

信号与系统— MATLAB 综合实验之音乐合成¹

谷源涛 应启珩 郑君里

二〇〇九年九月一日

¹摘录于草稿，可能和纸质出版物有出入。

目录

第一章 音乐合成	1
第一节 背景知识	1
1.1.1 乐音特征	1
1.1.2 乐音基波的构成规律	2
1.1.3 乐音谐波的作用—音色	4
1.1.4 乐音的波形包络	4
1.1.5 音调的持续时间	5
1.1.6 音符的迭接	6
第二节 练习题	6
1.2.1 简单的合成音乐	6
1.2.2 用傅里叶级数分析音乐	7
1.2.3 基于傅里叶级数的合成音乐	8

第一章 音乐合成

本章将基于傅里叶级数和傅里叶变换等基础知识，应用第一篇讲授的 MATLAB 编程技术，在电子音乐合成方面做一些练习。希望读者通过本章练习，可以增进对傅里叶级数的理解，并能够熟练运用 MATLAB 基本指令。本章包括两部分，第一部分介绍乐理和电子音乐的基本知识，第二部分给出详细的练习内容和编程步骤。相信读者对此会产生强烈兴趣。

第一节 背景知识

1.1.1 乐音特征

人类听觉可以感受到的声音大体上可划分为噪声、语音、乐音……几种类型。关于噪声和语音的特征及其性能分析将在本科高年级或研究生的课程中专门研究。本练习初步介绍乐音的基本概念，并利用 MATLAB 编程实现音乐合成系统。

音乐是乐音随时间流动而形成的艺术。用通信与电子技术的术语解释就是周期信号频率（某种指定规律的频谱结构）随时间节奏变化的一种表述。乐谱上的每个音符表达了此时此刻规定出现的信号频率和持续时间。

乐音的基本特征可以用基波频率、谐波频谱和包络波形三个方面来描述，下面将分别说明。

很明显，认识乐音的频谱规律之后，我们就可以借助电子系统从软件和硬件两种角度模仿各种乐器产生的声音，实现所谓电子音乐系统。

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

图 1.1: 钢琴键盘和相应频率

电子音乐系统是一门新兴的交叉技术科学，涉及计算机、集成电路、电子线路、信号处理、声学等多种领域，研究与应用前景广阔。本练习只是非常简单的入门介绍。

1.1.2 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 1.1 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。

从图 1.1 可以看到，靠下边的 A 键称为小字组 A ，它的频率值 $f_{A0} = 220\text{Hz}$ 。而靠上面的另一个 A 键是小字一组 A ，它的频率值是 $f_{A1} = 440\text{Hz}$ 。两者为二倍频率关系，即 f_{A1} 相当于 f_{A0} 的二次谐波。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”)。

从小字组 A 到小字一组 A 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 f_{A0} 导出小字一组 C (中央 C) 的频率

为

$$f_{C1} = 220 \times 2^{\frac{3}{12}} \text{Hz} = 261.63 \text{Hz} \quad (1.1)$$

或利用 f_{A1} 也可导出同样结果。

人耳听觉辨识振动频率的能力大约在 $\pm 0.5\%$ ，上述规律刚好符合这一要求。

从图 1.1 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 $\sharp D$ ，……。

下面给出“唱名”的概念。所谓唱名是指平日读乐谱唱出的 do、re、mi、……。每个唱名并未固定基波频率。当指定乐曲的音调时才知道此时唱名对应的音名，也即确定了对应的频率值。

例如，若指定乐曲为 C 调（或称 C 大调），此时唱名与音名的对应关系如图 1.2(a)。图中将唱名 do、re、mi、……以简谱符号 1、2、3、……代替。可见，它的“1”对应“ C ”也即基波频率为 261.63Hz。全部唱名对应键盘的白键。还可看出，3、4 分别对应 E 、 F 二者之间是半音阶，7、 $\dot{1}$ 之间对应 B 、 C 也属半音。其它唱名之间都为全音。

如果改为 F 调（ F 大调），唱名与音名的对应关系如图 1.2(b)。它的“1”与 F 对应，频率值是 349.23Hz。为了保持 3、4 以及 7、 $\dot{1}$ 之间为半音之规律，只有 C 调的全部唱名都与白键对应。而其它各调都需要引用黑键。如上例 F 调的“4”对应 bB ，其它各调对应黑键的规律可作为练习请读者自行导出。

下面给出一个乐曲实例，练习写出每个唱名对应的基波频率值，如图 1.3。这是《东方红》的开头两句曲谱，用简谱写出。曲调定为 F ，即 $1 = F$ ，于是可查出第一个音 5 对应 C ，频率值为 523.25Hz，其它音之频率可依次写出。稍后，我们将以此为基础进行音乐合成练习。

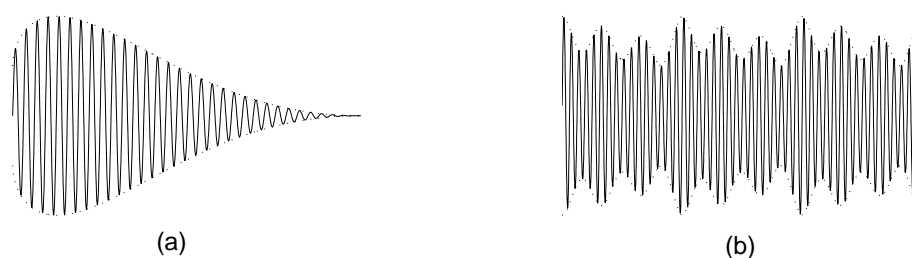


图 1.4: (a) 钢琴和 (b) 管乐的信号波形

- 连续型（风琴、手风琴、弦乐）
- 弹奏型（钢琴、吉他）
- 拨奏型（琵琶、月琴、曼多林）
- 击奏型（木琴、木鱼）
- 吹奏型（管乐）
- 颤音型（调频信号，如小提琴揉音）

图 1.4 示出钢琴与管乐的波形，可以看出二者包络之区别。其它类型省略。

在乐音合成实验中，为简化编程描述，通常把复杂的包络函数用少量直线近似，于是，乐音波形之包络呈折线。

1.1.5 音调的持续时间

在我们最简单的例子中，每个音调都可以用连续的一段正弦信号并带有一小段静音（停顿）来表示。停顿保证我们可以区分开连续的相同音调。每个音调的持续时间取决于它是全音符、二分音符、四分音符还是八分音符等等。四分音符的持续时间是八分音符的两倍。而每个音符之后的停顿时间应该是相同的，不随音符的长度而变化。在乐谱中，更长一些的停顿要用休止符来表示。图 1.3 中左侧的 $\frac{2}{4}$ 表示每小节有二拍和每个四分音符持续时间为一拍。对这段乐曲来说，一拍大约是 0.5 秒。

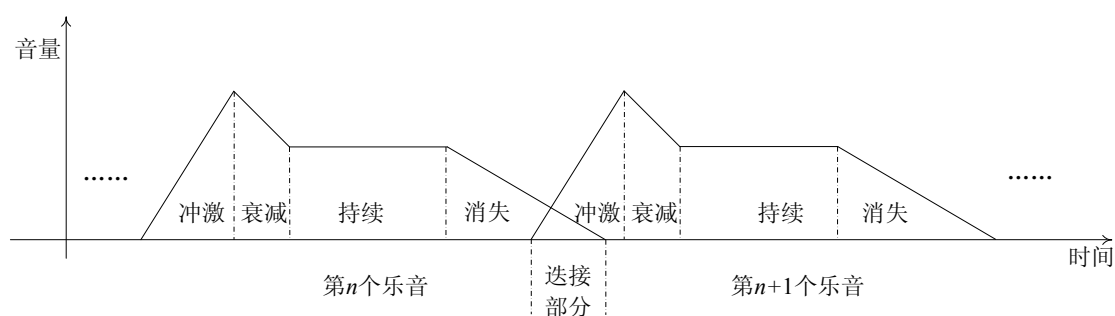


图 1.5: 音量变化

1.1.6 音符的迭接

某些艺术家演奏乐器时相邻的音符会有些重叠，就是当一个音调消失的时候，另一个被演奏出来。数学的表示就是两个信号有些重叠。这样听起来会更连续，较少断音。注意：这里说得迭接和前面要求的停顿并不矛盾，迭接时两个信号的幅度差别必须足够大以保证被区分开。

第二节 练习题

1.2.1 简单的合成音乐

(1) 请根据《东方红》片断的简谱和“十二平均律”计算出该片断中各个乐音的频率，在 MATLAB 中生成幅度为 1、抽样频率为 8kHz 的正弦信号表示这些乐音。请用 sound 函数播放每个乐音，听一听音调是否正确。最后用这一系列乐音信号拼出《东方红》片断，注意控制每个乐音持续的时间要符合节拍，用 sound 播放你合成的音乐，听起来感觉如何？

(2) 你一定注意到 (1) 的乐曲中相邻乐音之间有“啪”的杂声，这是由于相位不连续产生了高频分量。这种噪声严重影响合成音乐的质量，丧失真实感。为了消除它，我们可以用图 1.5 所示包络修正每个乐音，以保证在乐音的邻接处信号幅度为零。此外建议用指数衰减的包络来表示¹。

¹根据人耳听觉特性，当主观感受声强为线性变化时，声音信号的功率实际呈指数变化。

(3) 请用最简单的方法将 (2) 中的音乐分别升高和降低一个八度。(提示: 音乐播放的时间可以变化) 再难一些, 请用 `resample` 函数 (也可以用 `interp` 和 `decimate` 函数) 将上述音乐升高半个音阶。(提示: 视计算复杂度, 不必特别精确)

(4) 试着在 (2) 的音乐中增加一些谐波分量, 听一听音乐是否更有“厚度”了? 注意谐波分量的能量要小, 否则掩盖住基音反而听不清音调了。(如果选择基波幅度为 1, 二次谐波幅度 0.2, 三次谐波幅度 0.3, 听起来像不像象风琴?)

(5) 自选其它音乐合成, 例如贝多芬第五交响乐的开头两小节。

1.2.2 用傅里叶级数分析音乐

现在我们开始要处理真实的音乐信号了! 请用 `load` 命令载入附件光盘中的数据文件“guitar.mat”, 工作区会出现两个新的变量 `realwave` 和 `wave2proc` (可以用 `who` 查看变量名), 如图 1.6 和 1.7 所示。其中前者是从一段吉他乐曲 (附件光盘上的“fmt.wav”) 中截取下来的真实信号, 后者是用信号处理方法得到的这段信号的理论值, 它们的抽样率都是 8kHz。

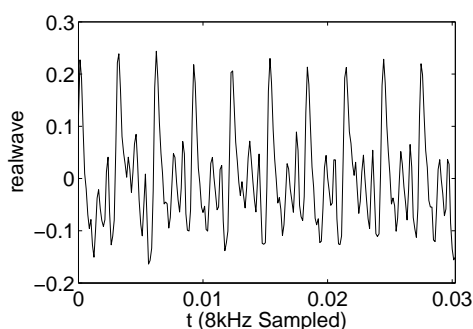


图 1.6: 真实吉他音

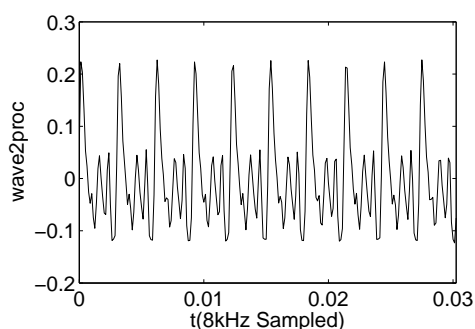


图 1.7: 待处理的吉他音

(6) 先用 `wavread` 函数载入光盘中的 `fmt.wav` 文件, 播放出来听听效果如何? 是否比刚才的合成音乐真实多了?

(7) 你知道待处理的 `wave2proc` 是如何从真实值 `realwave` 中得到的么? 这个预处理过程可以去除真实乐曲中的非线性谐波和噪声, 对于正确分析音调是非常重要的。提示: 从时域做, 可以继续使用 `resample` 函数。

(8) 这段音乐的基频是多少？是哪个音调？请用傅里叶级数或者变换的方法分析它的谐波分量分别是什么。提示：简单的方法是近似取出一个周期求傅里叶级数但这样明显不准确，因为你应该已经发现基音周期不是整数（这里不允许使用 `resample` 函数）。复杂些的方法是对整个信号求傅里叶变换（回忆周期性信号的傅里叶变换），但你可能发现无论你怎么提高频域的分辨率，也得不到精确的包络（应该近似于冲激函数而不是 `sinc` 函数），可选的方法是增加时域的数据量，即再把时域信号重复若干次，看看这样是否效果好多了？请解释之。

(9) 再次载入 `fmt.wav`，现在要求你写一段程序，自动分析出这段乐曲的音调和节拍！如果你觉得太难就允许手工标定出每个音调的起止时间，再不行你就把每个音调的数据都单独保存成一个文件，然后让 MATLAB 对这些文件进行批处理。注意：不允许逐一地手工分析音调。编辑音乐文件，推荐使用“CoolEdit”编辑软件。

1.2.3 基于傅里叶级数的合成音乐

现在进入了合成音乐的高级境界，我们要用演奏 `fmt.wav` 的吉他合成出一段《东方红》。

(10) 用 (7) 计算出来的傅里叶级数再次完成第 (4) 题，听一听是否像演奏 `fmt.wav` 的吉他演奏出来的？

(11) 也许(9) 还不是很像，因为对于一把泛音丰富的吉他而言，不可能每个音调对应的泛音数量和幅度都相同。但是通过完成第 (8) 题，你已经提取出 `fmt.wav` 中的很多音调，或者说，掌握了每个音调对应的傅里叶级数，大致了解了这把吉他的特征。现在就来演奏一曲《东方红》吧。提示：如果还是音调信息不够，那就利用相邻音调的信息近似好了，毕竟可以假设吉他的频响是连续变化的。

(12) 现在只要你掌握了某乐器足够多的演奏资料，就可以合成出该乐器演奏的任何音乐，在学完本书后面内容之后，试着做一个图形界面把上述功能封装起来。



MATLAB 知识点 (1) — 访问标准格式文件

为了充分发挥其强大的数据处理能力的影响，MATLAB 支持对工业界多种标准格式文件的读写访问。表 1.1 中列出了一些常用的函数。

表 1.1: 访问标准格式文件的常用函数

函数名	说明
dlmread dlmwrite textscan	
dlmread dlmwrite csvread csvwrite 等	访问各种 ASCII 文本格式文件
xlsread , xlswrite	访问微软 EXCEL 格式文件
wavread , wavwrite	访问微软 WAV 格式音频文件
auread , auwrite	访问 NeXT/SUN 格式音频文件
imread , imwrite	访问 JPEG, TIFF, GIF, BMP 等多种格式图像文件
aviread , avifile , 等	访问 AVI 视频文件

为了支持用户在程序中保存 Figure 窗口中的图形图像，MATLAB 提供了 saveas 函数，和用户在窗口中点击“File→Save As”一样，它支持的包括 pdf 在内的图像文件格式有 16 种之多！