

Media

TECHNOLOGY

传媒典藏

音频技术与声音艺术丛书

童雷 韩柯 著

交互式音频程序开发

CREATING INTERACTIVE
AUDIO APPLICATIONS
WITH PURE DATA



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

Media

TECHNOLOGY

传媒典藏

音频技术与声音艺术丛书

交互式音频程序开发

童雷韩柯著

CREATING INTERACTIVE
AUDIO APPLICATIONS
WITH PURE DATA

人民邮电出版社
北京

图书在版编目 (C I P) 数据

交互式音频程序开发 / 童雷, 韩柯著. — 北京 :
人民邮电出版社, 2018. 8
(音频技术与声音艺术丛书)
ISBN 978-7-115-47506-0

I. ①交… II. ①童… ②韩… III. ①音频技术—程
序设计 IV. ①TN912

中国版本图书馆CIP数据核字(2017)第307300号

内 容 提 要

这是一本关于如何为电子音乐、数字交互艺术开发音频程序的图书。该书从声学、数字音频以及计算机通信基础入手, 结合 Pure Data 示例程序分类讲解各种声音合成与实时处理技术的实现方法, 并介绍如何通过支持 MIDI、OSC 协议的交互设备为程序增加交互形式, 帮助声音设计师与艺术创作者完成交互式的声音作品。

-
- ◆ 著 童 雷 韩 柯
责任编辑 宁 茜
责任印制 彭志环
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京圣夫亚美印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 11.75 2018 年 8 月第 1 版
字数: 245 千字 2018 年 8 月北京第 1 次印刷
-

定价: 69.00 元

读者服务热线: (010)81055339 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

本书系北京电影学院“教学质量提高项目”研究成果

前言

计算机与网络技术为数字艺术带来了新的理念，在一些算法与交互艺术作品中，声音的设计就是一个音频程序的开发过程。本书正是为数字媒体作品中音频程序的开发教学而编写。

本书从声学与数字音频的基本知识入手，以程序设计语言 Pure Data 为平台，对音频处理与声音合成的实现方法进行讲解。Pure Data 是数字艺术领域广泛使用的声音设计工具，它采用图形化的操作方式，即使是没有计算机编程经验的声音设计师也可以使用 Pure Data 轻松地完成一个音频程序。

正如计算机音乐的先驱 Max Mathews 所说，“Max 与 Pure Data 可以让差不多所有人立刻合成出一个没有意思的音色。制作有趣的音色则要面临许多困难，需要更多的额外知识。”声音合成涉及一些数学与信号处理方面的理论，本书仅对这类技术的基本原理作出简要介绍，力图通过实用性的 Pure Data 程序示例帮助读者理解这些概念。如果需要深入掌握声音的处理与合成，可以阅读 Pure Data 作者 Miller Puckette 所著的《电子音乐技术》(*the Theory and Technique of Electronic Music*) 一书。此外，Andy Farnell 所著的《设计声音》(*Designing Sound*) 则从“过程式音频”(procedural audio) 的角度探讨了一些具体音响效果的程序实现，让声音设计师能够摆脱数字音频样本的束缚，将实时合成技术运用到自己的作品中。

交互是数字艺术的一种创作理念，交互作品通过创作者的规则设计让观众的行为能够改变作品的内容与形态。对交互作品中的音频程序而言，观众可能以各种行为去操控或者影响程序的执行，从技术上讲，这需要对音频程序获取反映观众行为的数据。本书的最后一章对交互式音频程序的实现技术进行了介绍，这包括 MIDI 与 OSC 两种主流的多媒体系统通信协议，以及一些易于使用的交互设备。这些技术可以让 Pure Data 开发的音频程序以多种形式与体验者进行交互。一些交互设备的使用可能涉及代码式的计算机语言，并要求使用者具有电子学方面的知识。不过，互联网与代码共享正在让交互式音频程序的开发变

2 前言

得轻松，相信科技发展将带来声音设计理念的再一次革新。

最后，感谢本书的编辑宁茜和熊一然，正是你们的细致工作让本书可以顺利出版。感谢研究生李子龙、胡晓同学为本书的程序插图进行整理与校对。感谢所有在 Pure Data 论坛为我提供帮助与建议的朋友。

目录

第 1 章	数字化的声音	1
1.1	声波与信号	1
1.2	乐音与傅里叶理论	4
1.3	时域分析与频域分析	5
1.4	数字信号与 PCM 音频系统	7
1.4.1	采样率	8
1.4.2	混叠失真	9
1.4.3	量化精度	10
1.5	音频编码与数据压缩	12
1.5.1	无损压缩	13
1.5.2	有损压缩	13
1.6	数据封装与信号重建	13
第 2 章	音频程序设计语言 Pure Data	17
2.1	Pd 的安装与设置	18
2.1.1	安装 Pd	18
2.1.2	Pd 操作界面	19
2.1.3	Pd 音频系统的设置与测试	19
2.1.4	Pd 启动参数	21
2.2	Pd 的基本元素	21
2.2.1	对象与程序	22
2.2.2	信息与信号	23
2.2.3	编辑模式与运行模式	23
2.2.4	帮助文档	24

2.3 使用 Pd 开发程序	24
2.3.1 示例 1 “hello world”	24
2.3.2 对象框与信息框	26
2.3.3 信息的类型与格式	26
2.3.4 对象的运行方式	28
2.3.5 图形化用户接口对象	29
2.3.6 注释框	29
2.3.7 示例 2 运算器	30
2.3.8 冷端与热端	30
2.3.9 对象的参数	31
2.3.10 “从右向左”与“深度优先”	31
2.3.11 示例 3 正弦信号发生器	33
2.3.12 “波浪号对象”	34
2.3.13 音量控制	35
2.3.14 示例 4 节拍器	37
2.3.15 示例 5 音序器	39
2.3.16 数值数组	40
2.3.17 数组的设置与操作	42
2.3.18 示例 6 包络发生器	45
2.3.19 “子程序”	50
2.3.20 “父程序开窗”	51
2.3.21 示例 7 模块化合成器	53
2.3.22 “\$”符号的用法	58
2.3.23 “抽象化”	61
第 3 章 信号分析与音频播放	63
3.1 信号分析工具	63
3.1.1 示例 8 示波器	63
3.1.2 示例 9 电平表	64
3.1.3 示例 10 频谱分析器	66
3.2 音频文件的播放与录制	72
3.2.1 示例 11 基于【readsf ~】的音频播放器	72

3.2.2	示例 12 基于【writsf ~】的录音程序	72
3.2.3	示例 13 基于“Pd 数组”的音频播放与录制	74
第 4 章	音频处理技术	77
4.1	滤波	77
4.1.1	滤波器的类型与参数	77
4.1.2	滤波器的实现	79
4.1.3	示例 14 直流偏移滤除器	81
4.1.4	示例 15 自动滤波器	82
4.2	延时	83
4.2.1	延时器的基本参数	83
4.2.2	延时处理的实现	84
4.2.3	示例 16 带反馈的延时器	86
4.2.4	示例 17 镶边效果器	87
4.2.5	示例 18 混响器	88
4.3	动态范围处理	91
4.3.1	压缩器的基本参数	93
4.3.2	示例 19 压缩器	95
4.4	傅里叶分析	97
4.4.1	快速傅里叶变换与逆变换	97
4.4.2	示例 20 基于快速傅里叶变换 (FFT) 的滤波器	99
4.4.3	示例 21 声码器	100
第 5 章	声音合成技术	103
5.1	基本波形的实现	103
5.1.1	正弦波	103
5.1.2	方波	104
5.1.3	锯齿波	105
5.1.4	三角波	106
5.1.5	白噪声	106
5.1.6	示例 22 基本波形发生器	107
5.2	减法合成	108
5.2.1	减法合成的实现	108

5.2.2	示例 23 减法合成器	108
5.3	加法合成	110
5.3.1	加法合成的实现	110
5.3.2	示例 24 加法合成器	112
5.4	调频合成	114
5.4.1	调频合成的实现	114
5.4.2	示例 25 调频合成器	115
5.5	波表合成	116
5.5.1	波表合成的实现	116
5.5.2	示例 26 矢量波表合成器	117
5.6	数字样本合成	118
5.6.1	样本合成的实现	118
5.6.2	示例 27 采样器	120
5.7	粒子合成	121
5.7.1	粒子合成的实现	121
5.7.2	示例 28 粒子合成器	124
第 6 章	开发交互式音频程序	127
6.1	通信系统的基本概念	127
6.1.1	通信链路	127
6.1.2	通信模式	128
6.1.3	数据传输速率	128
6.1.4	传输带宽	129
6.1.5	数据通信协议	129
6.1.6	校验方案	129
6.1.7	通信标准	129
6.2	音乐设备数字化接口 (MIDI)	130
6.2.1	MIDI 概述	130
6.2.2	MIDI 通信链路	130
6.2.3	MIDI 信息	134
6.2.4	在 Pd 程序中使用 MIDI	141
6.2.5	MIDI 通道信息的处理	142

6.2.7 MIDI 的局限性	146
6.3 开放式声音控制	147
6.3.1 OpenSoundControl 协议概述	147
6.3.2 在 Pd 程序中使用 OSC	153
6.4 常用交互设备	161
附录 1 MIDI 信息格式	169
附录 2 ASCII 代码表	173
附录 3 分贝值 - 功率比 - 振幅比对应关系	175

数字化的声音

本章将介绍一些声学与数字音频领域的基本概念，这些知识可以帮助你理解声音在数字世界的存在形式。

1.1 声波与信号

声音是由声源振动、声波传播和听觉感受三个部分共同构成的。我们把发出声音的振动物体称为声源，声源振动时，会引发周围弹性介质的波动，形成声波（sound wave），而声音则是人脑通过听觉器官对声波的感受。举例来说，扬声器的振膜振动时，会引发周围空气粒子的波动而产生声波，当我们所在的区域存在着声波时，人耳就会听到声音。

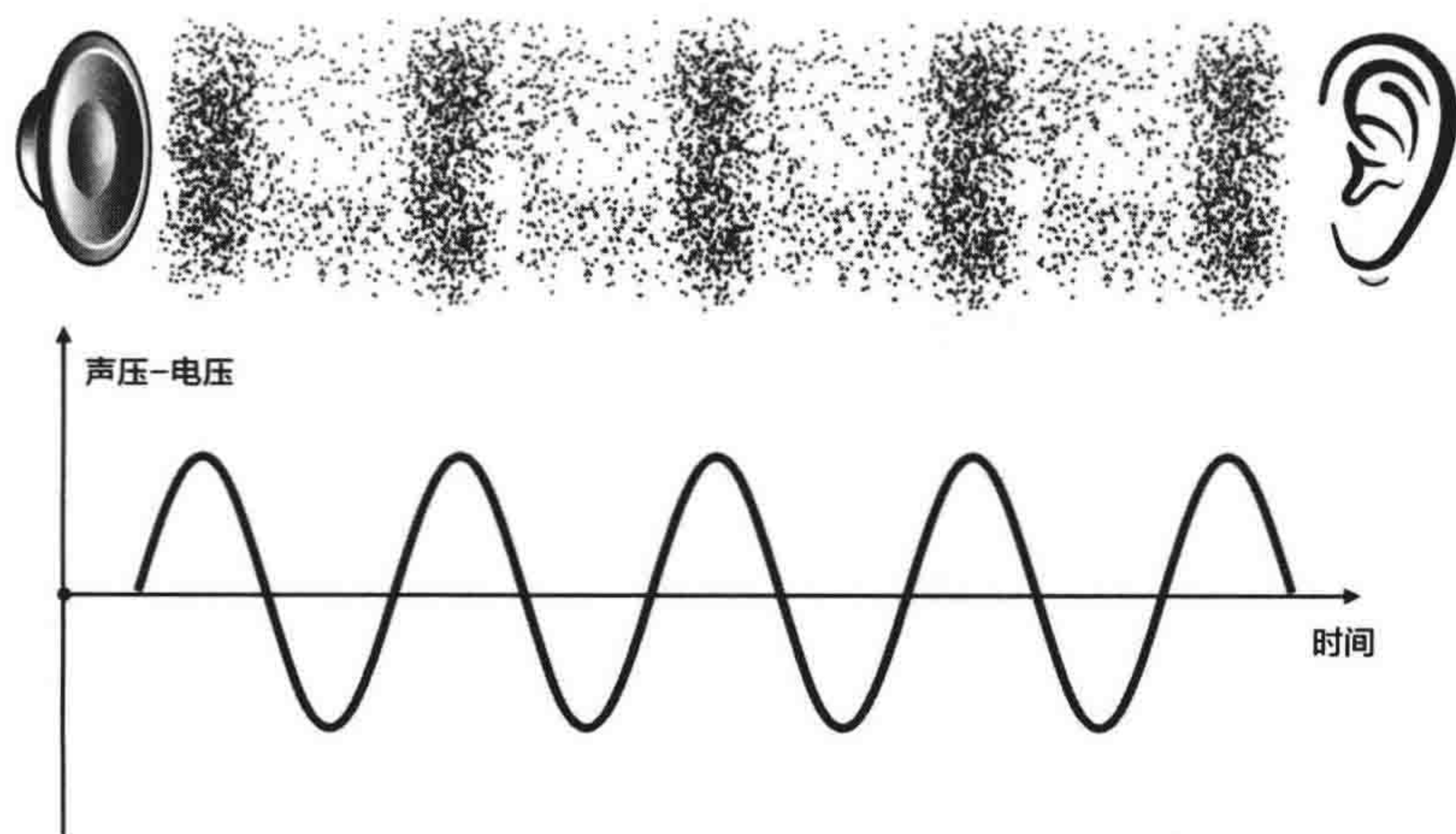


图 1.1 声波与信号

空气中的声波可以引发单位区域的气压变化,气压的变化量被称为声压(sound pressure)。如果沿着声波的传播方向把气压的变化记录下来,就会形成一个波形图(见图1.1),把波形图放入以“时间”为横轴,以“声压”为纵轴的直角坐标系中,声波就可以用一个声压随时间变化的函数“ $y = \text{Sound Pressure}(\text{time})$ ”来表示。当你使用类似话筒这样的换能器将气压变化转换成电能时,声波就表现为一个电压随时间变化的函数“ $y = \text{Voltage}(\text{time})$ ”,也就是一个音频信号(audio signal)。

如果声源的振动具有规律性,声波的波形就会呈现一个周期性重复的曲线。例如,音叉振动时的波形就类似一个周期性的正弦曲线,这让我们可以用一个包含频率(frequency)与振幅(amplitude)变量的方程来描述它的声波(见图1.3)。当然,并不是所有声波的波形都具有周期性,比如一个噪声的波形¹(见图1.2),不过我们通常以正弦波为模型来定义声波的物理特性。

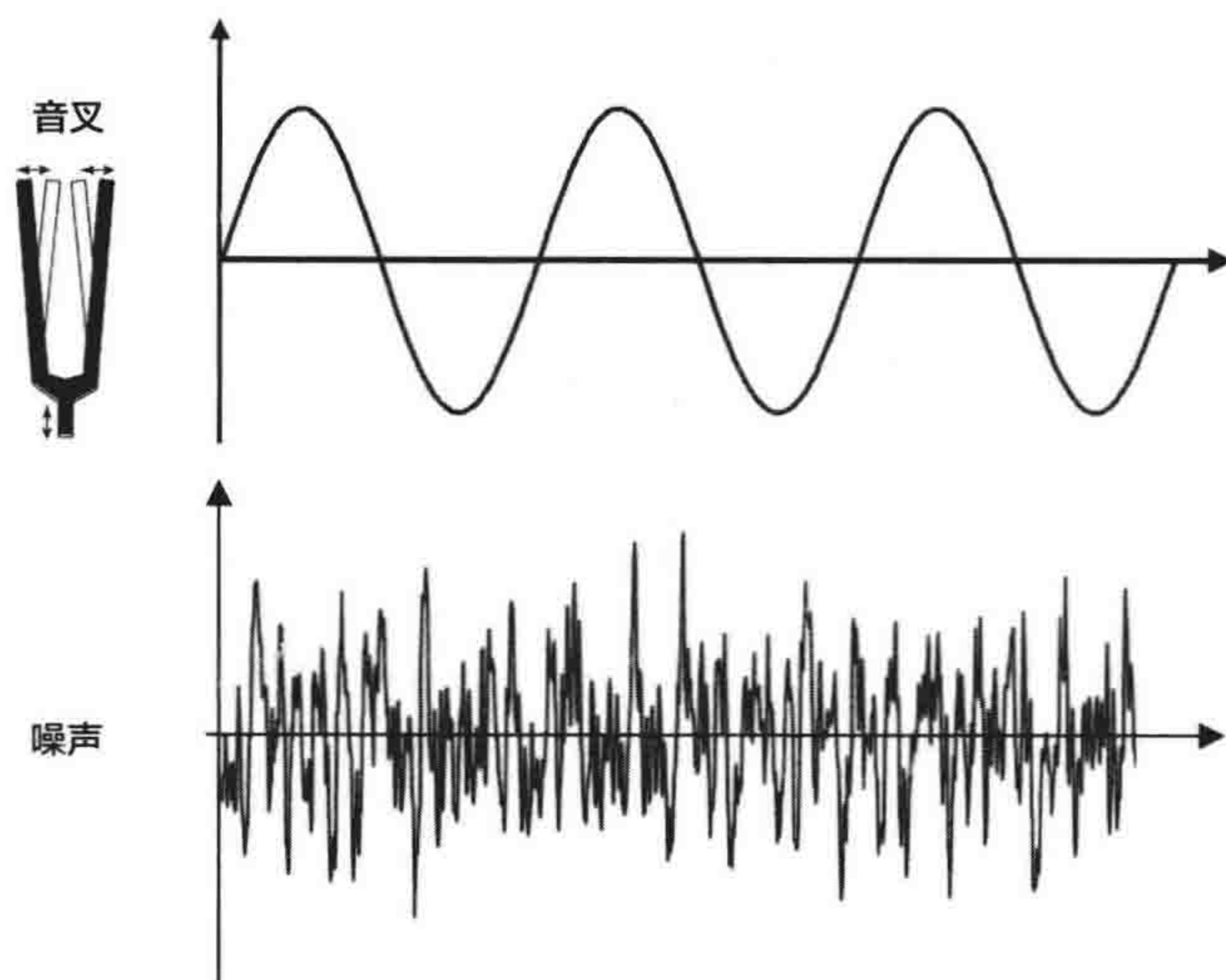


图 1.2 音叉与噪声的波形

声波的物理特性

声波的物理特性主要有频率、振幅、波长、周期,这里以图1.3为例来说明这些特性。

周期: 如果一个声波的波形由可分辨的重复图形组成,则每个重复图形的持续时间就称为这个声波的周期。

频率: 频率被定义为周期的倒数,单位是赫兹(Hz)。频率反映了声波在一秒钟内可以完成多少个完整周期。人耳可听到声波的频率在20赫兹(Hz)~20

¹ 这里的噪声是物理学上的定义,心理学上通常把干扰人们获取有效信息的声音称为噪声。

千赫兹 (kHz) 之间。

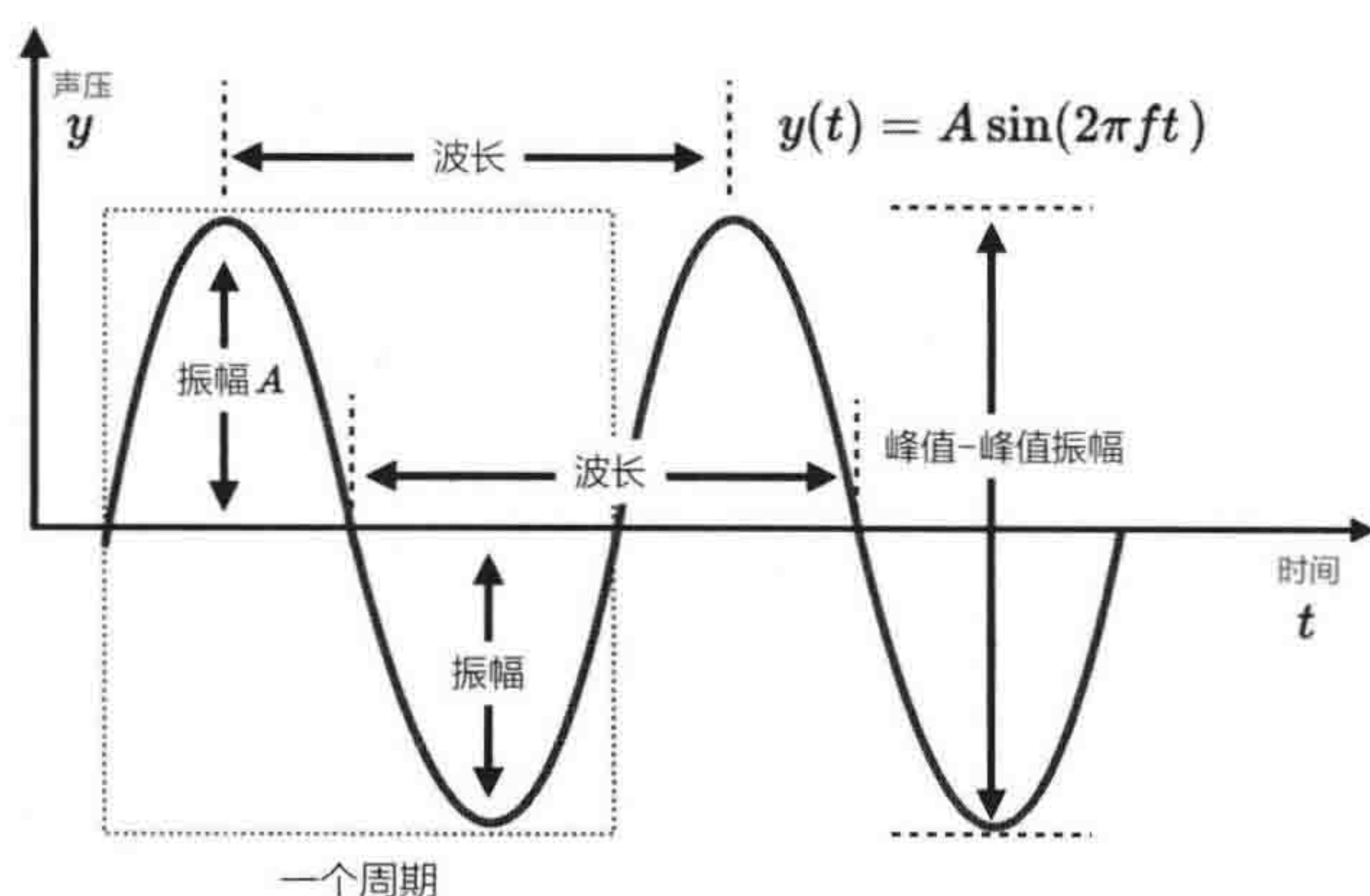


图 1.3 声波的物理特性：公式中 t 表示时间， A 表示振幅， f 表示频率

波长：波长是波形的一个周期在其传播方向的传播距离，它通常使用一个距离单位，比如“米”。

振幅：如图 1.3 所示，我们通常设 x 轴为振动的平衡位置，这时声音的波形在 x 轴的上下区域是对称的。声波的振幅【准确地说是峰值振幅 (amplitude)】被定义为波形曲线离开 x 轴的最大垂直距离。如果波形在 x 轴的上下区域不对称，我们可以使用峰峰值振幅 (peak-to-peak amplitude) 来描述这个波形最低点与最高点之间的距离。声波的振幅可以反映它改变气压的最大程度。

对于周期性的声波，我们还常使用相位来描述它在周期中的运行位置。当两个频率相同的声波在一起传播时，如果一个声波相对于另一个延迟了一些时间发出，它们之间就会出现相位差 (见图 1.4)。相位差可以反映两个同频率信

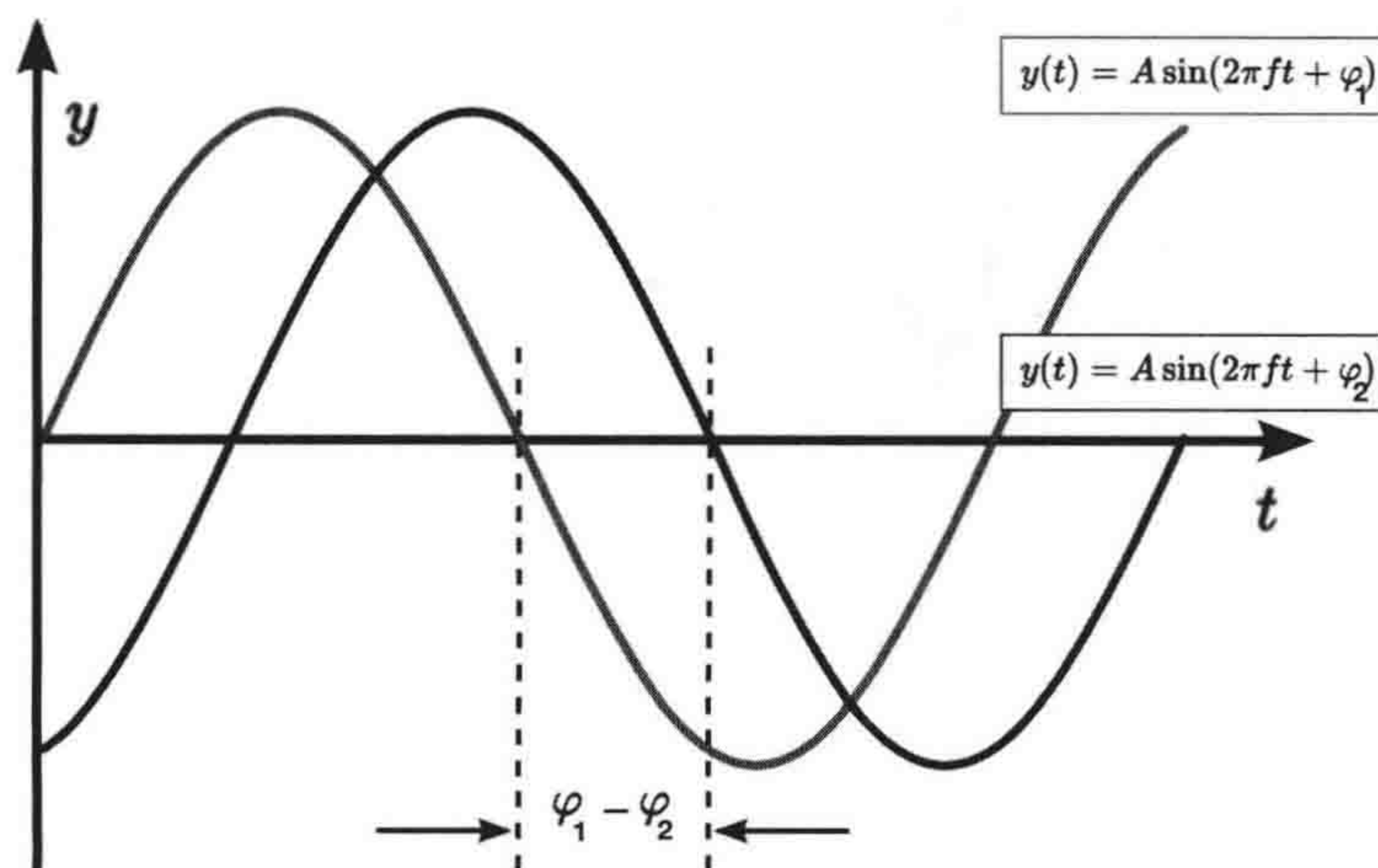


图 1.4 相位差

号之间的时间差，它的常用单位是度($^{\circ}$)。如果两个频率与振幅都相同的声波恰好相差 180° ，它们的音频信号将在叠加时相互抵消。

1.2 乐音与傅里叶理论

如果一个声音的波形呈现出稳定的周期性，就可能听出一个固定的音高(pitch，又称音调)。我们把具有固定音高的声音称为乐音(tone)。乐器所发出的声音基本上都是乐音，例如钢琴的每一个琴键都能发出不同音高的乐音。

最简单的乐音是纯音(pure tone)，它的波形是正弦波，而它的音高取决于正弦波的频率。严格地讲，现实世界中的物理乐器并不能发出真正的纯音，并且它们所发出声音的波形各不相同。而根据数学家傅里叶的理论，任何周期性的波形都可以用一系列具有特定频率、振幅以及相位关系的正弦波组合而成，这就意味着每一种乐音都可以由一系列不同音高的纯音组合而成(见图1.5及图1.6)。

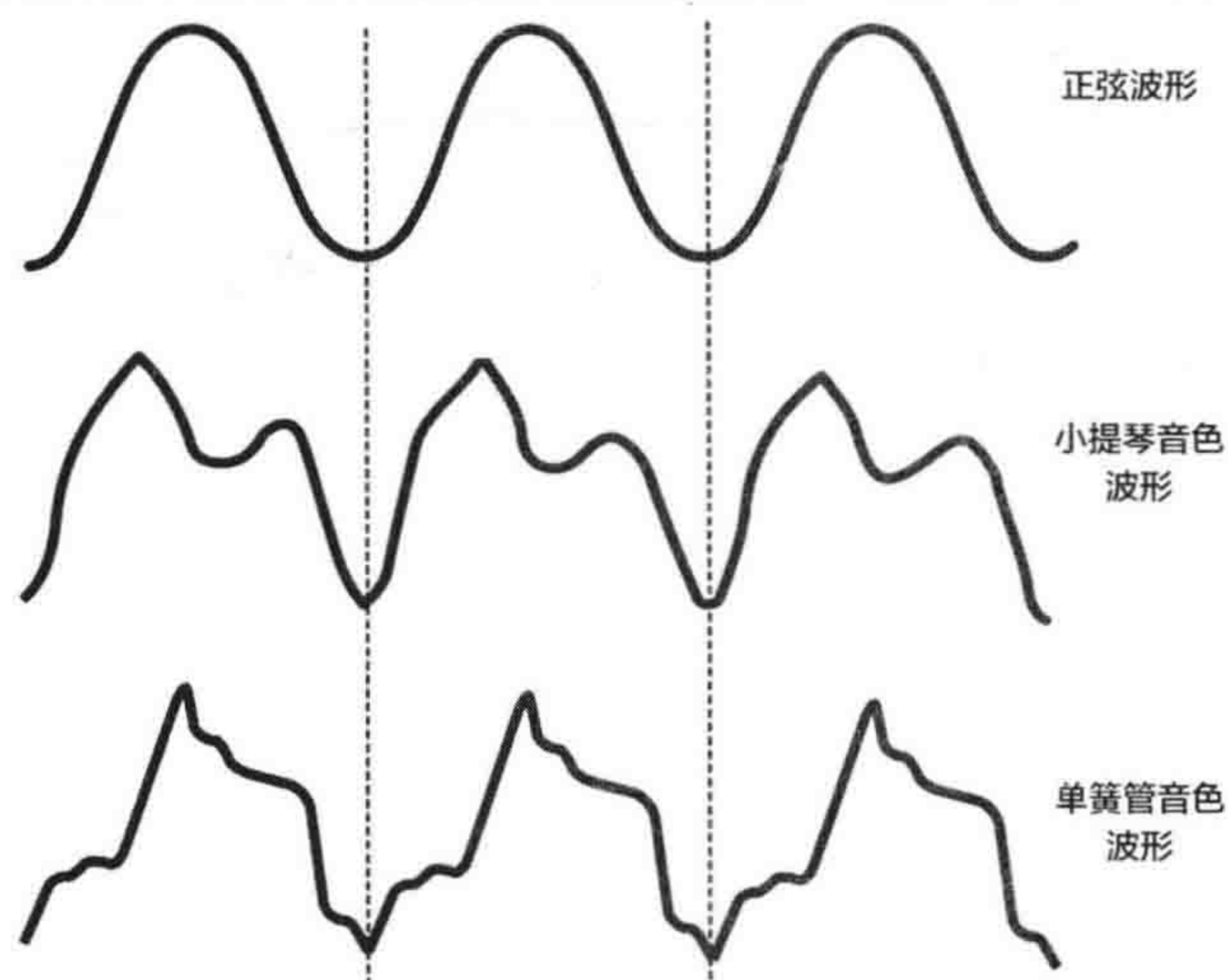


图 1.5 不同乐器的波形

基音、泛音、谐音

当我们把乐音看作一系列纯音的组合体时，其中决定乐音音高的纯音被称为基音(fundamental tone)，而其他的纯音被称为泛音(overtone)。通常，基音的响度比每一个泛音都大，并且它与泛音在频率值上存在着一定的比例关系。

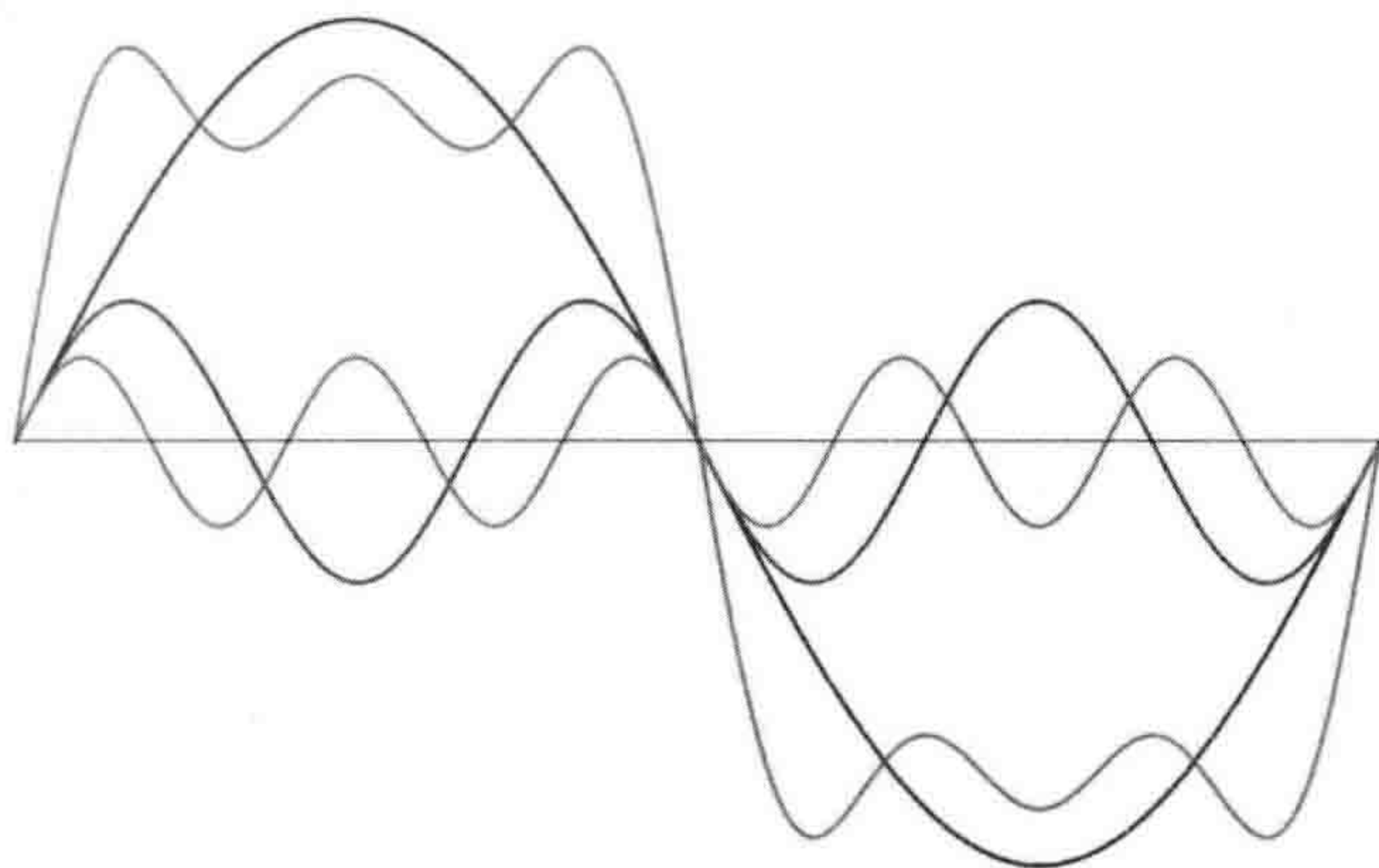


图 1.6 图中的浅色波形由三个正弦波组合而成，其中振幅最大的正弦波形对应基音，频率为基音 2 倍、3 倍的正弦波形分别对应第二谐音与第三谐音

基于上述规律，我们把频率等于基音的纯音称为第一谐音（first harmonic），频率等于基音两倍的纯音称为第二谐音（second harmonic），三倍的称为第三谐音（third harmonic）²，依此类推。如果一个声音仅由一些频率成整数倍增加的谐音组成，它的波形就会呈现出理想的周期性，音高也会清晰而稳定。当然，真实乐器的一系列泛音不可能像谐音一样符合严格的整数倍关系，因此，当小提琴和钢琴发出相同音高的音时，这两个乐音的基音频率是一致的，但它们的泛音不会相同，而这是造成钢琴与小提琴，甚至两把小提琴之间音色（timbre）不同的原因之一。

1.3 时域分析与频域分析

信号的波形反映了信号强度随时间的变化，使用波形图来描述与分析一个信号的方式被称为时域分析。根据傅里叶级数理论，周期性的波形可以由一系列特定频率的正弦波形组合而成，如果将一系列正弦波的频率记录下来，就形成了这个信号的频谱图（见图 1.7）。对周期性信号而言，组成它的一系列正弦波在频率上成整数倍关系，因此它的频谱图是一系列间断的点，这些点被称为信号的频率分量。我们把这种使用频谱图来分析信号频率分量的方式称为频域分析。

² 描述声音时，我们可能会混用泛音与谐音这两个词。严格意义上讲，谐音是用来抽象地指代频率为基音整数倍的纯音的，对真实乐器而言，它的泛音频率与谐音频率不一定相同，且泛音组合中不一定包含所有的谐音。

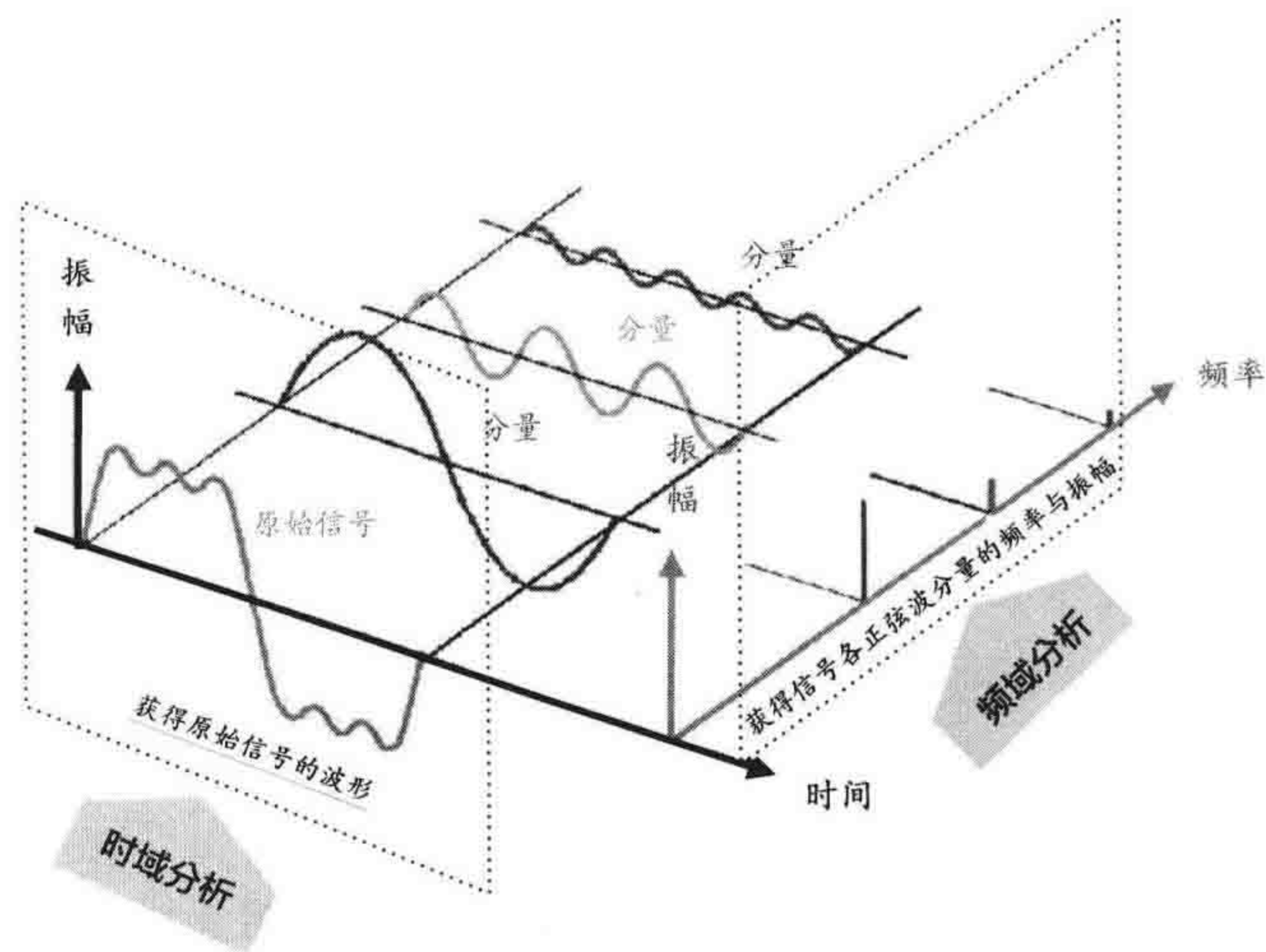


图 1.7 时域与频域

对一个音频信号进行时域分析可以获得其强弱变化的范围与速率，这也在一定程度上反映了声音的响度变化³，帮助我们判断信号的峰值出现在什么时刻。另外，对信号进行频域分析可以得到信号各频率分量的情况，对一个乐音而言，频率分量直接反映了其基音与泛音的频率和振幅，因此可以影响信号还放时的音色。需要说明的是，大部分音频信号并不具有周期性，但我们可以截取信号的一小段时间（比如 300ms）并把它不断重复，这样就能得到信号一小段时间上的频谱图。

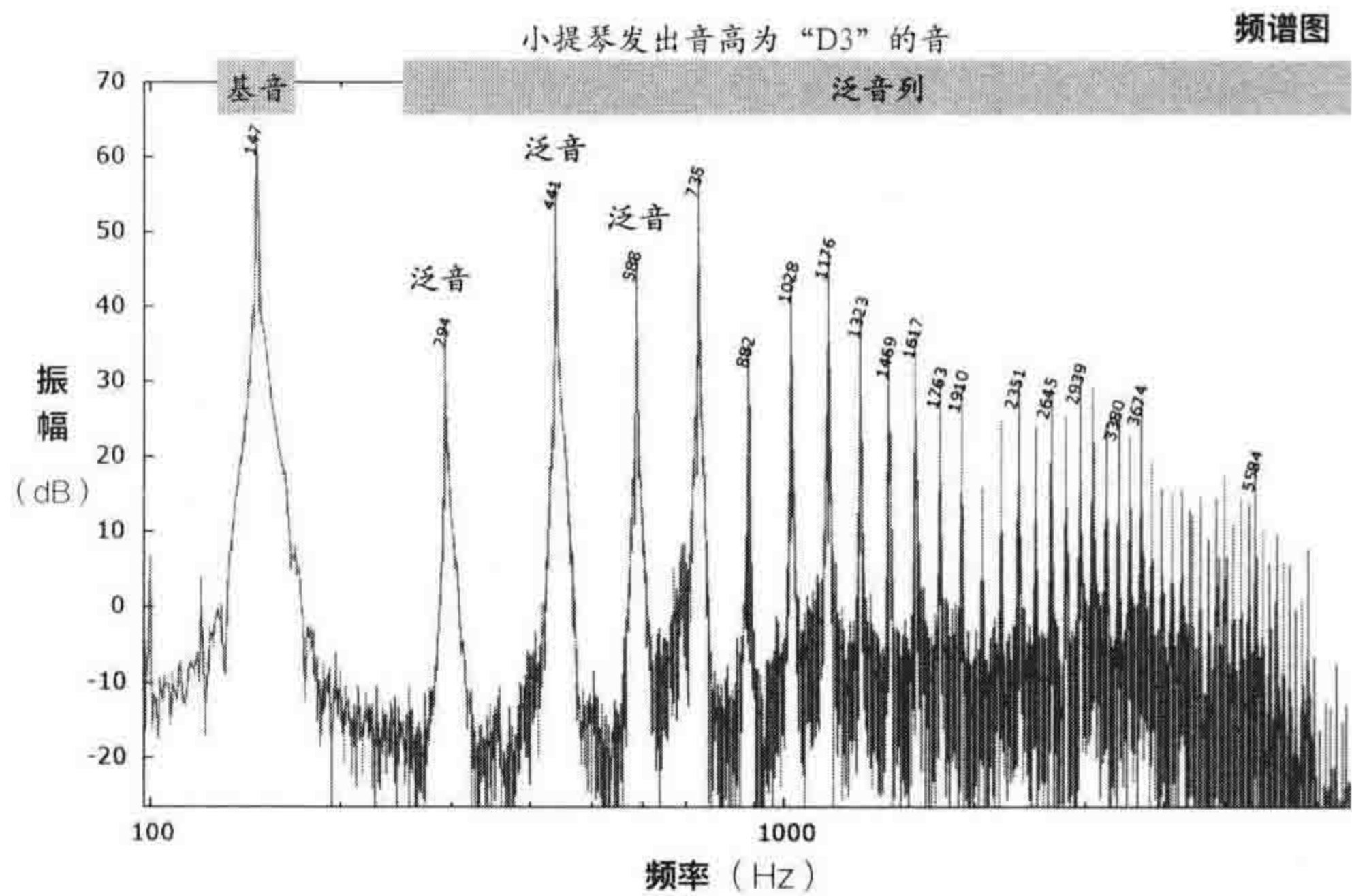


图 1.8 频谱图

分析音频信号时，我们也会组合使用时域分析与频域分析，让频谱图具有

3 声音的响度是一个主观量，事实上它的大小不只取决于信号的强度。