Python 与声音制造

第一部分: 声音与数学

Topic 1:声音信号与三角函数

数学最有用的地方就在于它可以量化地描述真实世界中发生的事情(尽管只能描述很小一部分)，如果能掌握这些表述方法，那么我们就具有了某种改造世界的能力，甚至合成出新的事物。声音就是这样一个例子，它源于物体的振动，由振动源产生的能量通过空气传播至人耳， 最后通过人的神经系统转化成我们感受到的所谓声音。

如果一个声音信号波动具有明显的周期性，那么我们就感觉到它有一个明显的音高(pitch)，例如一个具有固定频率的正弦波。弦乐乐器发出的声音都是周期信号，如果不是这样，那我们听到将会是噪音，当然这里的前提是演奏乐器的人具有基本的演奏能力。但是打击乐器，例如鼓或者镲，就不能演奏出明显的音符，这主要是因为它们的物理形状是平面的，所以很难形成占绝对优势的振动频率。

频率，或者周期是描述声音信号最基本的基本参数。频率(Frequency)描述振动的快慢，以振动次数除以时间为单位(赫兹)，周期(Period)则是频率的倒数，人类的听觉范围是 20Hz-20000Hz，频率高的声音往往给人以尖锐紧张的感觉。振幅(Amplitude)描述了一个波动相对中间位置可偏移的最远距离，体现在声音上就是音量。最后就是振动在时间维度上的图像，大量声学研究表明我们人类对于声音的感受与声音信号的波形有关 。在正式开始有关数学描述之前，请思考这么几个问题:

1. 如果钢琴和小提琴演奏同一份乐谱，你是否可以分辨出它们的不同？
2. 如果你爸爸和你妈妈先后喊你的名字，你是否可以分清他们的声音?

一个著名的数学定理阐述了这样一个现象，任何周期函数都可以由有限或无限多个正弦函数经过线性叠加而成. 回忆正弦型函数具有以下基本形式:

$$y = A\sin(\omega t + \varphi)$$

其中的 A，ω，φ为常数，它们决定这个正弦波的全部性质，即振幅(Amplitude), 频率(Frequency)和初相位(Phase)。

所以若一个波满足以下条件:

$$f(t) = f(t + T)$$

T 为这个波的最大周期，则有

$$f(t) = \sum_{k=0}^{\infty} A_k \sin(k\omega t + \varphi_k)$$

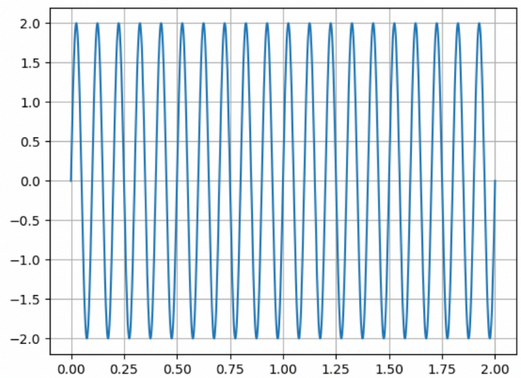
其中 k 为非负整数，这也就是说任何周期函数都是由一个最低频率的正弦波和其他频率为其整数倍的正弦波叠加而成的。每一个正弦波都有各自的振幅，频率和初相位而不受其他正弦波的影响。我们可以把这个系统想象成一个维度为有限(k 取有限个)或者无限的空间，每一个正弦波都代表了这个空间的基向量，就像平面直角坐标系中的 x,y 轴。这样我们就可以把任何周期波动看成是在这个由正弦波构成的空间中的一个向量。 这样就可以利用向量的数量积对波形进行分解，得到每一个基波的振幅，频率和初相等信息。这样做有什么意义呢？这就意味着，我们得到这些信息之后就可以重构原始波形。事实上所有软件音源都是基于这一工作原理。

下面我们使用一个简单的 Python 程序来呈现这个结果:

首先生成一个正弦波的图像

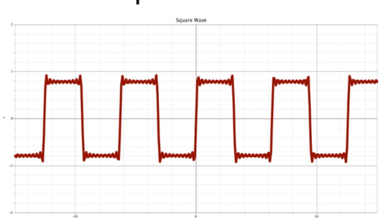
```python
import numpy as np
import matplotlib.pyplot as plt

#用户输入信号的特征信息
T=int(input("请输入时间间隔:"))
f=int(input("请输入正弦函数的频率:"))
a=float(input("请输入振幅:"))
Fs=44100# 采样频率
t=np.linspace(0,T,Fs*T)#产生时间序列
y=a*np.sin(2*f*np.pi*t)#产生正弦波
plt.plot(t,y)#绘制图像
plt.grid()#给图像增加网格线
plt.show()#显示图像
```

```
>>>
 RESTART: D:/onedrive_北大附中
记/课件/第一章代码/simple_sine
请输入时间间隔:2
请输入正弦函数的频率:10
请输入振幅:2
```



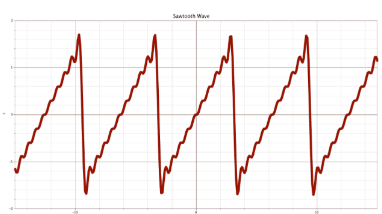
由多个频率不同的正弦函数按照一定的规则进行叠加可以产生一些有趣的基本波形。

square wave



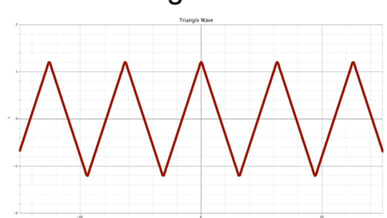
$$y=\sum \frac{1}{2n-1}\sin(2n-1)x$$

sawtooth wave



$$y=\sum \frac{(-1)^{n+1}}{n}\sin nx$$

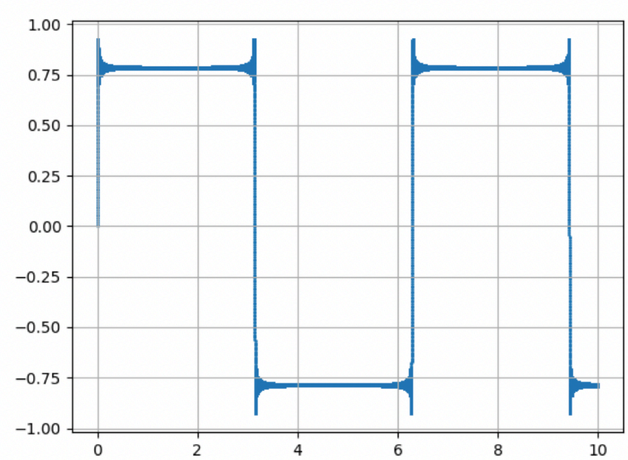
triangle wave



$$y=\sum \frac{1}{(2n-1)^2}\cos(2n-1)x$$

以矩形波为例:

```python
import numpy as np
import matplotlib.pyplot as plt
Fs=10000 #采样频率
n=100 #级数的项数
T=10 #声音信号的时长
t=np.linspace(0,T,Fs*T) #时间节点序列
#sqaure wave
y=np.zeros(len(t)) #初始化信号序列
for i in range(1,n):
    y=y+1/(2*i-1)*np.sin((2*i-1)*t)
plt.scatter(t,y,s=1)
plt.grid()
plt.show()
```



对于上面的公式，我们还可以利用三角函数和角公式对其展开:

$$\mathrm{f(t)} = \sum_{k=0}^{\infty} a_k \cos(k\omega t) + b_k \sin(k\omega t)$$

这是因为:

$$\begin{aligned} \mathrm{A}\sin(\mathrm{k\omega t} + \varphi) &= \mathrm{A}(\sin(\mathrm{k\omega t})\cos(\varphi) + \cos(\mathrm{k\omega t})\sin(\varphi)) \\ &= \mathrm{A}\sin(\varphi)\cos(\mathrm{k\omega t}) + A\cos(\varphi)\sin(\mathrm{k\omega t}) \\ &= \mathrm{a}\cos(\mathrm{k\omega t}) + b\sin(\mathrm{k\omega t}) \end{aligned}$$

其中：$\mathrm{a} = \mathrm{A}\sin(\varphi), b = A\cos(\varphi)$

这样就有以下关系式:

$$\mathrm{A} = \sqrt{a^2 + b^2}$$

$$\tan(\varphi) = \frac{\mathrm{a}}{\mathrm{b}} = \frac{\mathrm{A}\sin(\varphi)}{A\cos(\varphi)}, \varphi = \tan^{-1}\left(\frac{a}{b}\right)$$

接下来将要介绍被誉为最美数学公式的欧拉恒等式，首先存在这样的关系

$$\mathrm{e^{ix}} = \cos(x) + isin(x)$$

其中 i 为单位虚数，$\mathrm{i} = \sqrt{-1}$。

这是根据泰勒展开式:

$$\sin(\mathrm{x}) = \mathrm{x} - \frac{\mathrm{x}^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \frac{(-1)^{n-1}x^{2n-1}}{(2n-1)!} + o(x^{2n+1})$$

$$\cos(\mathrm{x}) = 1 - \frac{\mathrm{x}^2}{2!} + \frac{x^4}{4!} + \cdots \frac{(-1)^n x^{2n}}{2n!} + o(x^{2n+2})$$

$$\mathrm{e^x} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \frac{x^n}{n!} + o(x^{n+1})$$

可以把以上公式中的最后一项认为是跟前 n 项比可以忽略不计的量。

这个公式将三角函数和指数函数联系起来，下面来看一下它的威力:

$$\mathrm{e^{i2x}} = \cos(2x) + isin(2x) = \left(\cos(x) + isin(x)\right)^2 = \cos^2 x - \sin^2 x + i2sin(x)\cos(x)$$

一个由实部和虚部构成的数叫做复数, 基本形式是$\mathrm{a} + \mathrm{bi}$。两个复数相等的充要条件是它们的实部和虚部分别相等，即得到倍角公式:

$$\cos(2\mathrm{x}) = \cos^2 x - \sin^2 x$$

$$\sin(2\mathrm{x}) = 2\sin(\mathrm{x})\cos(\mathrm{x})$$

由此得到一般性结果

$$\cos(\mathrm{nx}) + \mathrm{i}\sin(\mathrm{nx}) = \left(\cos(\mathrm{x}) + \mathrm{i}\sin(\mathrm{x})\right)^{\mathrm{n}}$$

下面我们将π带入欧拉等式,

$$\mathrm{e^{i\pi}} = \cos(\pi) + isin(\pi) = -1 + 0$$

即得到被称为最美数学公式的，它将数学中最常见的常数和运算都浓缩进一个公式里。

$$\mathrm{e^{i\pi}} + 1 = 0$$

有了这一层关系，我们就可以将前面的周期信号以复指数的形式表现出来，这里我们将波动位移看作关于时间的函数：

$$\mathrm{x(t)} = \mathrm{a} * \cos(\omega \mathrm{t}) + \mathrm{b} * \sin(\omega \mathrm{t})$$

$$\mathrm{x(t)} = \sqrt{a^2 + b^2}\left(\frac{a}{\sqrt{a^2 + b^2}}\cos(\omega t) + \frac{b}{\sqrt{a^2 + b^2}}\sin(\omega t)\right)$$

$$x(t) = A * \cos(\omega t + \varphi)$$

其中:

$$A = \sqrt{a^2 + b^2}, \varphi = -\tan^{-1}\left(\frac{b}{a}\right)$$

设$I = Ae^{i\varphi} = Acos(\varphi) + iAsin(\varphi)$, 那么原始信号就可表示成

$$x(t) = Re(Ie^{i\omega t}) = Re(Ae^{i\varphi} * e^{i\omega t}) = Re(Ae^{i(\omega t+\varphi)})$$

即这个信号为复指数函数$Ae^{i(\omega t+\varphi)}$)的实数部分。

那么啰嗦了这么多到底有什么用？第一，大家可以看到在高中学到的有关数学知识经过整合可以用来表现和分析一个具体的事物，所以证明学好数学是有用的！其次，以上的数学语言是我们接下来利用计算机制造声音的基本理论工具，所以如果你希望将你的电脑变成一个声音制造工厂，那么你最好牢牢记住以上概念。

在下次的推送中我们将介绍有关傅里叶级数的内容，从而掌握如何将一个比较复杂的波形进行解构和重塑。

本文的内容编排主要参考以下资源:

F. R. Moore, An Introduction to the Mathematics of Digital Signal Processing: Part I: Algebra, Trigonometry, and the Most Beautiful Formula in Mathematics Author(s): F. R. Moore, Computer Music Journal, Vol. 2, No. 1 (Jul., 1978), pp. 38-47