



中国科学技术大学
University of Science and Technology of China

语音信号处理基础 课程实验报告

学生姓名：金桥
学 号：PB19061272

提交时间：
报告评分：

说明：

1. 每一道实验题目都需要按照第2-3页的模板填写实验报告；例如，第一次编程实验有2道题目，需要填写2次；
2. 中英文皆可；
3. 不是每一项都需要填写，根据实验难度选择。但实验分析，实验步骤，实验结果不可少。

第一次实验报告

Task 1

题目分析

- 概括题意
 - 本次实验要求用 MATLAB 读取一段采样率 f_s 为 16kHz 的音频信号，并将其带宽滤波到 6、4、2kHz。之后保存滤波后的文件，并比较不同带宽的音频信号在听觉上的差异，从而判断低通滤波对声音清晰度与音质的影响。
- 理论分析
 - 先用读取 .wav 文件，对其进行 fft 变换获取频谱，再对频谱滤波后逆 fft 变换并保存即可。
- 预期结果
 - 声音的频率对应着音调，滤去高频部分会使音频信号音调变低。
 - 高频部分信号也包含部分信息，滤去高频部分可能会使声音变得模糊。

实验过程

- 实验代码（部分）

```
% 读取音频文件
[y,fs]=audioread('test_16k.wav');
N=length(y);
Y=fft(y); % 快速傅里叶变换获取频谱

% 频谱滤波，并还原为语音信号
fs1=6000; index1=floor(fs1*N/fs);
Y1=[Y(1:index1); zeros(N-index1,1)]; y1=real(ifft(Y1));
fs2=4000; index2=floor(fs2*N/fs);
Y2=[Y(1:index2); zeros(N-index2,1)]; y2=real(ifft(Y2));
fs3=2000; index3=floor(fs3*N/fs);
Y3=[Y(1:index3); zeros(N-index3,1)]; y3=real(ifft(Y3));

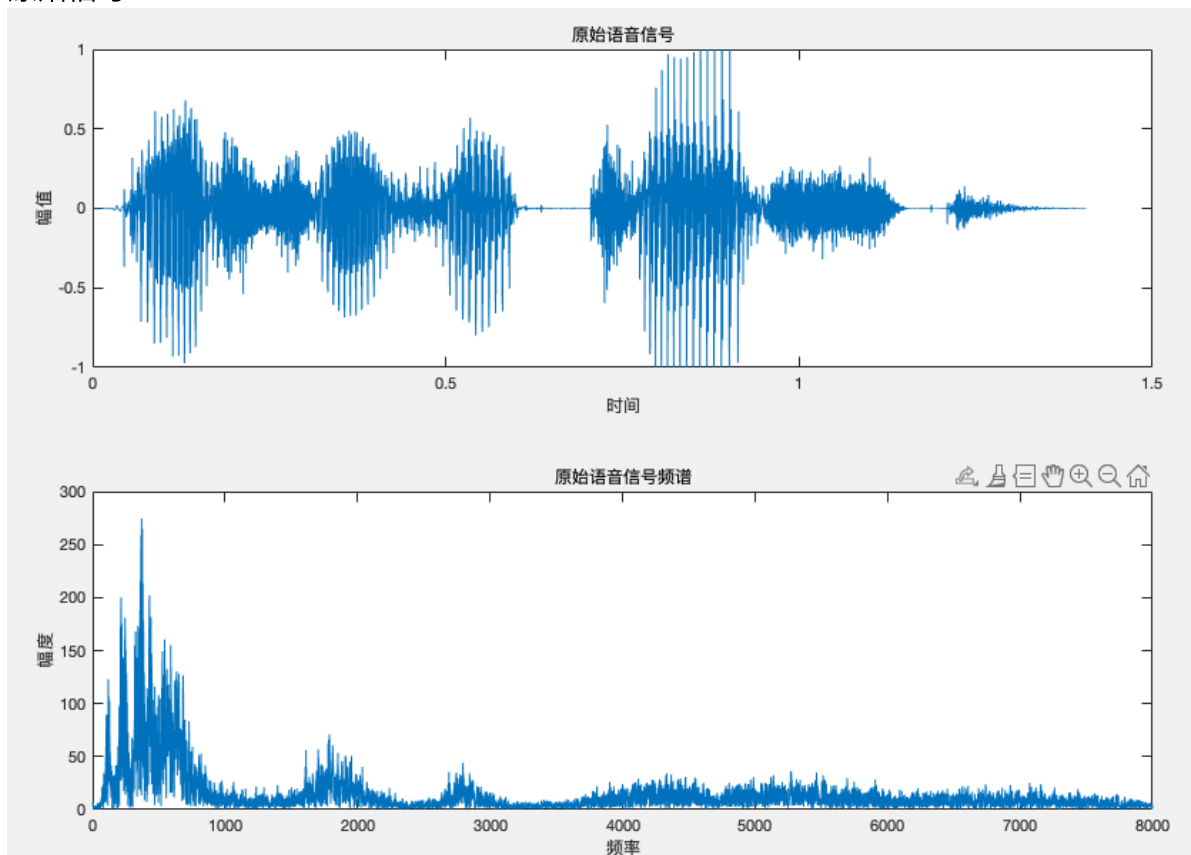
% 保存文件
audiowrite('test_6k.wav', y1, fs);
audiowrite('test_4k.wav', y2, fs);
audiowrite('test_2k.wav', y3, fs);
```

- 相关参数
 - N ：原始音频信号的长度（采样点数）
 - y 、 f_s 、 Y ：原始音频信号、采样率与频谱。
 - y_i 、 f_{s_i} 、 Y_i ($i=1, 2, 3$)：滤波后的音频信号、采样率与频谱，序号 123 依次对应 6k、4k、2k。

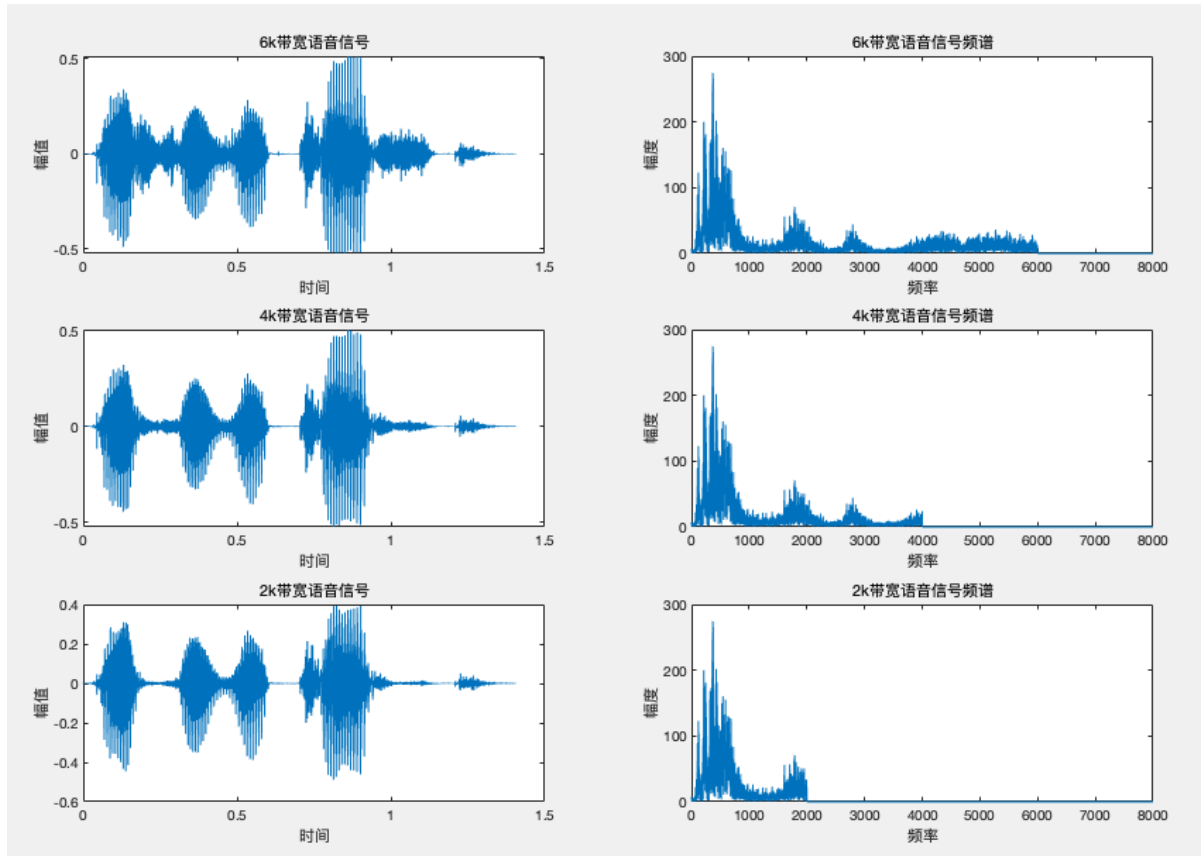
- y 、 Y 、 y_i 、 Y_i 长度均为 N ， Y 左右对称， Y_i 不对称。
- 实验步骤
 - 先用读取 .wav 文件到 y ，对其进行 fft 变换获取频谱 Y ，其中 Y 左右对称，长度为 N 。
 - 之后进行滤波
 - $\text{index1}=\text{floor}(fs1*N/fs)$ 获取带宽在 Y 中对应的下标。
 - $Y1=[Y(1:\text{index1}); \text{zeros}(N-\text{index1},1)]$ ，将 $Y(1:\text{index1})$ 的部分赋给 $Y1$ ， $Y1$ 其余部分补零。即只取原频谱的 0~6k，其它取零，完成了比较粗暴的滤波。
 - $y1=\text{real}(\text{ifft}(Y1))$ ，逆 fft 变换并取实部获取时域信号。
 - 保存文件

实验结果

- 实验结果
 - 原始信号



- 滤波信号



- 音频见附件
- 结果分析
 - 频域部分的改变比较明显，时域部分可以看出低带宽的信号比高带宽的信号少了不少高频细节，尤其在 1~1.2s 较明显。
 - 听滤波后的音频，可以很明显的感觉到音调与清晰度随带宽的降低而降低，符合预期结果。
- 疑惑与问题
 - 在 4k 带宽的信号中能听到细微的蜂鸣声，猜测原因
 - 播放器解析方式导致蜂鸣
 - 4k 频率较特殊，人耳会解析出蜂鸣

Task 2

题目分析

- 概括题意
 - 用 MATLAB 模拟双管声道模型，画出对数频率谱并标出共振峰。
 - 改变相关参数，观察不同参数的影响。
- 理论分析
- 预期结果

实验过程

- 部分代码

```
% 获取数据
[fr,vm]=get_log(rG,rL,l1,l2,A1,A2);
[pos, loc]=findpeaks(vm);
=====
```

```
=====
% 计算预备
c=35000;
l=l1+l2;
r1=(A2-A1)/(A2+A1);
tau1=l1/c;
tau2=l2/c;
tau=l/c;
numc=0.5*(1+rG)*(1+rL)*(1+r1);
denc1=r1*rG;
denc2=r1*rL;
denc3=rL*rG;

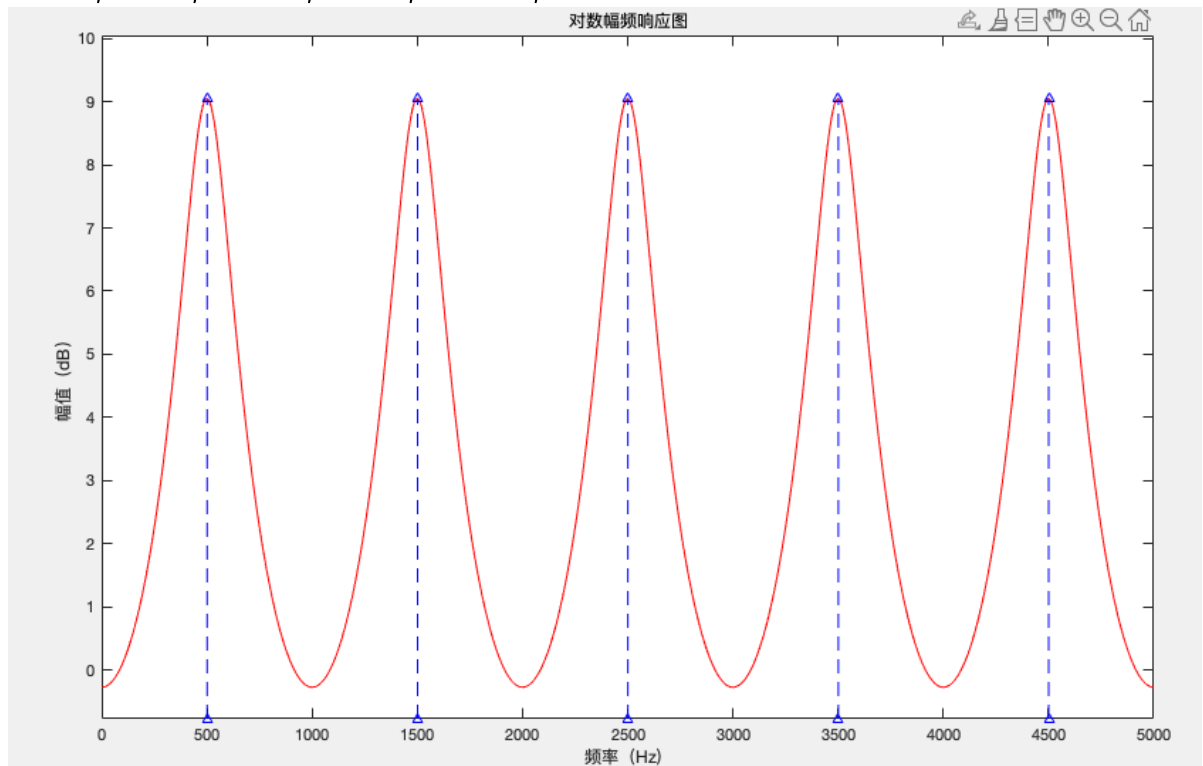
% 计算对数功率谱
freq=0:1:5000;
num=complex(numc*cos(2*pi*freq*tau),...
-numc*sin(2*pi*freq*tau));
den=complex(1+denc1*cos(2*pi*freq*2*tau1)+denc2*cos(2*pi*freq*2*tau2)...
+denc3*cos(2*pi*freq*2*tau),...
-denc1*sin(2*pi*freq*2*tau1)-denc2*sin(2*pi*freq*2*tau2)...
-denc3*sin(2*pi*freq*2*tau));
v=num./den;
vm=20*log10(abs(v));
fr=freq;
```

- 相关参数
 - c: 声速, 取 35000cm/s
 - li、Ai (i=1、2) : 通道 i 的长度 (cm) 与面积 (cm²)。
 - rG、rL: 声门、嘴唇处的反射系数。
 - fr: 绘制功率谱的频率范围为 0: 5000
 - vm: 得到的功率谱
 - loc、pos: 共振峰峰值的横、纵坐标。
 - get_log: 获取 vm 的自定义函数
- 实验步骤
 - 先设定各参数值
 - 再用 get_log 获取对数频谱 vm
 - 先对设定的参数预处理, 以简化表达式
 - 频率范围设为 0:5000, 步长为 1, 代入频谱响应得到 v
 - 取幅值的对数得到 vm
 - 最后用 findpeaks(vm) 获得共振峰的位置

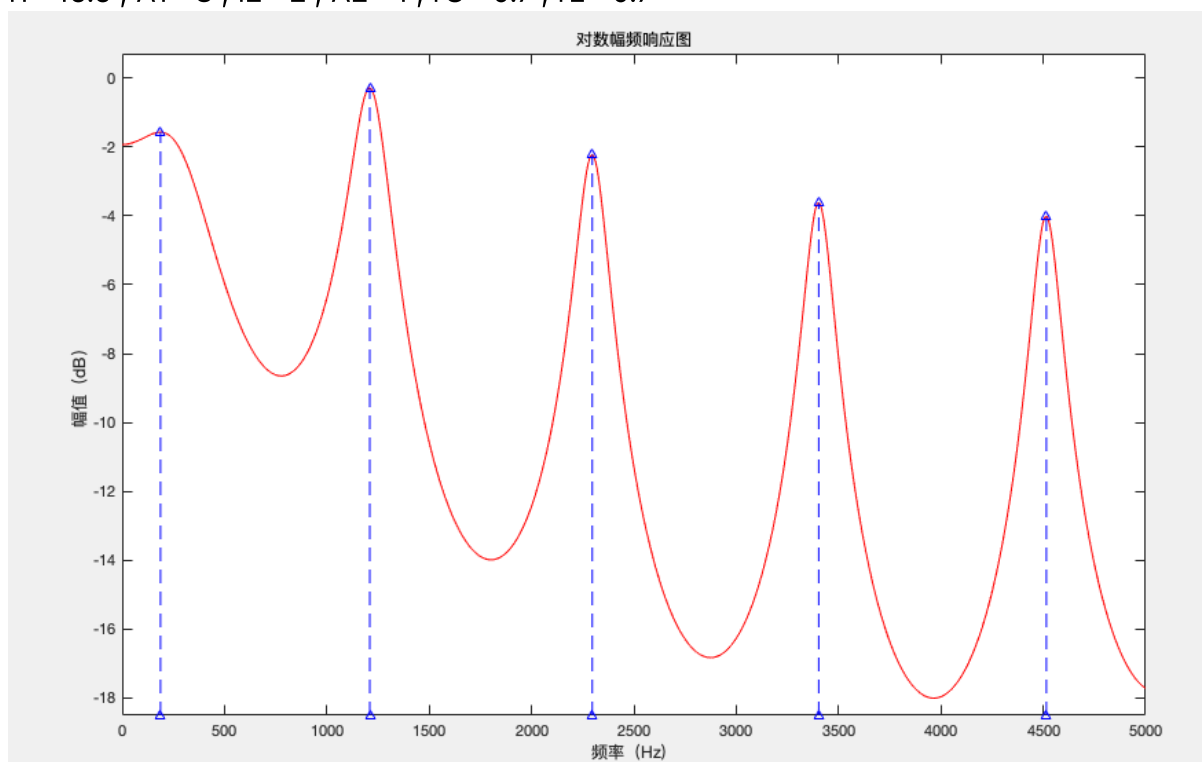
实验结果

- 实验结果

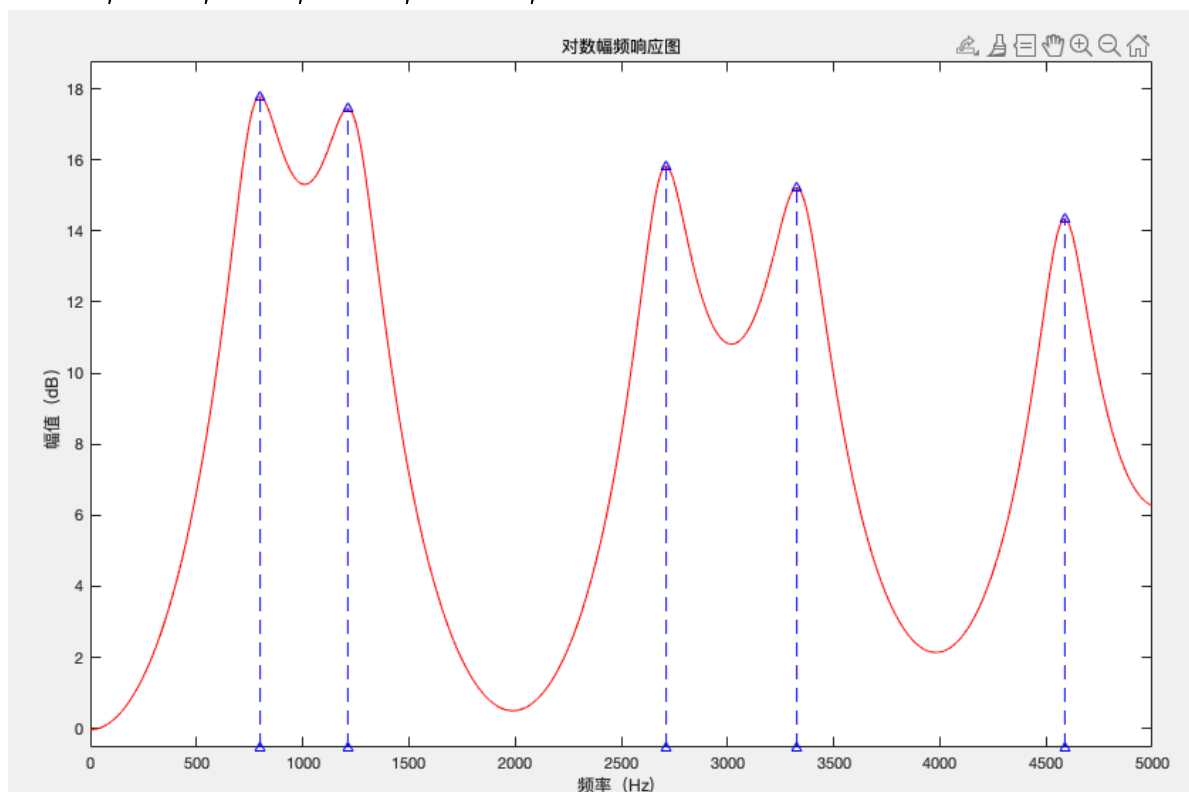
1. $l_1=10$, $A_1=1$, $l_2=7.5$, $A_2=1$, $r_G=0.7$, $r_L=0.7$



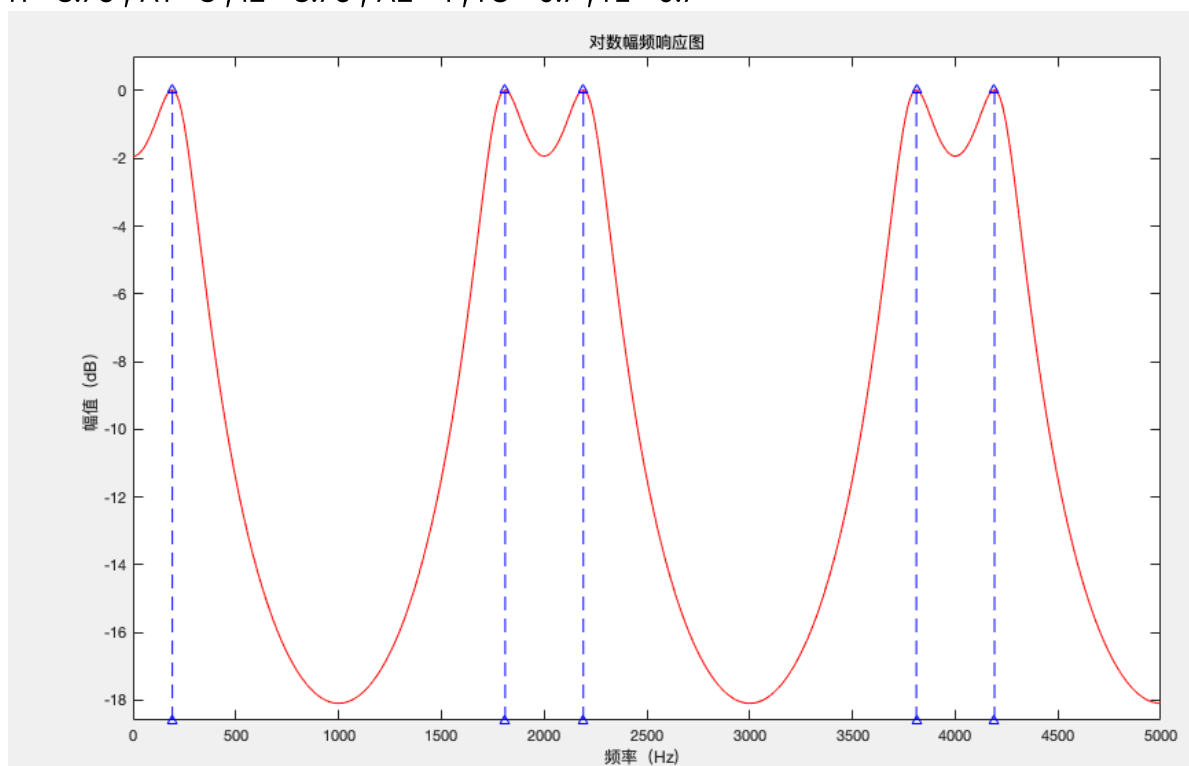
2. $l_1=15.5$, $A_1=8$, $l_2=2$, $A_2=1$, $r_G=0.7$, $r_L=0.7$



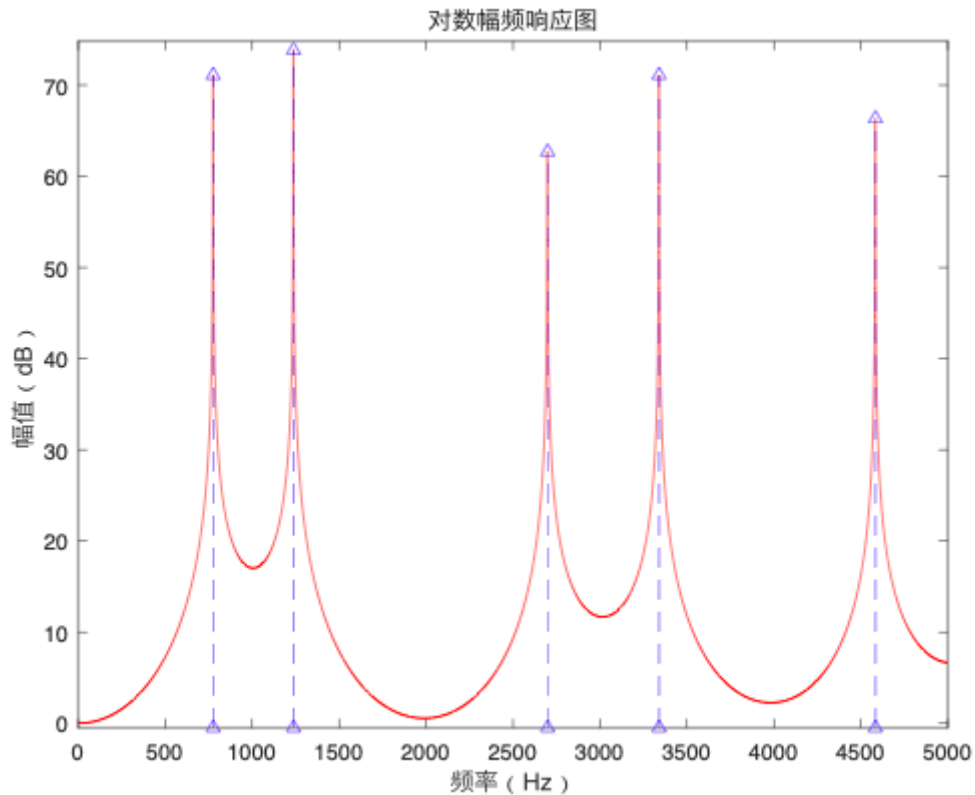
3. $l_1=9.5$, $A_1=1$, $l_2=8$, $A_2=8$, $r_G=0.7$, $r_L=0.7$



4. $l_1=8.75$, $A_1=8$, $l_2=8.75$, $A_2=1$, $r_G=0.7$, $r_L=0.7$



5. $l_1=9.5$, $A_1=1$, $l_2=8$, $A_2=8$, $r_G=1$, $r_L=1$ (仅取一例)



- 结果分析

- 可以看出 l_1 与 l_2 或 A_1 与 A_2 相等时, 代表双管系统具有较好的对称性, 此时图像也具有较好的对称性与周期性 (图 1 与图 4)。二者越相近, 图像对称性越好。
- 将图 2 图 4 与图 3 比较发现, $A_1:A_2=8:1$ 时, 幅频响应会有较大的衰减, 比例为 1:8 时, 会有较大的提升。这表明声音传播管道的面积大小变化具有很好的控制声音强度的功能, 这也与我们日常生活中的经验相符。
- 观察图 2 图 3 可以发现, l_1 大于 l_2 且 A_1A_2 不等时双管模型的幅频响应会随着频率的增高而不断衰减, 且 l_1 与 l_2 的差距越大, 衰减的越快。
- $r_G=r_L=1$ 时, 以图 5 为例, 图 5 与图 3 其他数据相同, 但图像有非常明显的区别
 - 反射系数变为 1 后, 幅频响应的增幅非常的高, 峰值比 0.7 时高了 50~60dB。
 - 共振峰十分突出, 峰值十分明显, 且左右峰的大小关系发生了变化。
 - 观察 2700Hz 与 4600Hz 处可以发现, 幅频响应不再严格地随频率增大而衰减。
- 反射系数改变带来的差距提醒我们在高反射系数的中, 双管模型的一般规律可能不再适用。