

第一节 声音是什么

本节中我们将介绍声音的物理本质以及与我们制作电声音乐密切相关的，声波在介质中传导的一些特性。了解声音的物理本质有助于我们理解合成器或音源中诸如振荡器、滤波器、包络发生器的概念与作用，对我们将来学习音色制作有很大的帮助。学习声音的传导特性有助于我们在电声环境中安排声音时做到井井有序，层次分明，主次清晰。

一. 声音的本质

下面我们将讨论声音的物理本质以及乐音的两大要素：音高及响度。

1. 振动和波

声音的物理本质非常简单，其实就是两个字：振动。我们对振动应该不陌生，振动的钟、振动的鼓膜、振动的音叉等等。那么到底振动如何引起的声音，又怎样能被我们听到呢？让我们以膜振动为例，见图 4.1.01。

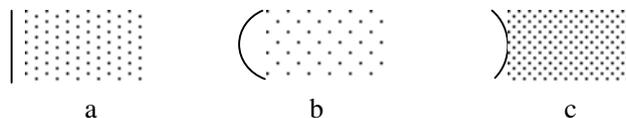


图 4.1.01 振动源引起的空气密度变化。

静止的薄膜与其右侧的空气密度如图 4.1.01a 所示。当薄膜向左摆动时，其右侧的空气密度减小，即变得稀疏，那么空气压力相对静止状态时就小(图 4.1.01b)；当薄膜向右摆动时，其右侧的空气密度增大，即变得稠密，那么空气压力相对静止状态时就大(图 4.1.01c)。当薄膜周而复始地由左至右摆动时，则使空气从稀薄到稠密，气压由小到大周而复始地变化。见图 4.1.02 所示：



图 4.1.02 振动的空气传播。

图 4.1.02 中位于左边振动源导致的空气波动象大海波浪一样，所以声音也叫做声波。声波传至图右的感知体，使感知体产生与振动源模式相同的振动。这样，振动源所产生的能量通过空气传给了感知体。人耳的骨膜正是类似这样的感知体。骨膜感受到的振动经神经传输给我们的脑，经过一系列复杂的神经传导过程，我们就听到了该声音。

由上我们得知：声音就是人耳感知到的物理的机械振动。

2. 音高

让我们再以图 4.1.01 为例：薄膜有规律地，周而复始地由左至右摆动。我们将由左至右的每一次摆动叫做一个循环，单个循环所占用的时间被称做周期。单位时间内完成周期重复的次数叫做频率。频率的测量单位是赫兹(Hz)，即每秒钟周期重复的次数。当频率达到一定数值时，人耳就会产生音高的感觉。可见，音高是人耳对周期性振动频率的主观反映。人类所能感知的频率是有限制的，其范围是 20 至 20000 赫兹。低于 20 赫兹的声音被称为次声波，高于 20000 赫兹的声音被称为超声波。

我们所熟悉的标准音 A 的频率为 440 赫兹。本书附录六《音名、音符号与频率对照表》列举了常用乐音与频率的关系。至于钢琴键盘中的每个音高与频率的换算关系牵涉到复杂的律学问题，枯

燥的计算公式以及诸如纯律、五度相生律等生疏的专有名词会令我们头晕目眩，故在此不加赘述。但做为常识我们应该知道，目前我们使用最普遍的是十二平均律。顾名思义，十二平均律既将一个八度平均分成了十二等份，但请注意，十二平均律中每个半音之间的频率差并不相等，相等的是它们之间的频率比！这一点提醒大家切记，切记。

人耳对音高的感知不是单一的，它受许多其它条件的影响。如相邻乐音的频率差、声音的振幅、音色、感知时间、演奏速度、年龄等诸多因素都会影响人耳对音高的感知。

音高是人耳对周期性振动的频率的主观反映。那么非周期性振动呢？当非周期性的振动被人耳接收时被感知为噪音，亦称噪声。噪声在声学应用中又有白噪声与粉红噪声之分。

3. 响度(Loudness)

图 4.1.01 中薄膜有规律地，周而复始地由左至右摆动。其摆动的幅度叫做振幅(Amplitude)。其摆动越大则振幅越大，其摆动越小则振幅越小。振幅代表振动引起的周围空气的变化大小，亦即声学能量的大小。我们人耳所感知到的响度就是振幅产生的声学能量。

强度(Intensity)是测量实际接触到某区域如耳鼓的能量大小的单位。由于表示能量的单位非常不利于计算，加之与音高的感知类似，人耳对声音强度的感知亦可呈对数特性，所以人们采用了分贝(dB)——对数比值做为比较强度的单位。当使用分贝比较声音的绝对强度时，我们将听觉门限时的声压级(SPL)定义为 0 分贝，其它声音强度比较都以此为基准。例如人类正常语言大约在 60 分贝左右，叫喊声能达到 80 分贝，而交响乐队全奏可以达到 120 分贝以上。

对一般人来说，对频率分别为 100 和 1000 赫兹，声压级都为 100 分贝的两个声音的主观感受是响度相同。但是人们感觉 40 分贝的 1000 赫兹要比相同声压级的 100 赫兹响度要大许多，大约强 20 分贝。有兴趣的读者可以通过其它专业声学书籍中列举的弗莱彻—芒森等响度曲线中获得更多的信息。另外，遮盖效应影响人耳对信号振幅的感知。当弱响度音和强响度音同时发声时，很可能听不到弱响度音。这之间的响度差大约为 15 至 20 分贝。

与音高的感知相同，人耳对振幅的感知也受许多其它条件的影响。

二. 音色(Timbre)

定量描述一个声音的音色要比描述声音的音高或振幅要复杂得多。因为对声音的音色而言，不存在象频率的高/低或振幅的大/小这样单维度的标量。在我们介绍音色的定量描述方法之前，我们必须先引入波型的概念。

1. 波型(Waveform)

波型是声音的一种定量的图形表示方法。即以时间为横坐标，以振幅为纵坐标，将振幅随时间变化模式的轨迹以图形方式记录下来。从某种意义上说，波型可以代表声音本身的“形状”，是声音在某一段时间内的图形定格。

声音的音色差别巨大，其波型亦各千姿百态。其中我们经常能够在各种相关书籍中见到的正弦波是最简单的波型之一。让我们来绘制一个正弦振动的波型。将一张绘有坐标系的纸平放于桌面上，手中的笔保持与桌面垂直。当我们匀速地向左方(以 X 轴为中心)移动该纸张时，沿 Y 轴做纵向匀速运动的笔将在纸上留下一条平滑的曲线。见图 4.1.03。其中 T 表示周期。

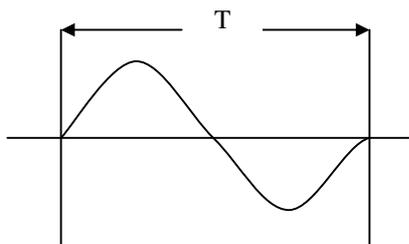


图 4.1.03 正弦波型示意图。

图中所示为一个周期的正弦波，如果在一秒钟内有 440 个这样的周期，那么该波型表示的结果将是一个以标准音 A 为音高的正弦音。除正弦波外，常见的简单波型还有方波(Square Wave)、锯齿波(Sawtooth Wave)、脉冲波(Pluse Wave)等。图 4.1.04 为其它简单波型的图示。

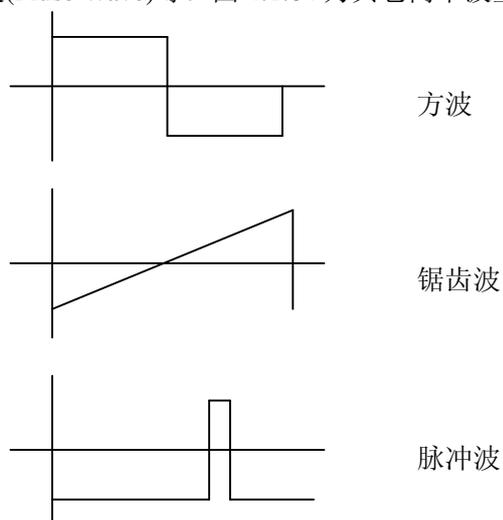


图 4.1.04 简单波型示意图。

自然界中不存在以上图示的简单波型，这些波型都是通过电气或电子方法得到的。再者，以上提及的波型都是周期性波型，那么，非周期性波型是什么样呢？见图 4.1.05，白噪声(White Noise)的波型图示。

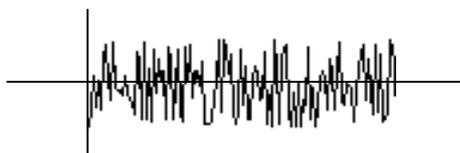


图 4.1.05 白噪声的波型。

图 4.1.05 为一段白噪声波型的放大显示。由图中可见，没有任何周期性的因素出现。前文中提到，波型是声音在某一段时间内的图形定格，它可以只包含一个循环周期，也可以是整个声音甚至整个音乐片段的造影。见图 4.1.06：

该波型表示的声音是我们比较熟悉的视窗默认的开始音乐。由图中可见，其振幅是由弱至强，然后逐渐减弱直至消失的。

有了声音的波型，我们可以通过这些声音的“照片”分析声音随时间变化的振幅模式，波型是量化表示音色的重要维度之一。



图 4.1.06 视窗开始音乐的波形。

2. 简单包络(Envelope)及 ADSR 包络

声音的音色多种多样，有的铿锵有力、掷地有声；有的绵绝悠长、回肠荡气；有的呆板单调，有的灵活易变。例如钢琴的音色清脆具颗粒性；风笛的声音绵软且悠长。非洲弯音鼓通过特殊的演奏方法可以改变音高。这些音色上的差异意味着声音的力度(振幅)响应或音高随着时间而变化。为了定量地表示音色在这些方面的差异，人们采用绘制该音色的包络线的方法来实现这一目的。见图 4.1.07，简单包络线示意图。

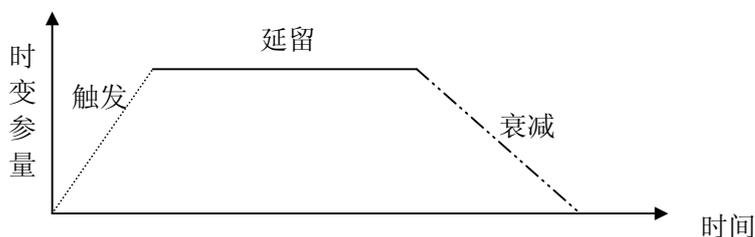


图 4.1.07 简单包络线示意图。

横坐标为时间，纵坐标为随时间变化的参量。虚线表示参量从最小值到最大值的过程，其延横坐标所用的时间叫做触发(Attack)时间或上升(Rise)时间；实线表示参量在某一取值延留或延持(Sustain)的情况，其延横坐标所占用的时间叫做延留时间；虚实相间的部分表示参量从最大值逐渐衰减(Decay)的过程，其延横坐标所用的时间叫做衰减时间。如果纵坐标的时变参量是振幅的话，该包络即叫做振幅包络，那么前文提及的钢琴和风笛的振幅包络如图 4.1.08 所示：

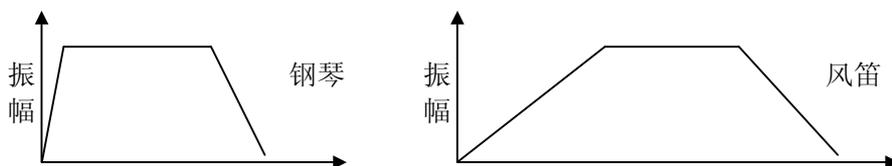


图 4.1.08 钢琴与风笛的振幅包络示意图。

就上例而言，随着手指敲击琴键的同时，钢琴声音立即做出回应，风笛声音的力度则是在吹奏开始后缓慢提升的；钢琴音色的力度在保持一定时间后很快地衰减下去，而风笛力度的衰减梯度则比较平缓。

对音高随时间变化的音色来说，就需要通过音高包络(Pitch Envelope)来表示了。见图 4.1.09，弯音鼓的音高包络示意图。

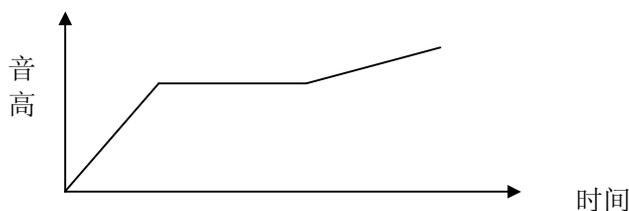


图 4.1.09 弯音鼓的音高包络线示意图。

弯音鼓的音高在触发阶段呈上升的形状，经过延留阶段后，其衰减阶段中音高还有轻微的上扬。

在数字电子乐器中，我们常用的包络是基本包络的扩展——ADSR 包络。它除了与基本包络同样具有触发、衰减和延留外，还引入了一个新的阶段，即在衰减阶段后面的释放(Release)阶段。见图 4.1.10，ADSR 包络示意图。

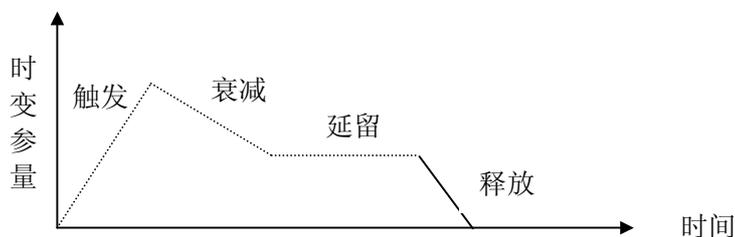


图 4.1.10 ADSR 包络线示意图。

释放意味着接收到音符关信息之后的参量变化。图 4.1.10 中实线部分表示参量在释放阶段时的变化，延横坐标所经过的时间叫做释放时间。

振幅包络做为定量描述音色的方法，与其它包络一起，同波型一样，在量化表示声音方面具有普遍的应用意义。

3. 泛音(Harmonic Partial)、声谱(Spectrum)及声谱包络

通过声音的振幅与音高包络，我们可以定量地分析声音力度、音高的响应差别。但振幅、音高包络在解释声音的某些特性如音色的明暗度时就显得无用武之地了。科学家们在不断的探索中发现了声音另一个维度的属性——泛音、声谱及声谱包络。

假设我们有一根长度为 1.2 米的琴弦。当这根弦在振动时，其 $1/2$ (0.6 米)处、 $1/3$ (0.4 米)、 $1/4$ (0.3 米)、 $1/5$ (0.24 米)等处都在振动。也就是说，虽然我们听到的音高是一个，但该弦的各部分的振动同样产生音高，只不过它们的音量很小，因此我们通常听到的音高就是一个，其它音高需要特意地用心听才能听到。声音在振动时的所有分量被称做分音(Partial)。其中由整根音弦振动时发出的，含声学能量最多的，既振幅最大、声音最响的分音在声学上被称做该声音的基频。

上例中，各部分振动频率分别为整个弦振动频率的 2、3、4、5 倍。假设基音是 C，其 2、3、4、5 倍音的音高分别为高八度的 C、一个八度以上的 G、两个八度的 C 和两个八度以上的 E。我们将这些整数倍的分音叫做谐波分音(Harmonic Partial)或者和声分音，亦称为泛音。其实，当声音的基频和泛音在振动时，还有许多不协和的分音在一起振动，只不过它们所包含的声学能量，对某些音色而言可能比谐波分音更小从而更不被人独立地感知。谐波分音与所有的分音对声音的音色有着重要的决定作用。

理论证明，任何周期性的波型都可以分解为多个正弦波的叠加，这就是著名的傅立叶理论。见图 4.1.11：基频与倍数分音以及含不同分音数的合成波型。

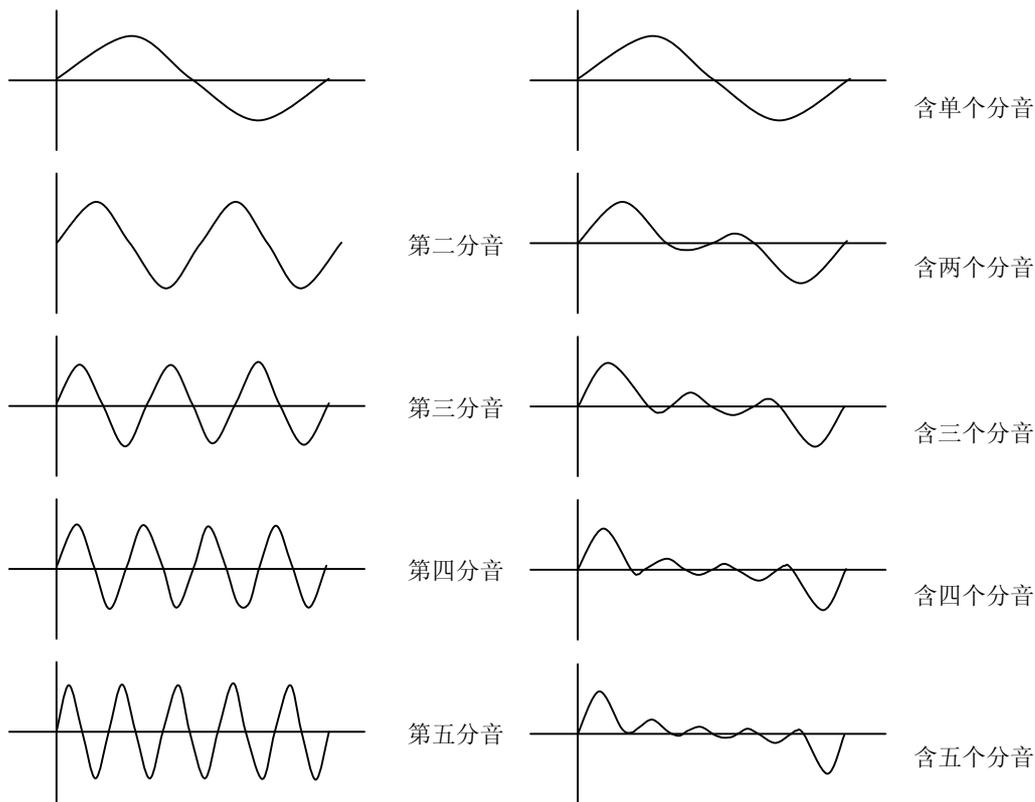


图 4.1.11 基频与倍数分音以及含不同分音数的合成波型。

图中左边为基频及其四个泛音的波型，右面第一个为只含基频的波型；第二个为含有两个分音(基频与第二分音)的波型；第三个为含有三个分音(基频、第二分音和第三分音)的波型，依次类推所示为分别含有四个及五个分音的合成波型。

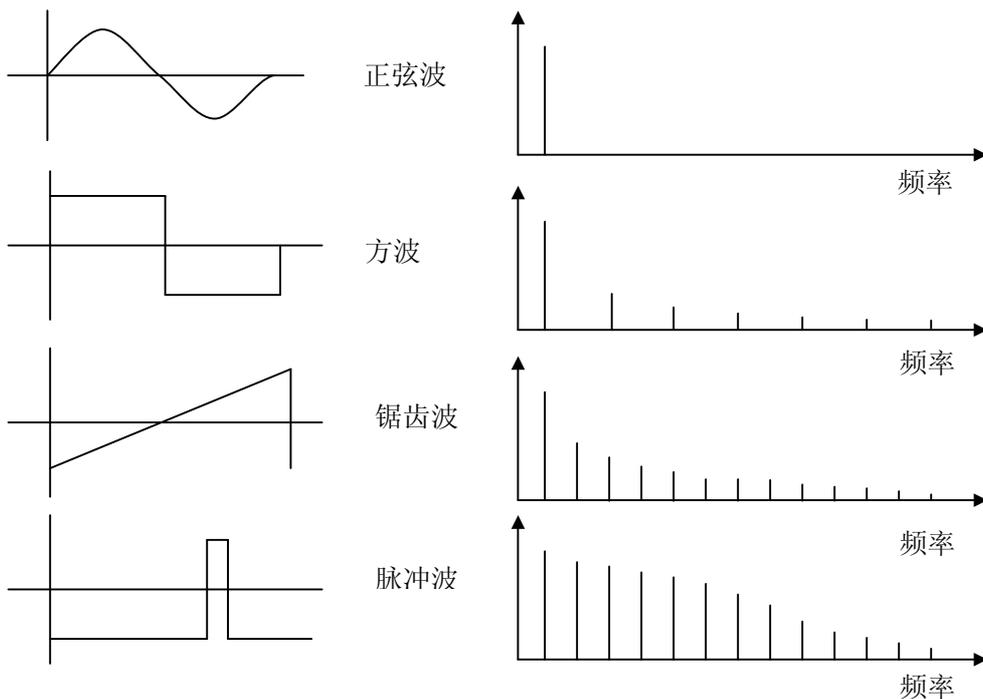


图 4.1.12 简单波型声谱示意图。

以图形方式表示出的某声音的声学能量在基频与各分音之间的分配模式即为该声音的声谱。在声谱图中，横坐标表示分音序号，纵坐标代表振幅。见图 4.1.12：简单波型的声谱示意图。

图中，左边为简单波型的图示，右边是它们的声谱示意图。正弦波只含一个分音，其全部声学能量集中在基频上；方波的声谱显示，其只含奇数分音，且分音的振幅是基频振幅的奇数分之一，如第三分音的振幅为基频的三分之一，第七分音的振幅为基频的七分之一等等；锯齿波几乎在所有分音中都含有声学能量，其分配形式与泛音号呈指数递减；脉冲波的声谱表明其在各频段广泛的能量分布。

将声谱图中表示分音振幅的竖线顶端相连即可得到该声谱的包络图示，见图 4.1.13：声谱包络示意图。图中虚线所表示的即为该声音的声谱包络曲线。

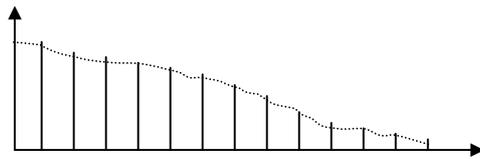


图 4.1.13 声谱包络示意图

通过声谱包络，我们就可以定量地表示某声音的声学能量分配形式。如白噪声的声谱包络如图 4.1.14 所示：

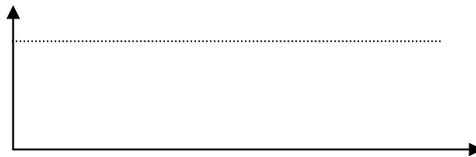


图 4.1.14 白噪声的声谱包络示意图

由该图中平直的声谱包络线我们得知，白噪声在各频段的声学能量分布都很大。通过对不同声音的声谱分析，从声音的声谱包络即声音能量的分配模式中计算机音乐家们就可以掌握声音的某些奥秘，比如我们已经知道铜管乐器嘹亮的音色是因为其声学能量在高频区的广泛分布造成的。掌握声音的能量分布后，就可以利用计算机及电子乐器进行声音的合成与再合成。同时音乐家就可以通过实时地改变声音声谱包络模式的方法从而实现音色的实时变化。

4. 共振频率(Formant Frequency)

共振峰(共振频率)是决定声音音色的另一个重要维度。人声与大多数声学乐器都具有共振峰。人的口腔、声学乐器的共鸣腔等由于材质、体积等的不同致使激发的声音无论其基频是多少，其声谱中声学能量都相对集中在某个频段，该频率被称为该发声体的共振频率。因此在电子音乐领域有专门或大量应用共振峰进行声音合成的技术，如 FOF 计算机直接数码合成语言、FS 音源、VP 音源等。

目前对音色的现代研究中，科学家们还在不断地探索着声音的奥秘。随着人们对声音不同侧面的发现，不断有新的声音合成理论的诞生，相继就有新的声音合成工具供人们研究、使用。相信有一天，音乐家们可以拥有简单易用、功能强大的声音制作工具供音乐家们在电子音乐世界中任意驰骋。

至此，我们了解了声音的物理本质，以及部分有关音色的基础理论。由于本书的主要目的在于学习 MIDI，因此只介绍有关声学的基本常识，有兴趣的读者不妨从其它专业书籍中获得更深、更广的相关知

识。

三. 声音的传播特性

上一个单元中，我们拿着“手术刀”，通过“显微镜”将声音解剖开来进行了研究。现在，让我们将声音作为一个整体，看一看它在空气中是怎样传播的。

1. 声音传播的衰减

声音在空气中传播的速度即声速大约为每秒钟 1130 英尺。声音在空气中的传播呈衰减的特性，在没有任何障碍物的室外空间中，当我们离开一个室外声源时，起初会觉得声压下降得很快，之后就觉得声压降低的速度变慢了；每当我们距离声源的距离加倍时，声压级便下降大约 6 分贝，这种关系被称为倒数平方规律。

2. 反射、吸收、折射、绕射

声音在遇到障碍物或空气变化时，会产生反射、吸收、折射和绕射等现象。

声音在遇到坚硬平滑的障碍物如玻璃时，会产生反射现象。反射角度以及反射量随障碍物材质的不同而略有差异。

与反射相反，当声音遇到吸音材质的障碍物如吸音石棉网时，只有很少的声音被反射出来，大部分的声学能量被“吸”走了——转化成热能等其它形式的能量。吸收量的大小也随着材质的不同而不同。

在室外，声音经过温度不同的区域时会产生折射或改变方向。另外风速对声音的传播也有影响，尤其是强风。

声音在遇到某些障碍物时会产生绕射。如声音可以绕过墙角传到另一侧，再如声音可以通过墙上的小孔传到另一侧。通常低频(波长较长)比高频(波长较短)的绕射现象更为显著。

3. 自然声场

前文中已经叙述了当声音在空旷的室外传播时遵循倒数平方规律。当在室外环境中遇到较大的反射物时，经过其反射的声音也同样会到达听者的位置，但时间要晚一些。经反射延迟到达的声音被称做延迟声或回声。如果反射物不止一个，而且它们的位置各不相同的话，听者有可能听到多次的反射声。例如山谷回声。

当声音在室内传播时，情况要复杂的多。听者最先听到的应该是直接由声源直线距离传来的声音，即直达声。由于环境四周的反射，听者其次听到的是从墙壁反射回来的声音，即非直达声。因为反射材质对声音能量的反射大于其吸收的能量，故经反射后的声音在损失了部分能量后仍然可以被继续反射。所以在非直达声中，又有只反射过一次的单次反射声与多次反射后的多次反射声之分。

直达声、单次反射声与多次反射声一起，构成了声音的残响，即所谓的混响声。一般来说，面积越大的房间其混响时间越长。如我们在教堂、大厅、楼道中说话时就比在宿舍中的混响大。

声音的延迟与混响是声音在自然声场中表现出的主要特性。

至此，我们从声音的物理本质到它在自然声场中的特性等方面已经对传统声音世界有了一定的了解，下一节我们将进入电声的世界。