H_{e,e} \cdot H_{d,d} - H_{e,d} \cdot H_{d,e}} \quad (4.28)$$

$$E_{af,d} = \frac{H_{e,e} \cdot P_{g,d} - H_{e,d} \cdot P_{g,e}}{H_{e,e} \cdot H_{d,d} - H_{e,d} \cdot H_{d,e}}$$

Assumindo que o ouvinte está simetricamente disposto em relação aos dois emissores sonoros, que (P_1/E_{af}) é a função de transferência de *campo livre* dos alto-falantes, que $(P_r/P_1)_{li}$ é a *HRTF* referente a orelha de lado igual ao do emissor, e que $(P_r/P_1)_{lo}$ é a *HRTF* referente à orelha do lado oposto, temos (4.29):

$$H_{e,e} = H_{d,d} = \frac{P_1}{E_{af}} \cdot \left(\frac{P_r}{P_1}\right)_{li} \quad (4.29)$$

$$H_{e,d} = H_{d,e} = \frac{P_1}{E_{af}} \cdot \left(\frac{P_r}{P_1}\right)_{lo}$$

4. Auralização Inteiramente Computacional

Substituindo estes termos (4.29) nas equações (4.28) e efetuando algumas manipulações algébricas, temos (4.30):

$$E_{af} = \frac{P_{g,li} - P_{g,lo} \cdot \frac{(P_r/P_1)_{lo}}{(P_r/P_1)_{li}}}{(P_r/P_1)_{li} \cdot (P_1/E_{af}) \cdot \left(1 - \frac{(P_r/P_1)_{lo}^2}{(P_r/P_1)_{li}^2}\right)} \quad (4.30)$$

onde o termo,

$$\frac{(P_r/P_1)_{lo}}{(P_r/P_1)_{li}}$$

significa a *função de transferência interaural (ITF)*, e o termo,

$$(P_r/P_1)_{li} \cdot (P_1/E_{af})$$

é a função de transferência de um alto-falante para um *ponto de referência* no interior do canal auditivo $(P_r/E_{af})_{li}$. Assim, simplificando a equação (4.30), temos (4.31):

$$E_{af} = \frac{P_{g,li} - P_{g,lo} \cdot ITF}{(P_r/E_{af})_{li} \cdot (1 - ITF^2)} \quad (4.31)$$

A equação (4.31) representa o processamento de sinal que deve ser efetuado nos sinais a serem *auralizados*. É importante lembrar que os sinais binaurais (S_{be} e S_{bd}) devem estar equalizados. Para tanto as medidas das *HRTF* precisam de um processamento adequado para compensar as respostas dos alto-falantes e dos microfones de medição.

Dois observações devem ser feitas a respeito das técnicas transaurais. A primeira delas é que a posição na qual o ouvinte deve estar localizado é muito específica, simetricamente entre os emissores. Este ponto de funcionamento otimizado é conhecido como *sweet-spot* e como consequência direta temos o fato de que o sistema permite apenas um usuário por vez. A segunda observação refere-se a problemas de *reverberações frente/costas* que costumam existir com o uso desta técnica. Para contornar esta anomalia podem ser implementados sistemas mais sofisticados que utilizam mais dois alto-falantes dispostos atrás do ouvinte [47].

4.7.2 Binaural

No capítulo 2, foi discutida a reprodução de *sinais binaurais* por intermédio de fones de ouvido. Viu-se a necessidade de *calibração* dos fones para que estes não introduzam colorações nos sinais a serem reproduzidos. As vantagens em sua

4. Auralização Inteiramente Computacional

utilização são a independência em relação ao ambiente e a apresentação exclusiva dos sinais para suas respectivas orelhas.

A escolha de um par de fones de ouvido deve ser feita baseada no conhecimento de algumas particularidades dos diversos tipos de dispositivos oferecidos no mercado [10]. Entre as características que devem ser avaliadas está a resposta em frequência, que deve ser a mais plana possível. O formato dos fones e a forma como estes se adaptam às orelhas também são importantes. Os fones de ouvido circumaurais (compartimento para orelhas), além de possibilitar o uso de transdutores mais planos, permitem um isolamento acústico muito mais eficiente que os fones supra-aurais (posicionamento sobre as orelhas) e também um acoplamento mais adequado.

Porém, mesmo que o transdutor do dispositivo possua uma resposta em frequência plana, a função de transferência que relaciona a pressão dos canais bloqueados ($p_2(t)$, ver capítulo 2) e os sinais de tensão nos terminais dos fones não será tão comportada. Em [10] foram levantadas estas funções de transferência de 14 fones comerciais para 40 indivíduos. A magnitude nas frequências baixas possui um comportamento relativamente suave enquanto que na região de frequências altas, principalmente acima de 8 KHz , as respostas são dominadas por diversos picos e vales bastante estreitos. Por isso, a melhor solução para garantir transparência na reprodução e uma consequente formação fidedigna do evento auditivo é a *calibração* dos fones e a compensação para essas funções de transferência.

Em relação à compensação das impedâncias acústicas (eq. 2.42), pode-se considerar que alguns fones comerciais possuem, de maneira aproximada, propriedades *FEC* [10, 48] (capítulo 2). Portanto o uso deste critério pode ser levado em conta para que a razão entre os divisores de pressão da equação (2.42) seja desconsiderada, evitando, assim, compensações para os acoplamentos acústicos.

CAPÍTULO 5

Modelagem da Fonte e do Ouvinte

Após a compreensão dos princípios que regem um sistema de *auralização* e dos métodos de modelagem mais comuns utilizados para calcular a *resposta impulsiva do ambiente*, se faz necessária uma discussão sobre os aspectos e as considerações mais importantes a respeito das modelagens da fonte e do ouvinte. Esses elementos também têm importância substantiva na constituição de ambientes acústicos virtuais já que a fonte é o objeto que se pretende *auralizar*, e o receptor é o ouvinte, usuário final do sistema. A modelagem do ouvinte, determinação da *resposta impulsiva relacionada à cabeça*, é utilizada para construção da *resposta impulsiva binaural do ambiente*.

5.1 Fonte

A modelagem da fonte sonora, para aplicações em ambientes acústicos virtuais, se refere à introdução de um sinal sonoro em um espaço determinando-lhe propriedades como diretividade e resposta em frequência. As abordagens mais simples para *auralização*, em relação ao sinal, são as gravações anecóicas ou a síntese sonora, que costumam ser utilizadas como sinal de áudio de entrada em alguns sistemas [5].

Para que uma fonte esteja bem representada dentro de um ambiente virtual é necessário que suas características direcionais sejam implementadas. É importante, também, que o sinal desta fonte obedeça a alguns requisitos básicos. O primeiro deles refere-se ao fato dos sinais serem *anecóicos* (*secos*), ou seja, não possuírem reverberação ou propriedades direcionais. Porém, em alguns casos, é desejado que estes sinais possuam algumas informações, como por exemplo, quando o objetivo é *simulação binaural* de sistemas estereofônicos comerciais. Normalmente, para os *métodos geométricos*, as fontes são pontuais e isto nos leva ao segundo requisito que implica em sinais monofônicos. Assim, sinais estéreo podem ser modelados por duas

5. Modelagem da Fonte e do Ouvinte

fontes pontuais. A qualidade dos sinais também deve ser garantida para que a *auralização* não seja afetada. Por isso são importantes, por exemplo, relação sinal/ruído, quantização e taxa de amostragem adequadas.

Desse modo, a modelagem da fonte sonora é constituída de dois pontos básicos: síntese sonora e características de radiação sonora. Porém, é preciso estar ciente das diferenças entre os tipos de fontes sonoras que costumam ser o alvo das *auralizações* de ambientes acústicos.

5.1.1 Natureza da fonte

Em ambientes acústicos há dois tipos de fontes sonoras: as *naturais* e as *não-naturais* [4]. As fontes não-naturais são os alto-falantes. Este tipo de fonte sonora possui características de diretividade e resposta em frequência bem determinadas que descrevem o seu comportamento deste dispositivo de maneira bastante precisa. Já a classe de fontes naturais se refere à voz humana e aos instrumentos musicais que não necessitam de transdutores eletroacústicos. Apesar de serem dispositivos técnicos, os instrumentos musicais tradicionais são denominados de fontes naturais devido à sua evolução técnica insignificante nos dias de hoje. Estes emissores se caracterizam, principalmente, pela complexidade acústica que envolve a sua geração e radiação sonora. A modelagem de fontes naturais é um dos grandes desafios em termos de implementação de sistemas acústicos virtuais.

5.1.2 Geração do sinal sonoro

O sinal da fonte pode ser implementado, em ambiente digital, a partir de gravações realizadas em câmara anecóica ou por intermédio de um método de *síntese sonora*. Dentre os métodos de síntese aqueles que são baseados em simulações físicas de instrumentos musicais passaram a ser muito utilizados na simulação de ambientes acústicos virtuais [25,26]. O *digital waveguide* é utilizado para síntese em tempo-real de instrumentos musicais unidimensionais tais como os instrumentos de corda [25]. Os requerimentos para a execução de modelagens físicas se baseiam em três pontos:

- Compreensão dos aspectos físicos dos instrumentos musicais.
- Entendimento das necessidades perceptuais básicas do sistema auditivo para uma síntese qualitativamente satisfatória.
- Capacidade de realização prática do modelo, de preferência em tempo real, para aplicação em ambientes virtuais.

O modelo geral para representação de um instrumento de cordas pode ser dividido em dois blocos principais [28] (fig. 5.1). O primeiro deles refere-se às cordas que são as fontes sonoras e estão conectadas ao segundo bloco, o corpo do instrumento, por intermédio da *ponte*. Fisicamente, a *ponte* permite que haja, também, uma interação entre as cordas.

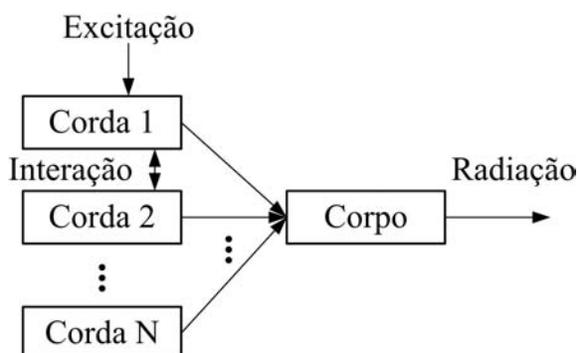


Figura 5.1: Diagrama de blocos da modelagem física de instrumentos de corda "puxada".

A solução geral da equação de onda, para cordas vibrantes [3], é composta de duas ondas transversais que viajam em sentidos opostos e refletem em suas extremidades com fase invertida formando ondas estacionárias. A modelagem deste sistema assume que a corda é um guia de onda bidirecional sem perdas (*digital waveguide*) e que as terminações serão representadas por filtros que modelam as reflexões e acumulam as perdas inerentes às cordas. A modelagem do corpo do instrumento, em geral, é feita utilizando uma única função de transferência que simula a radiação sonora num ponto frontal à *boca* do instrumento. Para a inclusão de características direcionais ao padrão de radiação, utiliza-se abordagens separadas para facilitar o processamento.

5.1.3 Direcionalidade

A abordagem mais simples em relação a este quesito refere-se ao uso de fontes pontuais omnidirecionais. Porém, para *auralizações* fidedignas, a diretividade em função da frequência deve ser implementada. Os fabricantes de alto-falantes costumam fornecer informações sobre os padrões de radiação de seus dispositivos. Em geral, instrumentos musicais tradicionais, em conjunto com o músico, emitem a maior parte das frequências altas pelo hemisfério frontal ao executor. Já as frequências baixas possuem um padrão omnidirecional.

5. Modelagem da Fonte e do Ouvinte

Dos modelos que implementam características de diretividade em ambientes virtuais e que são capazes de lidar com requerimentos para simulações em tempo-real, podemos citar dois dos mais importantes: *filtros direcionais* e *conjunto de fontes elementares*. O método do *conjunto de fontes elementares* se refere a fontes sonoras cujos padrões de radiação e diretividade encontram-se distribuídos e, desta forma, não podem ser modelados por um único ponto de emissão [26]. Exemplos de tais fontes são pianos grandes, devido às dimensões, e flautas, devido aos dois pontos de radiação sonora, a embocadura e a abertura principal. A abordagem por filtros implementa estes dispositivos, em meio digital, nos quais estes descrevem o espectro emitido em uma direção determinada [25, 26, 36]. Algumas medidas de emissões da fonte sonora devem ser feitas, em câmaras anecóicas, para obtenção dos dados que definem estes filtros. Assim, define-se uma direção de referência, normalmente com elevação e azimute iguais a zero, e coletam-se os sinais desta e da direção desejada ao mesmo tempo. O *filtro de diretividade* referencial é unitário. Os outros filtros representarão a variação da direção medida em relação à referência [26]. Esses métodos funcionam bem para modelagens de ambientes segundo os *métodos geométricos* já que cada direção de emissão representa um espectro do sinal da fonte.

Os *filtros diretivos* costumam ser implementados segundo condições de fase-mínima. As informações fornecidas pelos fabricantes dos alto-falantes sobre diretividade não incluem dados sobre a fase. Para as *fontes naturais* existem muitos dados de medidas disponíveis, porém, no que se refere à fase, também há uma grande deficiência de informação devido à dificuldade de medição [36].

5.2 Ouvinte

A modelagem do ouvinte envolve a compreensão das transformações lineares impostas pelo conjunto torso, cabeça e orelhas aos sinais acústicos incidentes que são transferidos aos tímpanos. As *HRTF* são as funções que representam esta transferência. Os filtros digitais são utilizados para implementar as funções de transferência relacionadas à cabeça a partir do conhecimento do espectro destas.

5.2.1 Medidas de respostas impulsivas binaurais (HRIR)

Medir um par de respostas impulsivas binaural ($h_{e\varphi\beta}$ e $h_{d\varphi\beta}$) é aferir, por intermédio de microfones de prova, as respostas de pressão sonora resultantes nas membranas timpânicas de determinado sinal analítico emitido a uma distância apropriada [2, 9]. Essas medidas devem ser realizadas para várias direções de incidência com o intuito

de estabelecer a dependência das *HRTF* em relação ao azimute (φ) e a elevação(β) do sinal (fig. 5.2).

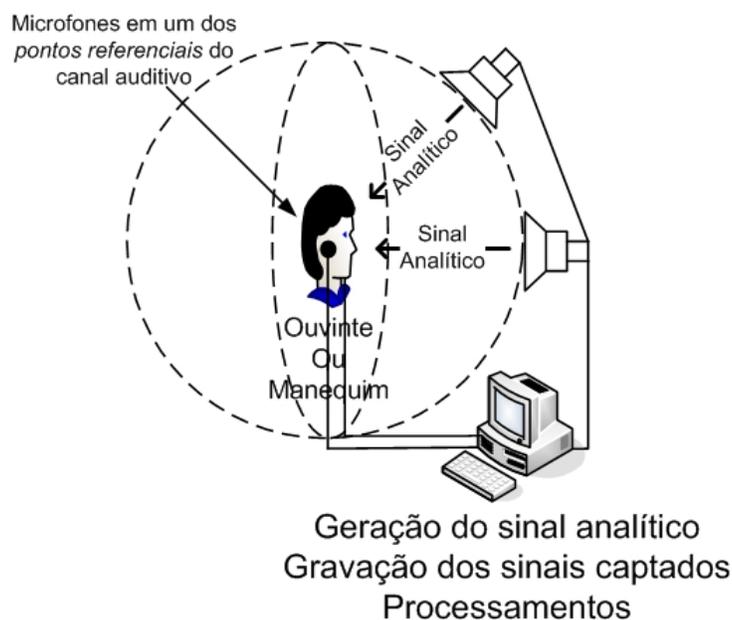


Figura 5.2: Medida de *HRTF*

Os sinais analíticos devem cobrir toda a faixa de frequências de interesse para o levantamento da função de transferência. As respostas nos canais auditivos são capturadas e comparadas com o sinal original. É importante ressaltar que sempre haverá algum ruído e por isto os sinais de excitação devem possuir energia relativamente alta para que a relação S/N seja elevada. O sistema usado para a realização das medidas, principalmente os alto-falantes, sempre introduzem não-linearidades em medidas de respostas impulsivas. Desta forma, o sinal de análise deve levar em consideração estes fatos para garantir a fidedignidade do processo de aferição das *HRTF*.

A função *Delta de Dirac* é o sinal de excitação que fornece de forma imediata a resposta impulsiva de qualquer *SLIT* [11]. Assim, os sinais captados no interior dos canais auditivos já serão as respostas binaurais. Contudo, para melhorar a relação *sinal/ruído*, ao invés de um único pulso, é necessário repetir a medida algumas vezes e efetuar uma média. Ao invés disso, pode-se optar por uma repetição periódica do sinal de forma que o pulso seja menor que o período. Assim, a resposta a cada período é somada e obtém-se a média. Pode-se, ainda, introduzir um desvio temporal para a esquerda eliminando o atraso de transmissão entre o alto-falante e os microfones [49].

5. Modelagem da Fonte e do Ouvinte

Os testes acústicos com impulsos são simples pois não requerem um processamento de sinal sofisticado. Porém, para obter relações S/N apropriadas à auralização, acima de 80 dB se possível [50], as câmaras anecóicas se tornam quase indispensáveis. Neste método as distorções devido às não-linearidades não podem ser identificadas nas respostas impulsivas.

Para além do *Delta de Dirac*, outros sinais analíticos podem ser utilizados com o intuito de melhorar a relação S/N e de remover distorções harmônicas da *HRIR* [34]. Dentre estes, podemos citar o *MLS* e o *SineSweep* [50]. O *MLS* (*Maximum-Lenght Sequences*) são sequências binárias periódicas e pseudorandômicas que possuem aproximadamente as mesmas propriedades estocásticas do ruído branco [49,50], apresentando uma magnitude espectral plana. Após a transmissão sonora, é realizada a deconvolução da sequência emitida com o sinal captado nos microfones. Este sinal possui boa imunidade em relação ao ruído branco porém não evita o aparecimento das não-linearidades distribuídas ao longo da *HRIR*. A sua aplicação em ambientes fora da câmara anecoica é aconselhada devido a esta imunidade.

O sinal conhecido como *SineSweep* consiste numa senóide com varredura contínua em frequência dentro da faixa a ser medida. O fato de ter seu espectro de frequências quase igual tanto para o sinal periódico quanto para um único *sweep*, agrega a este método velocidade nas medições e flexibilidade. Após a captação, a deconvolução é realizada por intermédio de um filtro processado previamente a partir do *sweep* emitido. Este filtro será uma função recíproca do espectro do *SineSweep* [11] de tal forma que a sua convolução com o sinal original resulte num *impulso*. Com o uso do *Sine Sweep* as distorções harmônicas inerentes ao sistema de medida não estarão inclusas na resposta impulsiva binaural. Ao invés disso, obteremos uma resposta para cada distorção, além da *HRIR*, devidamente separadas no tempo [50]. Esta, portanto, é uma técnica bastante robusta que permite relações S/N elevadas, rejeição a não-linearidades e flexibilidade.

O posicionamento dos microfones para a realização das medidas pode variar ao longo do canal auditivo sem que as informações espaciais entregues às orelhas sejam corrompidas [9]. Assim, qualquer ponto desde a entrada do canal até os tímpanos pode ser utilizado. Porém, as medidas realizadas com o canal bloqueado possuem algumas vantagens adicionais às demais. A primeira delas é o fato de que não há necessidade de introdução de microfone no canal auditivo. O transdutor é colocado num *plug* que veda o canal e não irá interferir na medida. A segunda grande vantagem

é que medições no interior do canal carregam muitas informações individuais do conduto auditivo. Assim, a pressão coletada com o canal bloqueado não só possui toda a informação direcional do campo sonoro como também o mínimo de informações individuais.

Com este *ponto referencial* de aferição, é necessário medir a relação $P_3(\omega)/P_2(\omega)$ (ver capítulo 2) que independe das direções do sinal acústico e, assim, pode ser medida para uma única direção. Utilizando o mesmo microfone de prova para medir tanto $P_3(\omega)$ quanto $P_2(\omega)$ elimina-se a resposta deste transdutor da relação desejada [48]. Para se tomar medidas em pontos imediatamente anteriores às membranas são necessários microfones de prova [2] de dimensões muito reduzidas. Nestas posições, as ressonâncias dos canais já estarão presentes na medida.

Para medir as *funções de transferência relacionadas à cabeça* levando em conta todas as direções de incidência, o sinal analítico precisa ser emitido de ângulos azimutais e de elevação variados (fig.5.2) [9]. Em alguns centros de pesquisas psicoacústicas existem gaiolas esféricas com alto-falantes distribuídos ao longo de sua superfície e com absorvedores que envolvem a estrutura metálica [34]. Seja qual for o formato do aparato de teste, um sistema de coordenadas esféricas relacionado à cabeça deve ser criado para referenciar as ondas incidentes. Este sistema formado pelos planos horizontal, médio e frontal definirá os grupos direcionais que serão medidos e posteriormente usados para a *auralização*. Desta forma, o par de *HRTF* para um determinado grupo direcional representado por φ_1 e β_1 , digamos $h_{e\varphi_1\beta_1}$ e $h_{d\varphi_1\beta_1}$, será convoluído com o sinal de reflexão correspondente àquele grupo de direções:

$$h_{e\varphi_1\beta_1}[n] * s_{\varphi_1\beta_1}[n] \quad (5.1)$$

$$h_{d\varphi_1\beta_1}[n] * s_{\varphi_1\beta_1}[n] \quad (5.2)$$

Nestas convoluções (eq. 5.1 e 5.2), as filtragens da onda incidente para a direção determinada já resultam nos sinais binaurais entregues aos tímpanos.

5.2.2 *HRTF generalizada*

Em programas computacionais que contemplem modelagem acústica de salas e que possibilitem simulação binaural, o uso de *HRTF* individualizadas é quase inviável pois o usuário, geralmente, não será objeto de medida para tal aplicação. Assim, surge a necessidade de *funções de transferência relacionadas à cabeça* generalizadas.

5. Modelagem da Fonte e do Ouvinte

Estas devem possuir, no mínimo, o comportamento médio de um conjunto de funções de transferência.

Para implementar uma *HRTF* generalizada pode-se proceder a análises físicas realizando-se médias de dados antropomórficos como tamanho do pavilhão auditivo e localização da entrada do canal para a confecção de *cabeças artificiais*. Outro procedimento é o levantamento de várias *respostas impulsivas binaurais*, realizar as suas transformadas de Fourier e realizar uma média destes espectros [34].

Pode-se, também, utilizar a técnica da *análise de componentes principais (PCA)* que serve para o reconhecimento de padrões em um conjunto que contenha muitos dados. Neste método, todas as *HRTF* em direções diferentes são avaliadas estatisticamente para mapear os comportamentos espectrais que se modificam de acordo com a direção e aqueles que não se alteram [34]. Como resultado, obtém-se um conjunto de *componentes principais* que podem originar qualquer *HRTF*.

Para além das médias que podem ser realizadas a partir de medidas com seres humanos, pode-se coletar as *HRTF* de *manequins* ou *cabeças artificiais* [51]. Há vários tipos de *manequins* disponíveis no mercado. Cada empresa confecciona o seu modelo de acordo com procedimentos próprios. Algumas, geram formas muito semelhantes ao torso, à cabeça e às orelhas de um homem real. Outras, constroem estes componentes com formas geométricas menos antropomórficas. Muitos destes já possuem os microfones de medida imbutidos nos canais auditivos e até mesmo um simulador para emissões sonoras bucais. A grande vantagem do uso desses dispositivos é a comodidade e a flexibilidade para medições e gravações binaurais.

Apesar da flexibilidade das *HRTF* generalizadas a auralização por intermédio destas é pior que a simulação efetuada quando a própria *função de transferência relacionada à cabeça* do usuário é utilizada [52]. A localização do *evento auditivo* é ainda mais pobre se as respostas binaurais derivaram de *cabeças artificiais*, mesmo daquelas de boa qualidade [51]. Na reprodução da *auralização*, através de fones de ouvido, o uso destas funções generalizadas podem agravar o aparecimento de *reversões frente/costas* e da *IHL*. Isto confirma a importância do conteúdo espectral dos *sinais binaurais* para que não ocorram anomalias relacionadas à *localização* (capítulo 2).

5.2.3 Processamento e equalização

No capítulo 2, foi discutida a necessidade de *calibração* dos fones de ouvido para que se possa proceder a uma equalização das *HRTF* em relação a estes e aos micro-

fonos utilizados nas medidas. Porém, o processamento e a equalização das *funções de transferências da cabeça* exigem outras intervenções além das correções para as distorções introduzidas por estes dois dispositivos.

Ao proceder às medições dos pares de *respostas impulsivas binaurais* para as direções de incidência predeterminadas, este conjunto de *HRIR* deverá sofrer intervenções no domínio do tempo e da frequência [34]. A primeira destas, é o descarte do lapso de tempo que o *signal analítico* leva para chegar do alto-falante até o ouvinte ou *cabeça artificial*. Para que isto seja feito, o sinal pertencente a orelha mais próxima do alto-falante de emissão, numa medição, determinará o lapso a ser desconsiderado. É interessante, também, promover um *janelamento* para eliminar erros introduzidos nas medidas devido ao cômputo de reflexões que ocorrem no aparato estrutural utilizado para o experimento [48].

Seguindo com as manipulações no domínio do tempo, executar-se-á uma normalização de amplitudes para que haja maior flexibilidade em relação à quantização. A amostra de referência para este procedimento será a de maior intensidade, obviamente a que possuía o menor lapso temporal, anteriormente citado. Desta forma, esta amostra deverá corresponder ao maior valor de quantização. Então, o coeficiente utilizado para normalizar a *HRIR* mais intensa, será aplicado à todas as respostas impulsivas.

As medidas das *HRTF* resultam em funções de transferência contaminadas pelos *sinais analíticos* e pelas informações espectrais do alto-falante, microfone e, se as medidas foram tomadas em algum dos *pontos de referência* com o canal auditivo aberto, pelas informações da ressonância deste conduto. Esta situação de aferição encontra-se exemplificada no diagrama de blocos da figura (fig. 5.3).

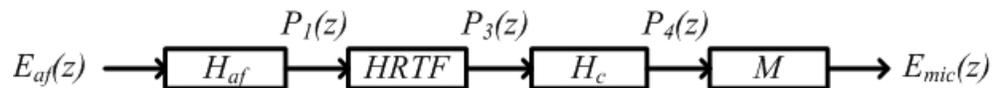


Figura 5.3: Diagrama de blocos: transmissão do alto-falante, nas medições, até a captação no tímpano.

Os espectros dos sinais e as funções de transferências são:

- E_{af} , sinal do alto-falante
- H_{af} , função de transferência do alto-falante
- $P_1[z]$, espectro da pressão de *campo livre* (centro da cabeça)
- $HRTF$, função de transferência da cabeça

5. Modelagem da Fonte e do Ouvinte

- $P_3[z]$, espectro da pressão na entrada do canal auditivo
- H_c , função de transferência do canal
- $P_4[z]$, espectro da pressão no tímpano
- M , função de transferência do microfone
- $E_{mic}[z]$, espectro da tensão nos terminais do microfone

Do diagrama percebe-se que a função de transferência total é:

$$\frac{E_{mic}[z]}{E_{af}[z]} = H_{af} \cdot HRTF \cdot H_c \cdot M \quad (5.3)$$

Isolando o sinal medido nos terminais do microfone:

$$E_{mic}[z] = H_{af} \cdot HRTF \cdot H_c \cdot M \cdot E_{af} \quad (5.4)$$

Da equação (5.4) vê-se a contaminação da $HRTF$ pelos dispositivos de medida e pela informação do canal auditivo. Esta será uma redundância que no momento da reprodução já estará presente na ressonância do canal auditivo do ouvinte. Deste modo, o sinal coletado $E_{mic}[z]$ deve ser multiplicado por,

$$\frac{1}{H_{af} \cdot H_c \cdot M \cdot E_{af}}$$

para que se obtenha uma *função de transferência relacionada à cabeça* que pode ser utilizada para *simulação binaural* de qualquer sinal, dentro da faixa de frequências medidas, com reprodução via fones de ouvido. O uso destes, pressupõe que a sua *calibração* e a conseqüente equalização entre os sinais do microfone e do fone seja efetuada para uma correta *auralização*, como visto no capítulo 2.

5.2.4 Modelagem com filtros digitais

As *funções de transferência relacionadas à cabeça* promovem modificações espectrais e atrasos temporais a uma frente de onda que atinge um ouvinte. Basicamente elas atuam num determinado sinal $x[n]$ provocando *ganhos* e *atrasos* em suas amostras. Como as $HRTF$ são funções que representam um $SLIT$, suas versões no domínio do tempo, as $HRIR$, podem ser representadas diretamente como *filtros FIR* [26].

As $HRTF$ possuem uma propriedade importante que torna as suas implementações factíveis: elas se comportam, aproximadamente, como sistemas de *fase-mínima*.

Algumas avaliações realizadas em implementações com base nesta propriedade demonstraram que, perceptualmente, não há inconsistências nas reconstruções de *HRTF* com esta abordagem [26, 34], ou seja, filtros de *mínima-fase* representam as *HRTF* de forma satisfatória. As *HRTF* incluem informações sobre o *ITD* e o *IID*. Assim, as *diferenças de tempo interaurais* estarão relacionados ao espectro de fase do filtro e as *diferenças de intensidade interaurais* à magnitude do espectro. Contudo, assumindo que a *HRTF* seja um *sistema de mínima-fase*, conseqüentemente, pode-se determinar o seu espectro de fase pela sua resposta em amplitude [11]. Isto, permite que o *ITD* seja tratado à parte da *HRTF* e possa até ser configurado, em situações de *HRTF generalizadas*, para a dimensão da cabeça de determinado usuário [34]. Os coeficientes para a construção dos filtros são obtidos diretamente das versões processadas das *respostas impulsivas binaurais*.

Begault [34], sugere um *filtro FIR* com 512 estágios de ganho e de atraso para a modelagem de uma *HRTF*. As funções de transferências podem ser processadas antes da implementação para a redução de dados e conseqüente diminuição de estágios do filtro. Uma dessas formas de redução pode ser conseguida por intermédio de *janelamento* temporal da resposta impulsiva. A *janela* deve ser aplicada a partir do mesmo ponto para todo o conjunto das *HRTF* com o intuito de preservar a informação de *ITD*. Outros métodos para redução de dados podem ser encontrados em [34].

Para além de implementações utilizando *filtros FIR*, existem outros métodos que visam aproximar respostas *HRTF* por outros tipos de filtros digitais. Tão explorado, na última década, quanto os filtros de resposta ao impulso finita [26], os *filtros IIR* também são aplicados em modelagens de *HRTF*.

CAPÍTULO 6

Considerações Finais

6.1 Conclusão

Tendo em vista a discussão e a análise das abordagens para modelagem e simulação binaural de ambientes acústicos virtuais, algumas conclusões a respeito da escolha dos métodos e das conseqüentes vantagens e desvantagens em relação à implementação computacional podem ser enumeradas.

A modelagem do campo sonoro e das interações deste com o ambiente precisa ser feita com abordagens diferenciadas. Essa necessidade fica evidente devido às relações de compromisso que precisam ser atendidas para garantir que o modelo contemple a faixa de frequências da audição, os fenômenos ondulatórios, as dimensões e as formas geométricas do espaço e ainda seja viável em termos computacionais.

Para que um sistema de *auralização* resulte em um *evento auditivo* convincente é preciso que o cômputo das *reflexões iniciais* seja efetuado com boa exatidão. Sendo assim, já que o *método de imagens da fonte* determina todas as direções de incidência desta partição da *RIR*, esta técnica torna-se uma opção robusta. Ele permite, ainda, uma flexibilidade interessante no que se refere aos movimentos do ouvinte, já que a *constelação de imagens* não precisará ser calculada novamente e isto é vantajoso para sistemas que pretendem interatividade em tempo real. Desta forma, porém, movimentos do ouvinte acarretarão em novos testes de visibilidade e, assim, um algoritmo eficiente de validação das *imagens*, como o *traçado de raios*, também pode ser utilizado.

Apesar de possibilitar boa flexibilidade e velocidade para sistemas interativos, o *ISM* possui carências em relação à difração e ao *espalhamento*. Para as *reflexões iniciais* a ausência de energia sonora resultante destes fenômenos pode causar uma percepção falsa do campo sonoro real. Para a difração é possível utilizar o *edge source* para minimizar a ausência deste fenômeno. Já para o *espalhamento*, sobretudo na fase

6. Considerações Finais

inicial da *RIR* na qual este tipo de energia não é predominante, podem ser utilizados métodos menos determinísticos como por exemplo um *RTM* de baixa resolução.

Os métodos baseados em soluções da equação de onda não são apropriados para simulações em tempo real devido às suas limitações em relação aos comprimentos de onda e às dimensões do ambiente. Porém, para processamentos efetuados previamente, como por exemplo da *cauda reverberante*, que é uma característica geral de um ambiente acústico e que independente da posição do ouvinte, o uso desses métodos para as frequências baixas pode ser vantajoso. Para as frequências mais altas da *cauda* o *RTM* pode ser utilizado.

Ainda em relação à *cauda reverberante*, os métodos estatísticos se apresentam como boas opções para aproximar o decaimento energético do campo sonoro no interior de um ambiente.

No que tange à modelagem da fonte sonora, as modelagens físicas possuem, apesar da sofisticação, a vantagem de não necessitarem de câmaras anecóicas como as gravações *secas*. Em relação ao ouvinte, a inviabilidade de se medir as *HRTF* de cada usuário nos leva às *HRTF* generalizadas. Esta opção permitiria que o sistema incluísse um conjunto destas funções de transferência para utilização. Desta forma, uma "imagem sonora" de calibração poderia ser utilizada para que o usuário pudesse escolher dentre as *HRTF* aquela que mais se aproxima da sua própria.

A apresentação dos *sinais binaurais* via fones de ouvido necessita que estes sejam *calibrados* para que a função de transferência dos fones seja compensada permitindo transparência dos sinais e a consequente formação correta do *evento auditivo*. O *ponto de referência* para medição destas funções que mais se adapta às necessidades práticas de utilização é à entrada do canal auditivo com este bloqueado, já que neste ponto as informações individuais, devido às ressonâncias do canal, estão minimizadas.

6.2 Trabalhos Futuros

Este trabalho mapeou os métodos de modelagem e simulação binaural que podem ser utilizados para a construção de sistemas de *auralização* com interatividade em tempo real. Como consequência direta deste trabalho surge a primeira sugestão de prosseguimento das pesquisas: implementação de um sistema de *auralização* com modelagem física da fonte sonora, modelagem híbrida para obtenção da *RIR* utilizando o *ISM* para as *reflexões iniciais*, e *HRTF* generalizadas. A audição seria executada com fones de ouvido devidamente *calibrados*.

As técnicas para modelar a *difração* e o *espalhamento* precisam de uma atenção especial e por isto é necessário um aprofundamento nas possibilidades de simulação destes fenômenos.

Como consequência de outras possibilidades vislumbradas no decorrer desta pesquisa podemos propor a implementação de sistemas de medidas de *respostas impulsivas de ambientes* e seus parâmetros acústicos, e de medidas da *BRIR* nestes espaços para avaliação perceptual da qualidade de sistemas de *auralização computacionais*. Isto é interessante pois seria uma espécie de referência perceptual e objetiva (comparação de parâmetros acústicos) para o sistema de simulação virtual.

Referências Bibliográficas

- [1] J. G. Roederer. *Introdução à Física e Psicofísica da Música*. Edusp, São Paulo, primeira edição, 2002.
- [2] J. Blauert. *Spatial Hearing - The Psychoacoustics of Human Sound Localization*. The MIT Press, London, England, second edition, 1997.
- [3] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders. *Fundamentals of Acoustics*. John Wiley & Sons, New York, USA, third edition, 1982.
- [4] H. Kuttruff. *Room Acoustics*. Elsevier Science Publishers LTD, London,UK, third edition, 1991.
- [5] M. Kleiner, B.-I. Dalenbäck, and P. Svensson. Auralization - an overview. *J. Audio Eng. Soc.*, 41(11):861–875, nov 1993.
- [6] T. Funkhouser, I. Carlborn, G. Elko, G. Pingali, M. Sondhi, and J. West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *ACM Computer Graphics, SIGGRAPH '98 Proceedings*, pages 21–32, jul 1998.
- [7] J. H. Rindel. The use of computer modeling in room acoustics. *Journal of Vibroengineering*, 4(3):219–224, 2000.
- [8] Tapio Lokki. *Physically-based Auralization - Design, Implementation, and Evaluation*. PhD thesis, Helsinki University of Technology, Espoo, nov 2002.
- [9] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36:171–218, 1992.
- [10] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen. Transfer characteristics of headphones measured on human ears. *J. Audio Eng. Soc.*, 43(4):203–217, April 1995.
- [11] A. V. Oppenheim and R. W. Schaffer. *Discrete-time Signal Processing*. Prentice-Hall International, Inc, New Jersey, 1989.
- [12] N. Xiang and J. Blauert. Binaural scale modeling for auralization and prediction of acoustics in auditoria. *Applied Acoustics*, 38:267–290, 1993.

REFERÊNCIAS BIBLIOGRÁFICAS

- [13] M. R. Schroeder. Digital simulation of sound transmission in reverberant spaces. *J. Acoust. Soc. Am.*, 47(2 Part 1):424–431, 1970.
- [14] M. R. Schroeder. Natural sounding artificial reverberation. *J. Audio. Eng. Soc. Am.*, 10(3):219–223, jul 1962.
- [15] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoustic Soc. Am.*, 65(4):943–950, apr 1979.
- [16] J. P. Vian and J. Martin. Binaural room acoustics simulation: Practical uses and applications. *Applied Acoustics*, 36:293–305, 1992.
- [17] B.-I. Dalenbäck, M. Kleiner, and U. P. Svensson. Audibility of changes in geometric shape, source directivity, and absorptive treatment - experiments in auralization. *J. Audio Eng. Soc.*, 41(11):905–913, nov 1993.
- [18] B.-I. Dallenback. Catt-acoustic. <http://www.negt.se/catt>. Gothenburg, Sweden.
- [19] G. M. Naylor. Odeon - another hybrid room acoustical model. *Applied Acoustics*, 38(2–4):131–143, 1993.
- [20] J. H. Rindel and C. L. Christensen. Room acoustic simulation and auralization-how close can we get to the real room? WESPAC 8, The Eighth Western Pacific Acoustics Conference, apr 2003. 7–9.
- [21] A. Farina. An example of adding spatial impression to recorded music: signal convolution with binaural impulse responses. In *Proc. of Acoustics and Recovery of Spaces for Music*, Ferrara, Italy, oct 1993.
- [22] A. Farina. Ramset - a new pyramid tracer for medium and large scale acoustic problems. In *Proc. of Euro-Noise 95*, Lyon, France, mar 1995.
- [23] T. Funkhouser, P. Min, and I. Carlborn. Real-time acoustic modeling for distributed virtual environments. In *In SIGGRAPH '99 Conference Proceedings. ACM SIGGRAPH*, jul 1999.
- [24] W. Ahnert and R. Feistel. Ears auralization software. *J. Audio Eng. Soc.*, 41(11):894–904, nov 1993.
- [25] Lauri Savioja. *Modeling Techniques for Virtual Acoustics*. PhD thesis, Helsinki University of Technology, Espoo, aug 2000.
- [26] J. Huopanieme. *Virtual Acoustics and 3-D Sound in Multimedia Signal Processing*. Doctoral thesis, Helsinki University of Technology, 1999. Report 53.
- [27] A. B. Carlson. *Communication Systems. An introduction to signal and noise in electrical communication*. Mcgraw-Hill, third edition, 1986.

- [28] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy. Physical modeling of plucked string instruments with application to real-time sound synthesis. *J. Audio Eng. Soc.*, 44(5):331–353, May 1996.
- [29] H. Kuttruff. Auralization of impulse responses modeled on the basis of ray-tracing results. *J. Audio Eng. Soc.*, 41(11):876–880, nov 1993.
- [30] D. T. Murphy, C. J. C. Newton, and D. M. Howard. Digital waveguide mesh modeling of room acoustics: Surround-sound, boundaries and plugin implementation. In *Conference on Digital Audio Effects*, pages 1–5, Limerick, Ireland, dec 2001.
- [31] R. Heinz. Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics*, 38(2–4):145–159, 1993.
- [32] Y. Korenaga and Y. Ando. A sound-field simulation system and its application to seat-selection system. *J. Audio Eng. Soc.*, 41(11):920–930, nov 1993.
- [33] T. Hidaka. Sound-field simulator for room acoustic design and assessment-introduction of wave theoretical treatment to synthesized sound. *J. Audio Eng. Soc.*, 41(11):914–919, nov 1993.
- [34] Durand R. Begault. *3-D Sound for Virtual Reality and Multimedia*. Ames Research Center, Mosfett Field California, april 2000.
- [35] U. Peter Svensson. Modelling acoustic spaces for audio virtual reality. In *IEEE Benelux Workshop on Model based Processing and Coding of Audio*, Leuven, Belgium, nov 2002.
- [36] H. Lehnert and J. Blauert. Principles of binaural room simulations. *Applied Acoustics*, 36:259–291, 1992.
- [37] H. Lehnert. Systematic errors of the ray-tracing algorithm. *Applied Acoustics*, (38):207–221, 1993.
- [38] M. Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoust. Soc. Ame.*, 86(1):172–178, July 1989.
- [39] U. R. Kristiansen, A. Krokstad, and T. Follstad. Extending the image method to higher-order reflections. *Applied Acoustics*, 38:195–206, 1993.
- [40] T. Lokki, P. Svensson, and L. Savioja. An efficient auralization of edge diffraction. In *AES 21st International Conference*, pages 166–172, ST. Petersburg, Russia, June 2002. AES.

REFERÊNCIAS BIBLIOGRÁFICAS

- [41] T. Lewers. A combined beam tracing and radiant exchange computer model of room acoustics. *Applied Acoustics*, 38:161–178, 1993.
- [42] E.-M Nosal, M. Hodgson, and I. Ashdown. Improved algorithms and methods for room sound-field prediction by acoustical radiosity in arbitrary polyhedra rooms. *J. Acoust. Soc. Am.*, 116(2):970–980, aug 2004.
- [43] E. Granier, M. Kleiner, B.-I Dalenbäck, and P. Svensson. Experimental auralization of car audio installations. *J. Audio Eng. Soc.*, 44(10):25–28, oct 1996.
- [44] D. Botteldooren. Finite-difference time domain simulation of low-frequency room acoustic problems. *J. Acoust. Soc. Ame.*, 98(6):3302–3308, December 1995.
- [45] R. R. Torres, M. Kleiner, and U. P. Svensson. Edge diffraction and surface scattering in auralization. In *ACM SIGGRAPH campfire*, may 2001.
- [46] T. Lokki and H. Järveläinen. Subjective evaluation of auralization of physics-based room acoustics modeling. In *International Conference on Auditory Display*, Espoo, Finland, July 2001.
- [47] H. W. Gierlich. The application of binaural technology. *Applied Acoustics*, 36:219–243, 1992.
- [48] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43(5):300–321, May 1995.
- [49] S. Müller and Paulo Massarani. Transfer-function measurement with sweeps. *J. Audio. Eng. Soc.*, 49(6):443–471, June 2001.
- [50] G.-B. Stan, J.-J. Embrechts, and D. Archambeau. Comparison of different impulse response measurement techniques. *J. Audio. Eng. Soc.*, 50(4):249–266, April 2002.
- [51] P. Minnaar, S. K. Olesen, F. Christensen, and H. Moller. Localization with binaural recordings from artificial and human heads. *J. Audio Eng. Soc.*, 49(5):323–336, May 2001.
- [52] C. I. Cheng and G. H. Wakefield. Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space. *J. Audio Eng. Soc.*, 49(4):231–249, April 2001.