“音色”背后的奥秘

2024.12

音乐是人类最伟大的艺术发明之一，现在让我们来探究一下其中蕴含的物理奥秘吧!

【一】描述音的基本物理量

我们一开始接触初中物理时就了解到，声音是由于物体振动产生的，这个振动产生声音的物体我们叫它声源。声源振动会带动其周围的介质振动，这种振动形式在空间中传播开来，传到我们的耳朵里，引起鼓膜振动，进一步产生听觉。

我们用如下名词去描述一种声音，或者说一种声波。

首先是声音的强弱，或者说它的大小。对应于一种物理量的话，就是振幅（amplitude），用 A 来表示。振幅即是说传递声波的介质振动的幅度。振幅越大，声音越响。

还有就是声音的高低——声音音调的高低，对应的物理量是频率（frequency），用 f 来表示。频率的单位是 Hz，某振动的频率是 100Hz，即是说传递这种振动的介质物质一秒钟完成 100 个全振动。需要补充的是，在本文后面的篇幅中经常会出现所谓角频率的概念，它的符号是ω。角频率有什么物理意义呢？是这样的，每一种振动都可以与一种匀速圆周运动产生关联，角频率就是这匀速圆周运动的角速度。角频率与频率的关系是:

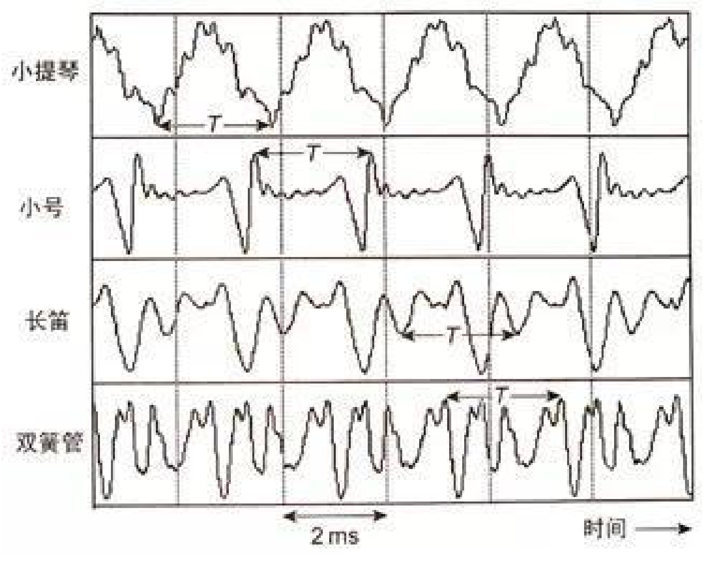
$$\omega = 2\pi f$$

跟频率对应的还有周期（period），用 T 表示。即完成一次全振动所需要的时间长短，周期和频率的大小互为倒数。

接着就是相位（phase），关于这个概念感兴趣的可以参看任何一本物理教材，在此就不再赘述了。

但是我们在初中就学习过，声音有三大特性——响度、频率、音色。前面两种特性就是上文中介绍的，对应于振幅、(角）频率这两个物理量。

那么该如何理解音色呢？一种朴素的理解方式是——我们正是凭借着音色不同来辨别这声音是什么东西或者什么人发出的。比如同样是一首曲子，我们轻松地就可以辨别出钢琴弹奏和提琴拉奏。更深入一点的话，我们可以从振动的波形来理解。我们用四种不同的乐器演奏一个乐音，比如说 do，那么这时候乐器附近的空间中的空气分子会发生振动。我们取一点，记为 P 点，来考察这一 P 点介质分子（这里是空气）的振动模式，就会得到左图一样的波形图。这里横轴是时间轴，纵轴代表了介质分子偏离了平衡位置的



位移。

仔细观察这一幅图片，不同乐器导致介质的振动模式看似大相径庭，其实是既有相似，也有不同的。这相似之处在于这些振动的频率是一致的，或者说周期是相同的——不同乐器引起 P 点的振动，在相同的时间内完成的全振动次数相同。这频率（或周期）一定等于声源振动的频率（或周期）。而这四种振动方式的不同点就很显然了，它们的波形是很不一样的。

我们说正是波形的特异性，决定了声音音色的特异性。

【二】周期函数的傅里叶级数展开

我们这篇文章就是要弄明白音色背后的秘密。上面用比较定性和朴素的方式对音色的物理图像有了一个基本了解。为了更深刻地探究其奥秘，我们需要一些数学上的工具和语言的预备。下面我将努力用比较易懂的语言完成这件事。

我们在高中学习三角函数的时候大抵有所耳闻——三角函数是“很基本”的函数，因为任何一个波形图（比如方波、三角波、脉冲波，甚至是上面哪些乐器演奏出的声波波形）都可以写成（数学上的术语叫做“展开成”）有限个或无限个正弦波、余弦波加在一起的形式。

举个例子吧，比如说上面那张图片里，用小提琴拉出的波形，放在平面直角坐标系中，它的函数为 $y = f(t)$，假设这个波形的最小周

期为 T。那么傅里叶级数展开理论宣称，在横轴任何一个长度为 T 的区间内（下面我们不妨取一个特例，即取[-T/2,T/2]作为展开区间），我们可以用有限或无限多个（一般情况下为无限多）正弦函数和余弦函数的叠加来表示函数 $y=f(t)$。写成等式，即为：

$$\begin{aligned} f(t) &= a_0 + a_1\cos\omega t + b_1\sin\omega t + a_2\cos 2\omega t + b_2\sin 2\omega t + \ldots \\ &\quad + a_n\cos n\omega t + b_n\sin n\omega t + \ldots \\ &= a_0 + \sum_{n=1}^{+\infty} a_n\cos n\omega t + b_n\sin n\omega t \end{aligned}$$

$$t\in[-\frac{T}{2},\frac{T}{2}]$$

可以看到，展开之后这些若干个正弦、余弦函数的角频率是不一样的，而且均为某一个确定角频率 ω 的整数倍。这个 ω 就是小提琴弦（声源）的振动角频率，其大小为 $\omega=\frac{2\pi}{T}$。此外，这一展开式里面还有许多待定系数 a_n、b_n（n=1，2，3…），它们等于多少呢？可以证明，只要左边的 $y=f(t)$ 表达式确定下来了，这些系数就都统统确定下来了。具体有：

$$a_0=\frac{1}{T}\int_{-T/2}^{T/2} f(t)\mathrm{d}t \quad ; \quad a_n=\frac{2}{T}\int_{-T/2}^{T/2} f(x)\cos n\omega t\mathrm{d}t \quad ; \quad b_n=\frac{2}{T}\int_{-T/2}^{T/2} f(x)\sin n\omega t\mathrm{d}t$$

注意，上面这些不过是一些积分式而已——它们运算的结果都是（和 n 有关的）常数。

对于上面的展开式，我们叫它为 $y=f(t)$ 的**傅里叶级数展开式**.我们不妨继续用辅助角公式把它的右边化成更加简洁的形式：

$$f(x)=a_0+\sum_{n=1}^{+\infty} A_n\cos(n\omega t+\varphi_n)$$

不难理解，波形 $y=f(t)$ 一旦确定，上述公式中的 A_n 、 φ_n 亦就随之确定了。

【三】复合音与分音列

我们平时听到的声音，一般都是由许多个声音复合而成的。这种声音，叫做**复合音**。有了第二节中的数学知识，就不难理解这一命题了——因为**任何**一种波形的声音波都可以看成是由 $A_n\cos(n\omega t+\varphi_n)$ 这些有限个或无限个的不同频的“基础的”声音波（简谐波）复合而成。这就导致了乐理中分音列的概念。

任何一种音波，一定可以把它分解成有限个或无限个不同振幅、频率（和初相位）的基础的简谐波的叠加，这一系列简谐波对应的音一起构成了这一声音的**分音列**。其中各个简谐波的角频率为声源角频率 ω 的整数 n 倍。在分音列中，角频率等于声源角频率 1 倍的（即就等于声源角频率的）那个声音，叫做**基础音**（基音），其余的叫做**泛音**（倍音）。

可以把分音列视为一个集合，基础音和泛音是分音列这一集合中的元素。

在实际中，我们认定某一声音的频率往往是其分音列中基础音的频率，尽管这一分音列中有众多更高频率的泛音。

【四】、弦乐器的音色

好的，让我们回到文章的主题——音色。是什么决定了音色呢？

让我们以弦乐器，比如钢琴、提琴、吉他、琵琶、二胡、古筝等等，为例，来考虑一下振动的源头——声源的振动。

弦乐器靠弦振动发声。现在想象——有一根密度ρ和粗细都均匀的长为l一维理想弦，作小振幅的高频振动，这一模型设定是比较符合实际的。把这根弦放在 x 轴上，原点对齐。引入一个函数$u(x,t)$，叫做这根弦振动的**波函数**，它的意义如下：认为弦由无数质点连成，那么在 t 时刻，坐标为 x 处的质点偏离平衡位置（这里是轴）的位移为$u(x,t)$。波函数 u 满足的偏微分方程为：

$$\frac{\partial^2 u}{\partial t^2}-\frac{T}{\rho}\frac{\partial^2 u}{\partial x^2}=0$$

其中 T 为振动过程中弦的张力大小，可以证明它是一个常数（即是说 T 不随着位置和时间而改变）。关于这一方程的导出方法（以及它的求解方法）比较繁琐，感兴趣的读者可以去翻看任何一本《数学物理方法》教材，这里我们先把它们接受下来。

在实际的乐器上，弦都是两端固定不动的，翻译成数学的语言，即是说对于波函数$u(x,t)$，它满足$u(0,t)=u(l,t)=0$这一**边界条件**。实际演奏过程中弦的波函数一定是在这一条件限制下的解。可以证明，这解形式上是：

$$u(x,t)=\sum_{n=1}^{+\infty}(a_n\cos\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t+b_n\sin\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t)\sin\frac{n\pi}{l}x$$

是一个无穷级数，其中a_n、b_n为待定系数，请注意将其与前面提到的a_n、b_n区分开来，前面是介质的振动，这里研究声源——弦的

振动。

如何确定这里的 a_n 、b_n？这要从弦振动的**初始条件**中寻找答案。所谓初始条件，就是指下面这两个函数：

(i) 在 t=0 的时刻（计时起点时刻），这根弦长什么样子？——即 $u(x,0)$ 这个函数的具体形式是什么？

(ii) 在 t=0 的时刻，它各个点振动的速度是如何分布的？——即 $\frac{\partial u}{\partial t}(x,0)$ 这个函数的具体形式是什么？

这两个初始条件就足以确定 a_n 、b_n，进而确定这弦的具体振动模式了。为什么需要有两个初始条件？因为描述弦振动的偏微分方程是二阶的。

如果我们已知了上面这两个初始条件，设它们是：

$$\begin{cases} u(x,0) = \alpha(x) \\ \dfrac{\partial u}{\partial t}(x,0) = \beta(x) \end{cases}$$

其中 $\alpha(x)$ 、$\beta(x)$ 为两个完全已知的函数（它们满足和波函数一样的边界条件）。那么就可以得到：

$$\begin{cases} a_n = \dfrac{2}{l}\displaystyle\int_0^l \alpha(x)\sin\dfrac{n\pi}{l}x\mathrm{d}x \\ b_n = \dfrac{2}{n\pi}\sqrt{\dfrac{\rho}{T}}\displaystyle\int_0^l \beta(x)\sin\dfrac{n\pi}{l}x\mathrm{d}x \end{cases}$$

这不过是两个积分式而已，它们运算的结果都是和 n 有关的常数。

好的，我们再仔细观察一下：

$$u(x,t) = \sum_{n=1}^{+\infty}\left(a_n\cos\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t + b_n\sin\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t\right)\sin\frac{n\pi}{l}x$$

这一方程，用辅助角公式将其简化，得：

$$
\begin{aligned}
u(x,t) &= \sum_{n=1}^{+\infty}(a_n \cos\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t + b_n \sin\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t)\sin\frac{n\pi}{l}x \\
&= \sum_{n=1}^{+\infty} A_n \cos(\sqrt{\frac{T}{\rho}}\frac{n\pi}{l}t + \varphi_n)\sin\frac{n\pi}{l}x \\
&= A_1 \cos(\sqrt{\frac{T}{\rho}}\frac{\pi}{l}t + \varphi_1)\sin\frac{\pi}{l}x + A_2 \cos(\sqrt{\frac{T}{\rho}}\frac{2\pi}{l}t + \varphi_2)\sin\frac{2\pi}{l}x + ...
\end{aligned}
$$

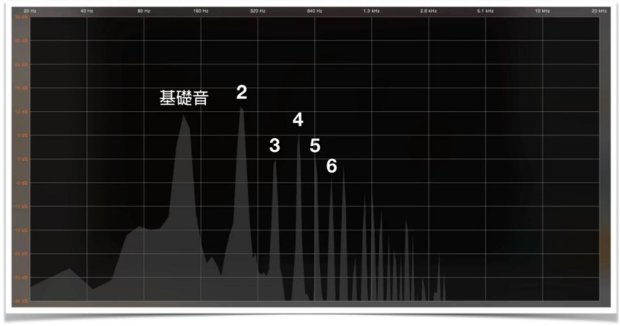
可见弦的振动波函数实际上是无穷多个**驻波形式**的波函数之叠加。这些驻波的频率都为某一个基础频率 $\omega = \sqrt{\frac{T}{\rho}}\frac{\pi}{l}$ 的整数 n 倍（n=1，2，3…）。我们把这些驻波按照 n 进行编号：1，2，3…。

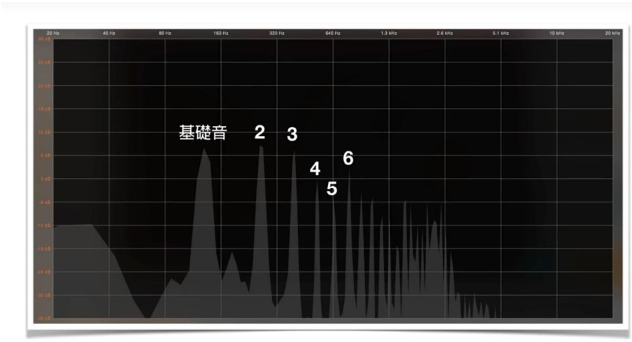
还记得我们前面曾提到的复合音的分音列吗？我们听到的声音都可以看成是由其分音列中的元素，即基础音和泛音，叠加而成的。**可以认为，分音列中的基础音和泛音就是分别由弦上的 1 号、2 号、3 号…驻波振动发出的。**这些驻波合成为弦的实际振动——相应地，基础音和泛音合成为我们听到的复合音。

从波函数的 a_n、b_n 的表达式中可以看出它们是和许多因素有关系的——弦密度、弦张力、初始条件，其中前两者与乐器的构造有关，后面的初始条件则取决于乐器的发音方式和演奏者操作的方式。

a_n、b_n 决定了弦振动波函数展开式中各个驻波的振幅和相位，而各个驻波又对应产生了分音列中的各个元素。正是分音列中各个元素（的波形）振幅和相位的特异性，导致了不同声音音色的特异性。所以我们可以得出结论：不同的弦乐器之所以有特异的音色，是因为其特异的构造和不同的演奏方式。

最后举个实例帮助大家理解——上下两幅图片依次是钢琴和大提琴弹奏同一个音 do 时声波的频谱图，横轴是频率。





图中有一个个“峰”，还有编号。编号代表着它是分音列中的第几个元素，峰的高低反应了这一元素的振幅的大小。可以看到，不同乐器虽然各个峰在横轴上的位置相同即对应频率相同 (这是当然的，因为弹奏的是同一个音 do)，但是它们的振幅比例并不相同，因此音色不同。而这两个 do 的频谱不同是由于产生它们的弦的构造以及振动模式不同。

参考文献及资料来源：
[1]梁昆淼.数学物理方法（第五版）.北京.高等教育出版社.2020.
[2]李重光.基本乐理通用教材. 北京.高等教育出版社.2004.
[3]@也许会弹琴的长颈鹿.《一次搞懂「泛音列」!（搬运自 NiceChord）需要多看几遍，但是很好理解！》.Bilibili 网站.2017-8-13.
[4]@柴知道.《科普：“音色”的本质是什么？为什么二胡拉欢乐颂都那么丧》.Bilibili 网站.2017-5-12.