

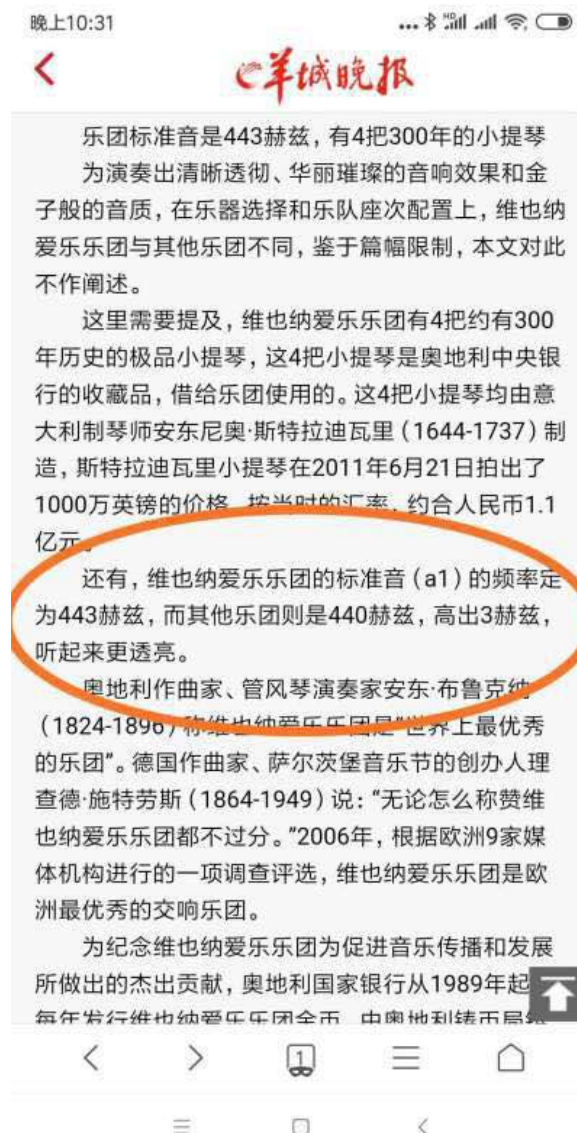
2019 春信号与系统大作业之“用科学支撑艺术”

谷源涛

2019 年 6 月 27 日

0 孟教授之问

2019 年 5 月 20 日，东南大学孟桥教授在“信号与系统教学群”中抛出一个问题：
“有老师能设法验证这个说法是真还是假？(墨镜脸)”



1 基础

1.1 乐音和频率 [2]

1.1.1 乐音基波的构成规律

我们用大写英文字母 $CDEFGAB$ 表示每个音的“音名”(或称为“音调”),当指定某一音名时,它对应固定的基波信号频率。图 1 示出钢琴键盘结构,并注明了每个琴键对应的音名和基波频率值。这些频率值是按“十二平均律”计算导出,下面解释计算规则。

		E	659.25 Hz	
bE		D	622.25 Hz	
bD		C	587.33 Hz	
		B	554.36 Hz	
bB		A	523.25 Hz	
bA		G	493.88 Hz	
bG		F	466.16 Hz	
		E	440 Hz	(f_{A1})
bE		D	415.30 Hz	
bD		C	392 Hz	
		B	369.99 Hz	
bB		A	349.23 Hz	
bA		G	329.63 Hz	
bG		F	311.13 Hz	
		E	293.66 Hz	
bE		D	277.18 Hz	
bD		C	261.63 Hz	(中央C, f_{C1})
		B	246.94 Hz	
bB		A	233.08 Hz	
bA		G	220 Hz	(f_{A0})
bG		F	207.65 Hz	
		E	196 Hz	
bE		D	184.99 Hz	
bD		C	174.61 Hz	

图 1: 钢琴键盘和相应频率

靠下边的 A 键称为 $A3$ (也称为小字组 A), 它的频率值 220Hz。而靠上面的另一个 A 键是 $A4$ (也称为小字一组 A), 它的频率值是 440Hz。两者为二倍频率关系, 也称为 8 度音程或倍频程 Octave (即我们画频响特性波特图时所用的术语“倍频程”。

从 $A3$ 到 $A4$ 共有 12 个键, 其中 7 个白色键, 5 个黑色键, 其频率值计算规律为相邻音倍乘系数 $K = 2^{\frac{1}{12}} = 1.05946309$ 。由此可求出图中各琴键对应之频率值。例如从 $A3$ 导出 $C4$ (也称中央 C) 的频率为

$$220 \times 2^{\frac{3}{12}} \text{Hz} = 261.63 \text{Hz}. \quad (1)$$

从图 1 可以看出 7 个白键之间插入了 5 个黑键。在 EF 之间和 BC 之间没有黑键, 也即这两组相邻的白键之间基波频率倍乘系数为 $2^{\frac{1}{12}}$, 也称为相隔半音, 而在其他白键之间都有黑键相隔, 因而他们的频率倍乘系数为 $2^{\frac{2}{12}}$, 也称为相隔全音 (如 CD 、 DE , FG 、.....之间)。若以白键英文字母为基准, 则升高半音以“ \sharp ”符号表示, 降低半音则以“ b ”符号表示。于是, 可以依次写出 12 个音名从低到高的字母表示为

$$C, ^bD, D, ^bE, E, F, ^bG, G, ^bA, A, ^bB, B$$

当然, 若改用“ \sharp ”号表示黑键, 则 bD 改为 $\sharp C$, bE 改为 $\sharp D$,。

1.1.2 乐音谐波的作用—音色

当指定音名（音调）之后仅指定了乐音信号的基波频率，谐波情况并未说明。对于各种乐器如钢琴或小提琴都可发出 440Hz 之乐音，而人的听觉会明显感觉二者不同，这是由于谐波成分有所区别，频谱结构各异。例如小提琴的三次、五次谐波成分很强，其它各种乐器都有自己的谐波分布规律。同种乐器不同音阶之谐波构成还可能略有区别。由于演奏技巧、方法之差异也可产生不同结构之谐波。在音乐领域中称谐波为“泛音”。谐波的作用是使音色发生变化。

1.2 信号处理方法

1.2.1 时域加窗

我们感兴趣的信号往往无限长，但受资源所限，我们只能截取它的一段（即用无限长信号乘以矩形窗）进行处理。在时域（无限长信号）和矩形窗相乘，在频域原（无限长信号的）频谱则和矩形窗的谱（Sa）卷积。Sa 的旁瓣很高，意味着把（无限长信号的）谱向两侧做（严重）混叠，所以直接计算有限长信号的谱并不能很好的估计原（无限长）信号的谱。

如果用某些非矩形的窗截取（在频域和能量更集中的谱卷积）则能极大地改善频谱估计的质量，这就是时域加窗。常见的窗函数包括升余弦窗、hanning 窗、blackman 窗、高斯窗等。

1.2.2 短时傅里叶变换

短时傅里叶变换（Short-time Fourier Transform, STFT）就是对信号加窗后再做傅里叶变换，而窗的位置可以变化，从而度量不同时间点上的能量在频域的分布，或者理解为不同频率点上的能量在时间上的分布，再或者，信号能量在时频平面上的分布。

$$\text{STFT}_w\{f(t)\} = F(\tau, \omega) = \int_{-\infty}^{\infty} f(t)w(t - \tau)e^{-j\omega t}dt,$$

其中 $w(t)$ 是一个以 0 为中心的时间窗，如果选用宽为 T 的矩形窗的话，则有

$$w(t) = \begin{cases} 1/T, & |t| < T/2; \\ 0, & \text{elsewhere.} \end{cases}$$

以 τ 为横坐标，以 ω 为纵坐标，我们就可以在“时频平面”上研究信号，即对信号进行时频分析。

1.3 用时频分析研究小提琴奏鸣曲

请大家打开手机上的音乐软件，搜索“降 E 调小提琴奏鸣曲第一乐章”，你将听到……

如果把刚才听到的音乐画出来，你将看到图 2 所示的样子。

看不清楚的话可以放大，如图 3 所示，有明显的周期性。这个周期就对应于基波频率，而一个周期波形里的“内容”就表示了谐波成分。

下面用时频分析法研究这段音乐。请看图 4。横坐标是时间，纵坐标是频率，平面上的颜色表示了“信号在这个时刻含有这个频率的强度”，越暖的颜色表示强度越大。

我们从左往右，把每个时刻的频率最强的地方描红，大约可以得到图示四个线段。它们的长度表示了四个音符的持续时间，纵坐标表示了“音调”，即“降 B3，D4，降 E4，D4”。

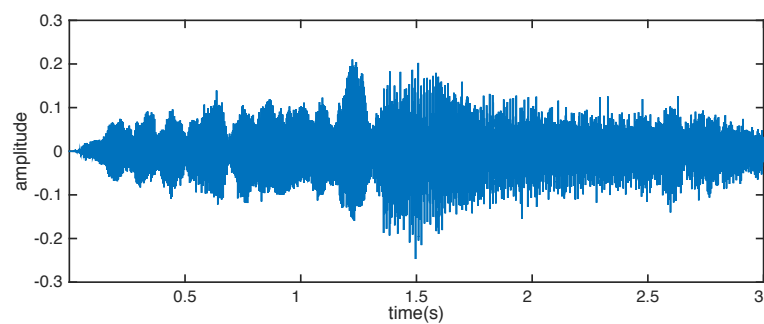


图 2: 小提琴奏鸣曲波形

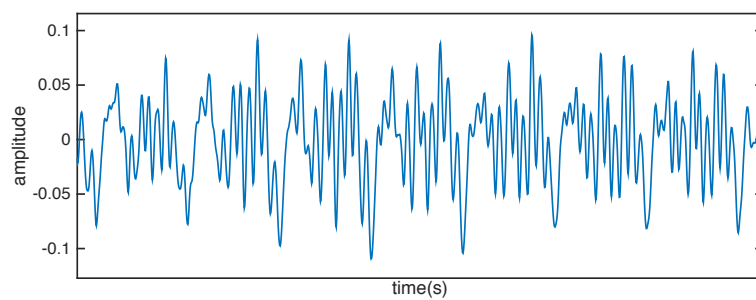


图 3: 小提琴奏鸣曲波形 (局部放大)

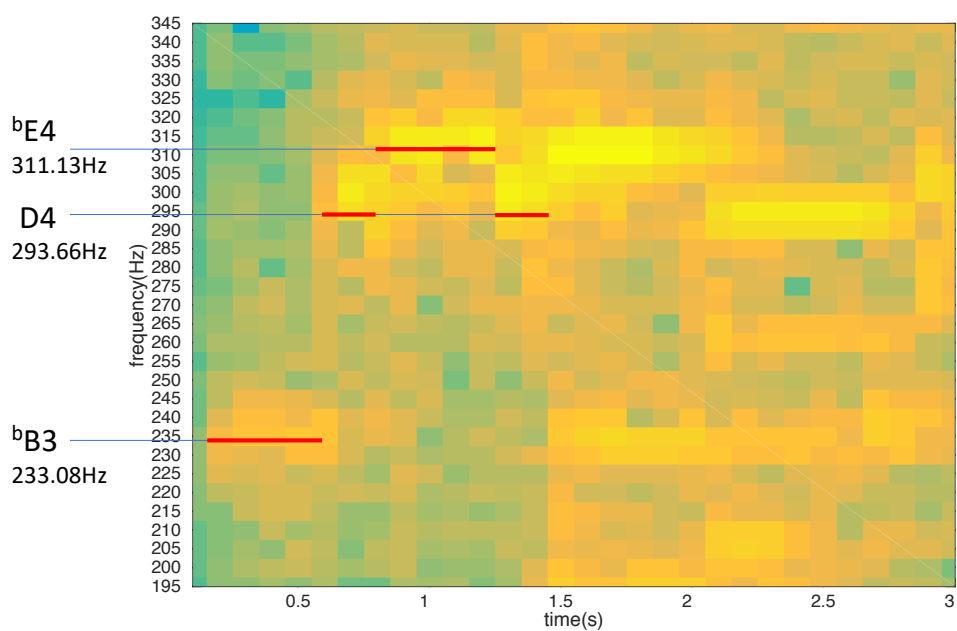


图 4: 小提琴奏鸣曲时频图

刚才听到的是卡米尔·圣桑的作品《降 E 调小提琴和钢琴奏鸣曲 2 号 Op.102 (1896)》¹，可以找到它的曲谱，如图 5 所示。



图 5: 小提琴奏鸣曲乐谱

图 4 中标记的是它的第一小节，放大了如图 6 所示。²



图 6: 小提琴奏鸣曲第一小节曲谱

请大家返回上页仔细观察图 4：在 0.7 秒和 1.4 秒附近，强度最大的频率值显然大于 293.66Hz。是演奏家没有“摠准”？还是 A4 不等于 440Hz？

2 艺术

- 为什么 A4 是 440Hz？

请参考“<https://www.zhihu.com/question/20277923>”

- A4 一定严格是 440Hz 吗？

请参考“<https://zhuanlan.zhihu.com/p/24701338>”

— “A=440 在流行乐最为通用。不过，在交响乐团中标准音高会有不同的规定。纽约爱乐乐团标准音高设为 A=442Hz，波士顿交响乐团使用 A=441Hz，而欧洲许多乐团则使用 A=443Hz。如果标准音高变了，那么其他所有音名的音高都会随之升高或者降低。”

¹Camille Saint-Saëns, Violin Sonata No.2, Op.102

²图 5 所示这部分是每小节四拍，但是第一小节只有一拍。其中第一个和第三个音各占了四分之三拍，第二个和第四个音各占了四分之一拍。从第二小节开始每小节四拍。

3 科学

本次大作业希望用科学理论（信号与系统）和工程技术（编程）回答孟教授之问。
开始工作之前，我们先做最苛刻的假设：即人耳只能根据单个乐音判定频率。

3.1 一个乐音必须持续多长时间才能测出 3Hz 的频差？

课上讲过：理想低通阶跃响应的上升时间和带宽之乘积是常量。这表明系统在时域的分辨能力与频域的分辨能力相互制约，为了提高频域分辨力（更准确估计瞬时频率），必须以减小时域分辨力（增大信号持续时间）为代价。

请认真阅读教材 [1]6.10 节“测不准（不定度）原理及其证明”，回答本小节问题。

3.2 前述小提琴奏鸣曲对应的 A4 是 440Hz 吗？

为了回答这个问题，可以考虑分析 1) 所有 A4 乐音，2) 所有乐音，3) 部分乐音。无论分析哪种对象，只要理论和方法恰当，都是很好的。

求解方法不限，鼓励大家探索任何解法。我大致想了这样三种方案，仅供参考。

3.2.1 直接计算

先把音乐信号按频率分成若干段，每段对应一个乐音。再计算这段乐音的基音频率，最后将它折算到 A4。

也可以只计算 A4 附近的乐音频率，但这样做限制了数据量和说服力。

3.2.2 利用先验信息计算

前述方案在分段和提取基音频率时未使用待分析信号是乐音的知识。如果应用这个先验信息，有可能显著提高这两个环节的正确率。具体说，A4 总是在 440Hz 左右，而所有乐音的基音频率总是约等于 $440 \times 2^{\frac{n}{12}}, n \in \mathbb{Z}$ 。

3.2.3 假设检验

再进一步，也许可以在假设检验的框架下更规范的解决这个问题。

关于假设检验，可以参考“<https://wenku.baidu.com/view/183f4d6869dc5022abea0098.html>”

具体的，本问题即从如下两个假设中选择一个接受：

$$\begin{cases} H_0 : A4 = 440\text{Hz}; \\ H_1 : A4 \neq 440\text{Hz}. \end{cases}$$

3.3 前述世界著名乐团的标准音高到底是多少？

基于收集到的数据，再利用上述技术，即可判定题面中著名乐团的标准音高是否真得比 440Hz 高。
一个重要步骤是收集数据：最好是小提琴独奏，其次只有钢琴伴奏，分析交响乐应该是很难的事情。

4 作业和评分规则

4.1 作业

第 3 节中三个小节标题里的问题。

网络学堂附件里包括了《降 E 调小提琴和钢琴奏鸣曲 2 号 Op.102 (1896)》完整乐曲的 mp3 文件和前 10 秒的 wav 文件。

4.2 评分规则

本次大作业采用“上传报告、代码和实验结果，由助教评判”的传统评分方式。具体流程如下：

- 本作业满分 100 分（第一题 25 分，第二题 50 分，第三题 25 分），计入总评成绩时加 3 分。
- 请在提交时把如下所有内容放在一个名为“学号 _ 姓名”的文件夹里，将文件夹压缩后上传到学堂作业区。
 - 读我：1 个名为 readme.txt 的文件，介绍当前目录下的所有内容；
 - 报告：1 个名为 report.pdf 的文档，描述你的答案、理由和解答过程、程序的运行方法、以及实验结果；
 - 代码：一个名为 code 的文件夹，内含解答过程中必要的代码和数据文件，包括源代码和可执行代码（如果有的话）；
 - 支持库：一个名为 support 的文件夹，内含在你的代码中用到但不是你亲自实现的支持性素材，包括开源代码或库等；如果这个文件夹下的内容超过 10M，请不要提供素材本身，而是提供下载地址，并注明安装和操作方法。
- 要求独立完成。禁止任何形式的抄袭。任何抄袭、剽窃等学术不端行为必将受到严厉打击。
- 不限解决方法，不限开发环境；可以用任何工具软件或开源代码（不包括其他人专门为解决本作业开发的工具），但必须注明出处。

5 文件操作和绘图工具介绍

可以用 MATLAB 编程实现，可能用到的专业功能函数如下表所示 [2]。

函数名	类型	说明
wavread	MATLAB 标准	将 wav 文件中的数据读入内存
wavwrite	MATLAB 标准	将内存中的数据写入 wav 文件
plot	MATLAB 标准	绘制波形
sound	MATLAB 标准	播放声音
window	MATLAB 标准	窗函数
spectrogram	MATLAB 标准	短时傅里叶变换并绘图

也可以用 Python 或其他任何语言实现。

6 致谢

1. 感谢东南大学信息科学与工程学院孟桥教授在“高校信号与系统教学群”里提出这个问题，启发我以此作为本学期的大作业。
2. 感谢无 76 班王威同学（小提琴业余十级）给我扫小提琴盲，现场演奏和答疑，让我下定决心以此为大作业。
3. 感谢我的研究生陶一嘉（钢琴业余十级）在我调试代码的过程中帮助我判定音准，以及告诉我乐曲详细信息、帮我寻找五线谱、告诉我第一小节的音调和节奏等。
4. 感谢助教博士生焦宇晨、李根和吴越审阅初稿并提出修改建议。

参考文献

- [1] 郑君里、应启珩、杨为理，《信号与系统》第三版，北京：高等教育出版社，2011.3
- [2] 谷源涛、应启珩、郑君里，《信号与系统——MATLAB 综合实验》，北京：高等教育出版社，2008.1

