

香港中文大學(深圳)  
The Chinese University of Hong Kong, Shenzhen

## EIE3510 Digital Signal Processing-Term Project

# Composition and Instrumental Music Generation from Bird Songs

Yang Jiao 118010121

Jifei Zhao 118010437

13 December 2021





# Introduction



## Motivation

In the AI generation, researchers started exploring the boundary between machine intelligence and artworks by using the machine to generate music. Previous studies mainly focused on deep learning methods, which need massive data.

- We hope to design an **explicit and lightweight model** with only a tiny amount of unmusical audio input.
- Inspired by nature, we investigate the similarity between the singing of birds and instrumental music.



## Goal

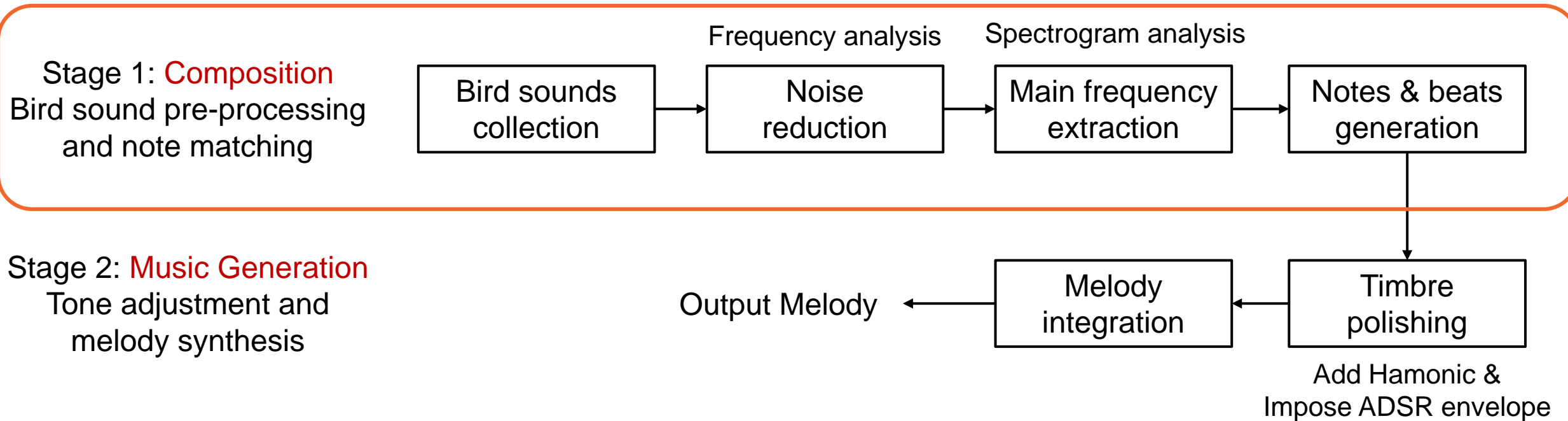
1. Map the bird sounds to a series of notes (Composition)
2. Play the melody by mimicking the timbre(音色) of piano (Music Generation)



# Methodology



## Framework of whole project :





# Stage 1—Bird sounds collection



## Dataset :

**xeno-canto**  
Sharing bird sounds from around the world

Search recordings... Search Advanced Search Tips

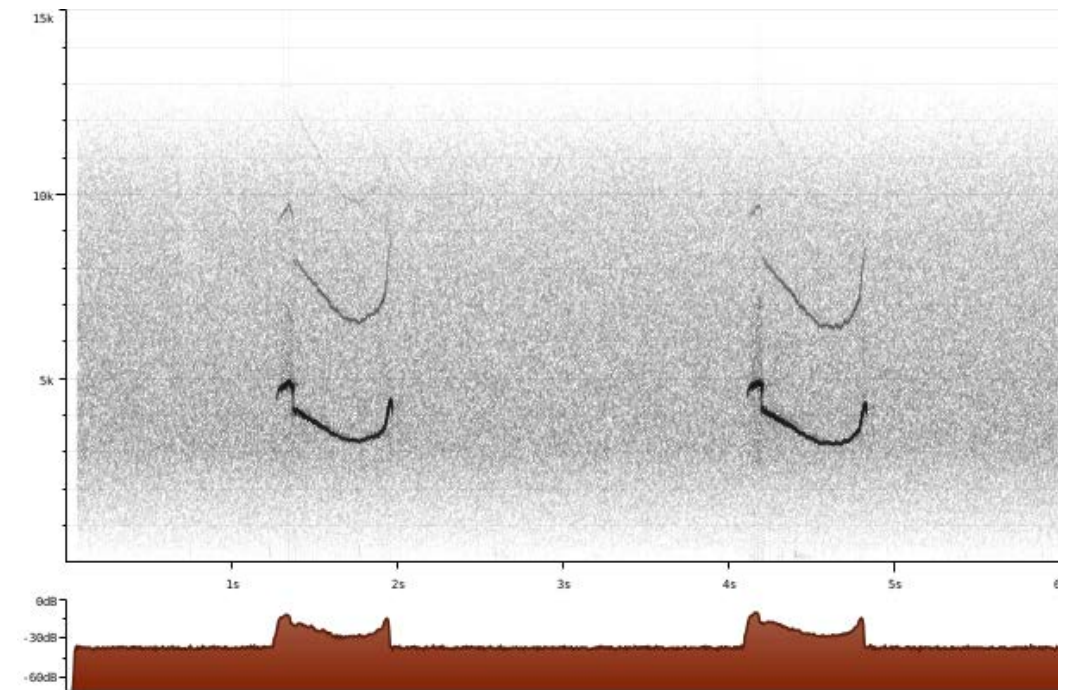
About Explore Upload Sounds Forum Mysteries Articles Log in / Register

**Recordings**  
669393 results from 10292 species (foreground species only) (0s)  
• Results format: detailed | concise | codes | sonograms

1 2 3 4 5 6 7 8 9 ... 22314 Next >

Common name / Scientific	Length	Recordist	Date	Time	Country	Location	Elev. (m)	Type	Remarks	Actions	Cat.nr.
太平洋潜鸟 (Gavia pacifica)	1:23	Jorge de Leon Cardozo & Susan Hochgraf	2021-07-30	01:00	United States	Gaswell Road, Utqiagvik, North Slope, Alaska	10	adult, call	Two adult birds calling from a small... more » [sono]		XC691192
太平洋潜鸟 (Gavia pacifica)	0:44	Jorge de Leon Cardozo & Susan Hochgraf	2021-07-30	01:00	United States	Gaswell Road, Utqiagvik, North Slope, Alaska	10	adult, call	Two adult birds calling as they flew... more » [sono]		XC691191
北极鸥 (Larus hyperboreus)	1:11	Jorge de Leon Cardozo & Susan Hochgraf	2021-07-30	00:00	United States	Gaswell Road, Utqiagvik, North Slope, Alaska	10	adult, call	Two adult birds calling as they flew... more » [sono]		XC691190
短嘴鸦 (Corvus brachyrhynchos)	0:30	Michael & Katie LaTour	2021-12-12	05:00	United States	Washington Township (near Sardinia), Brown County, Ohio	00	life stage uncertain, mimicry/imitation, sex uncertain	A murder of crows gathered in my... more » [sono]		XC691189
弗吉尼亚秧鸡 (Rallus limicola limicola)	1:19	Barry Edmonston	2021-12-12	15:00	Canada	Swan Lake, Victoria, Vancouver Island, BC	10	adult, call, sex uncertain	Calls from single individual moving in... more » [sono]		XC691188

<https://xeno-canto.org/>



Spectrogram information



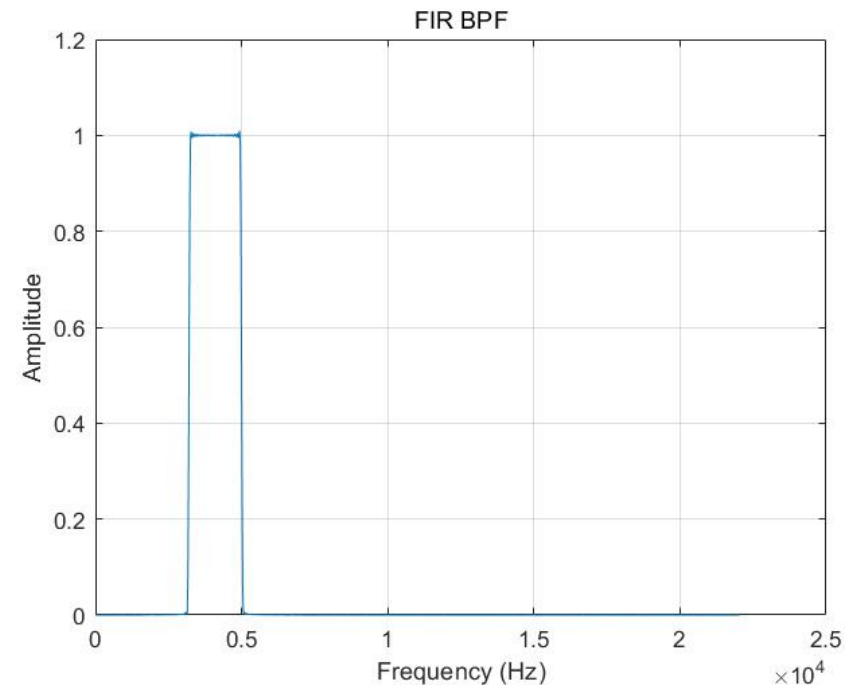
# Stage 1 — Noise Reduction

## 🎵 Ideal band-pass filter

- Perfect square wave
- Cutoff frequency:  $fc1 = 3200$  Hz,  $fc2 = 5000$  Hz

## 🎵 FIR band-pass filter

- Width of transition band: 100 Hz
- Ripples: [0.01, 0.01, 0.01]
- Cutoff frequency:  $fc1 = 3200$  Hz,  $fc2 = 5000$  Hz
- Kaiser window with  $M = 986$ ,  $\beta = 3.3953$



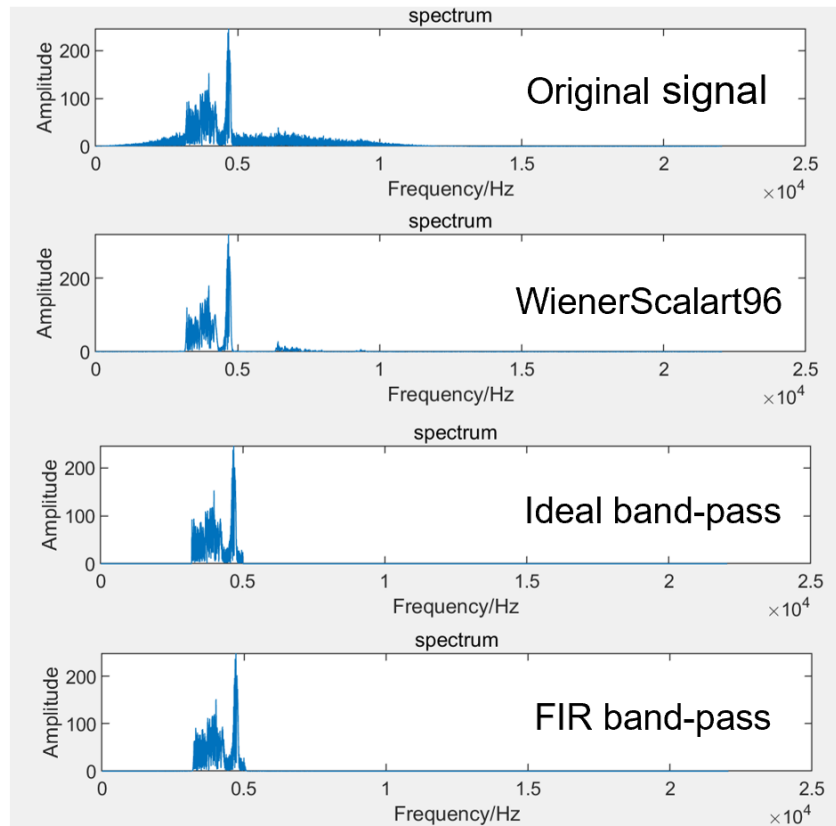


# Stage 1 — Noise Reduction



## Comparison between filters

- Filtering effect in frequency domain:



- Filtering effect on the basis of SNR and time complexity:

Filter	SNR (dB)	Time complexity
Ideal band-pass	10.8853	$\mathcal{O}(N \log_2 N)$
FIR band-pass	10.8257	$\mathcal{O}(N)$
[P. Scalart, 1996] Wiener estimation	13.2802	/



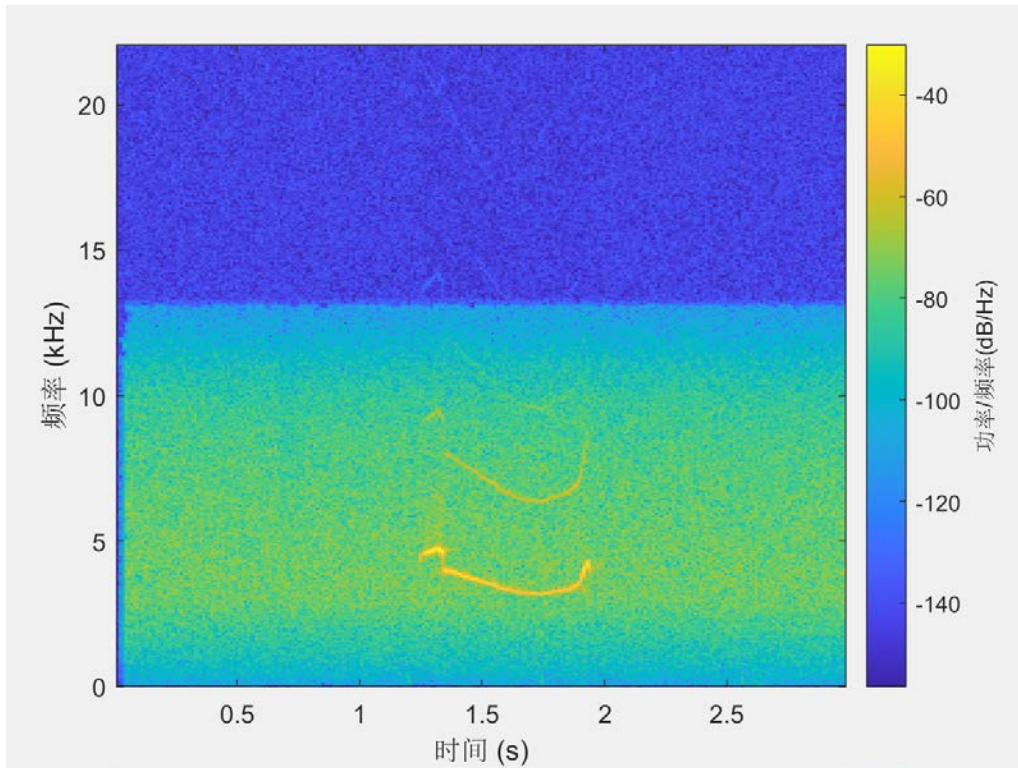
# Stage 1 — Main frequency extraction



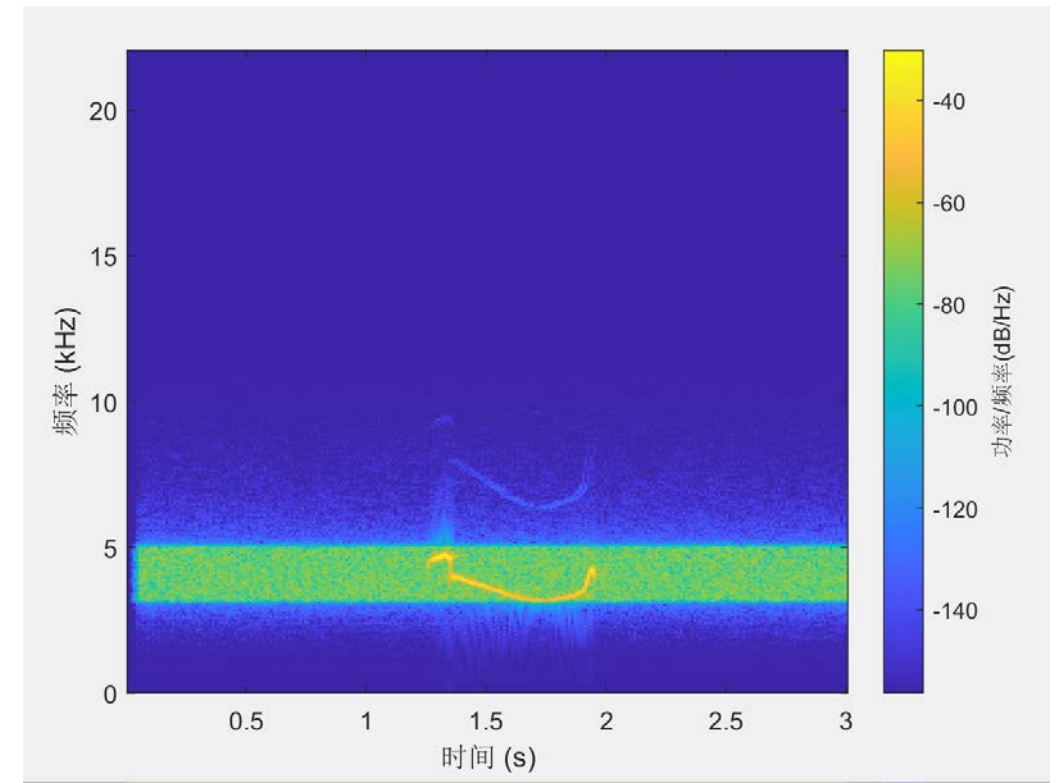
## Spectrogram analysis

We apply Short-time Fourier transform with:

- window length  $L=1000$  (23 ms Hamming window),
- window moves by  $R=500$  samples each time.



Original Spectrogram



Spectrogram after FIR BPF





# Stage 1 — Notes & beats generation

## Notes

- Starting from a frequency on the idealized standard piano tuned to  $A_4$  (440 Hz)
- Key number can be calculated from

$$Key(n) = \text{round}\left(12 \log_2 \left(\frac{f_{main}(n)}{440 \text{ Hz}}\right)\right) + 49 - 8$$

The 49<sup>th</sup> key  
on piano

Key number	MIDI note	Helmholtz name <sup>[5]</sup>	Scientific name <sup>[5]</sup>	Frequency (Hz) (Equal temperament) <sup>[6]</sup>
79	99	d <sup>♯</sup> '''/e <sup>b</sup> '''	D <sup>♯</sup> <sub>7</sub> /E <sup>b</sup> <sub>7</sub>	2489.016
78	98	d'''	D <sub>7</sub>	2349.318
77	97	c <sup>♯</sup> '''/d <sup>b</sup> '''	C <sup>♯</sup> <sub>7</sub> /D <sup>b</sup> <sub>7</sub>	2217.461
76	96	c''' 4-line octave	C <sub>7</sub> Double high C	2093.005
75	95	b'''	B <sub>6</sub>	1975.533
74	94	a <sup>♯</sup> '''/b <sup>b</sup> '''	A <sup>♯</sup> <sub>6</sub> /B <sup>b</sup> <sub>6</sub>	1864.655
73	93	a'''	A <sub>6</sub>	1760.000

Decrease the  
frequency by  
an octave





# Stage 1 — Notes & beats generation

## Beats

- **Whole note** duration = 1s
- Note usually divided by a factor of **powers of two**
- Whole note, half note, quarter note, and eighth note are considered
- Mapping rule between number of **repetitive frequency levels** & standard **duration** of notes

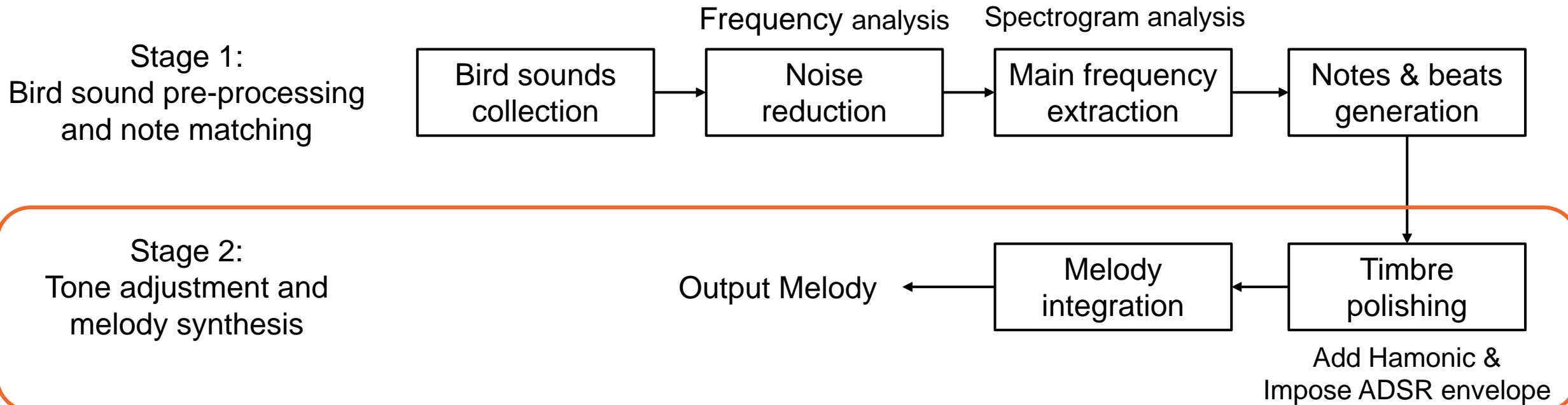
Number of repetitive frequency levels	Note value	Time duration (s)
1	Eighth note	0.125
2	Quarter note	0.25
3, 4, 5	Half note	0.5
6, 7, 8, 9, 10	Whole note	1





## Stage 2

# Tone adjustment and melody synthesis



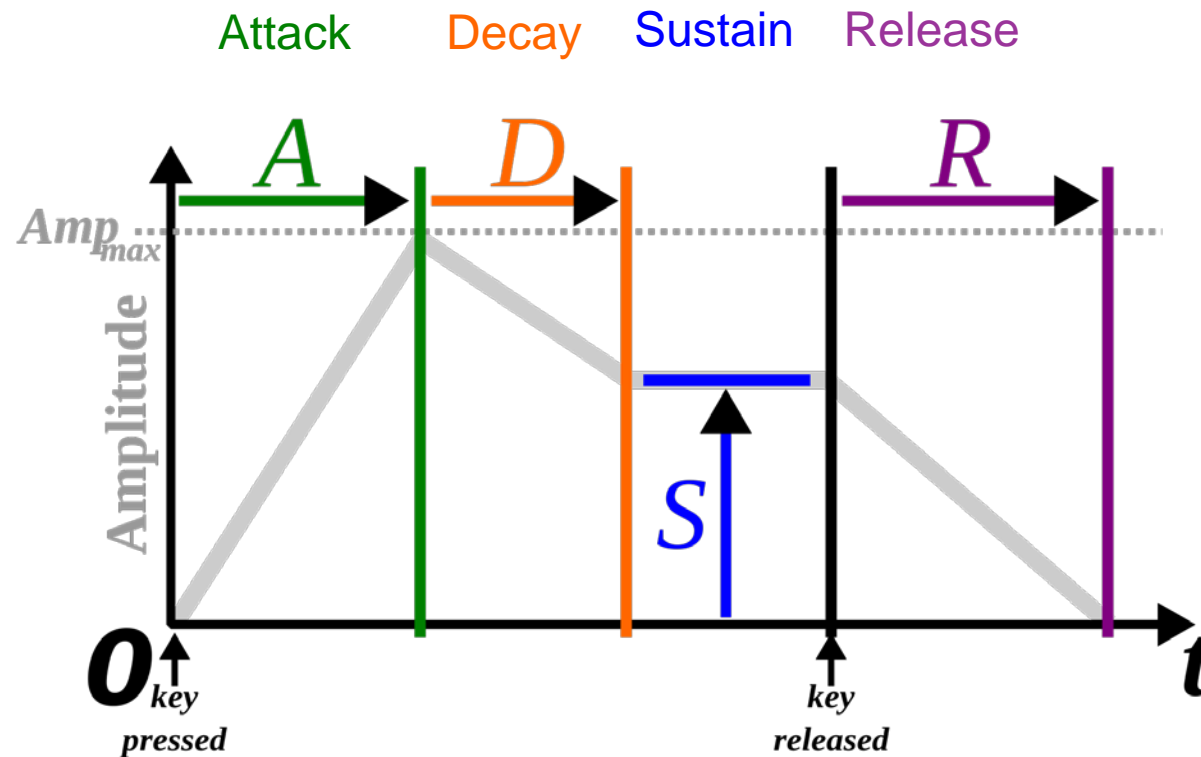


## Stage 2 — Timbre Polishing



*How is timbre determined for a kind of instrument?*

- Envelop: ASDR model
- Superposition of high order harmonics





## Stage 2 — Timbre Polishing

Idealized ADSR model

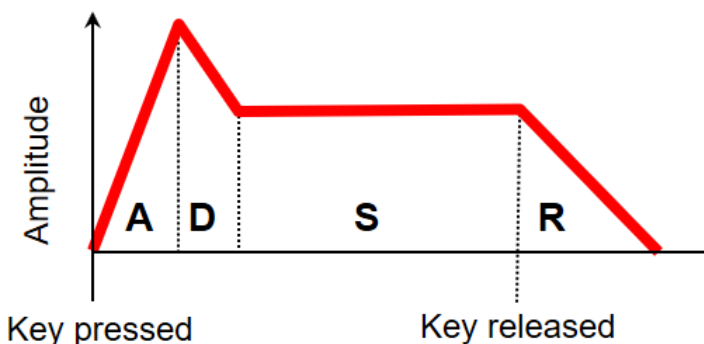
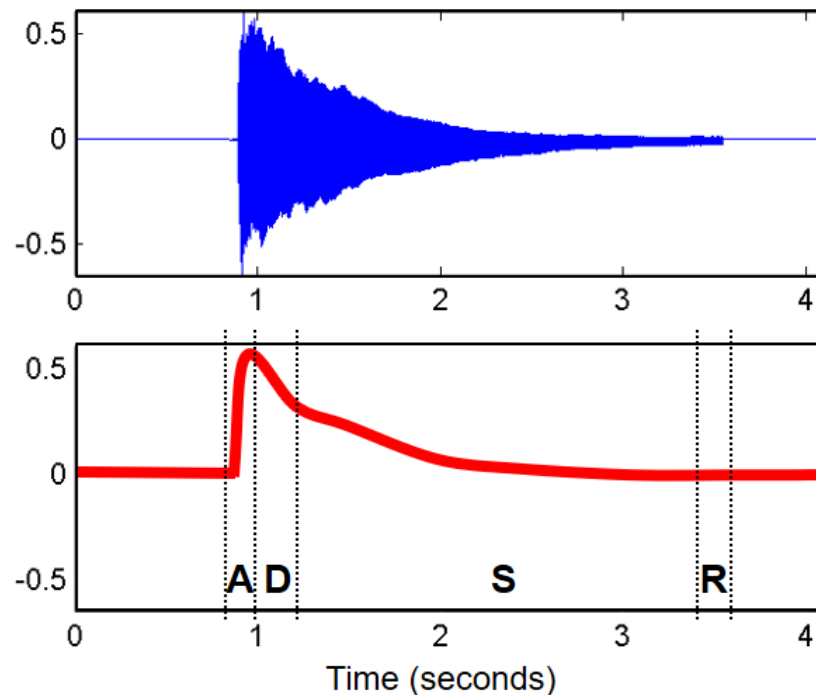
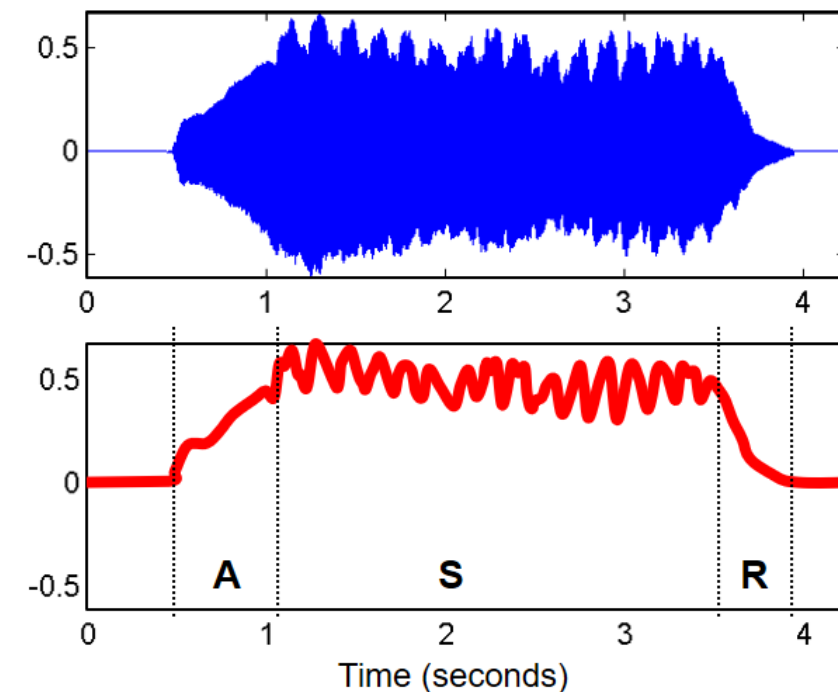


Figure 1.22b and Figure 1.23  
from [Müller, FMP, Springer 2015]

Piano sound



Violin sound



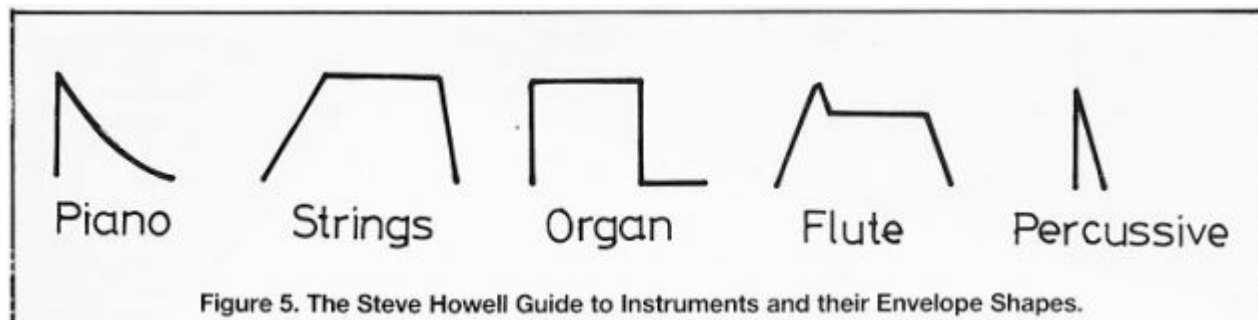
- Very **short attack** time
- **Decay drastically**
- Envelop approximated by an **exponential** function

- **Long attack** time for set up
- Long sustain with **vibration** and **tremolo**
- **Difficult** to mimic by an elementary function

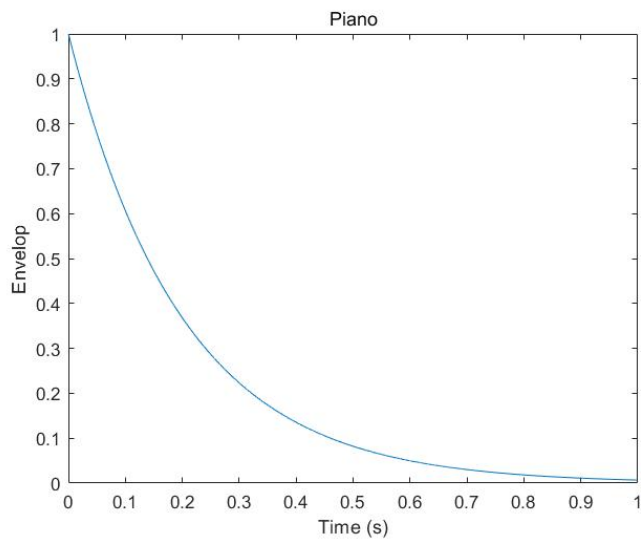


## Stage 2 — Timbre Polishing

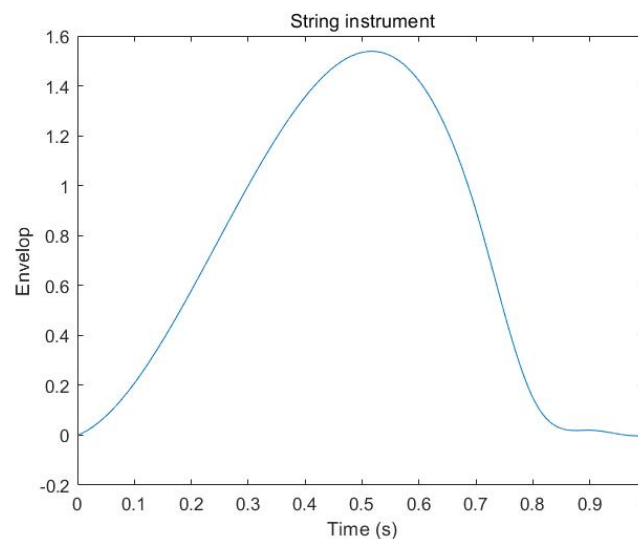
### Different envelop shapes



### Piano



### Violin



```
L = 1;  
y = [0, 1, 0.9, 0.15, 0.02, 0.005, 0];  
x = [0, 0.3, 0.7, 0.8, 0.9, 0.95, 1];  
fs = 44100; Ts = 1 / fs;  
t = 0:Ts:L;  
env = spline(x*L, y, t);
```

MATLAB code to generate  
the envelop of violin



## Stage 2 — Timbre Polishing

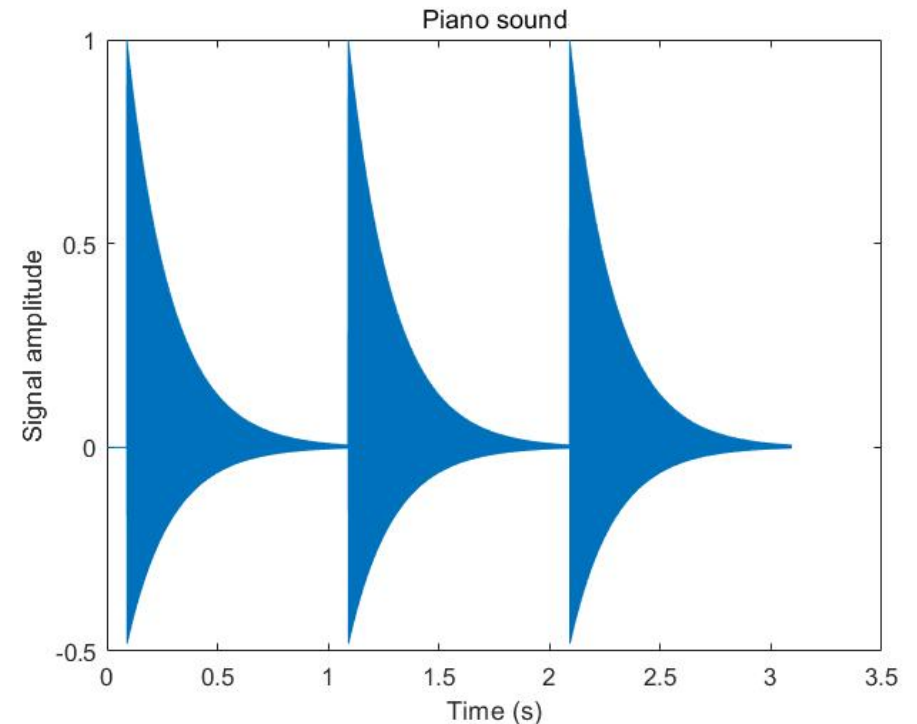
### High order harmonics

- Added by **empirical tuning**
- Just an approximate model. Professional music synthesis may sample and analyze the **real instrumental sound** first
- At fundamental frequency (main frequency)  $f_c$ , the **signal for one note** can be expressed as:

$$y = env(t) \times \left( \cos(2\pi f_c t) + 0.4 \cos\left(4\pi f_c t + \frac{2\pi}{3}\right) + 0.15 \cos\left(6\pi f_c t + \frac{4\pi}{3}\right) \right)$$

Envelop function

sinusoidal signal at  $f_c$  with high order harmonics





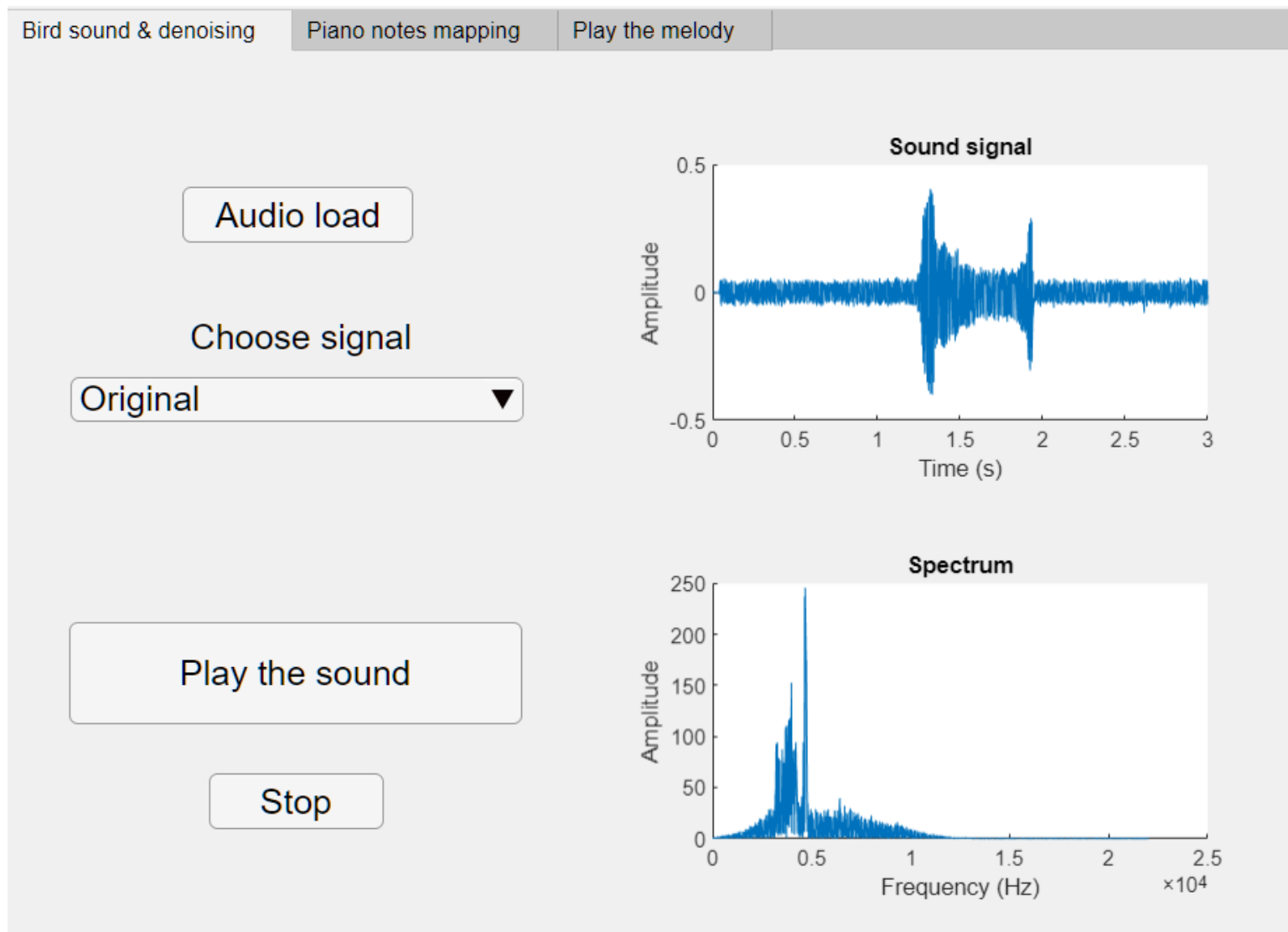
## Stage 2 — Melody Integration & demo

### Melody

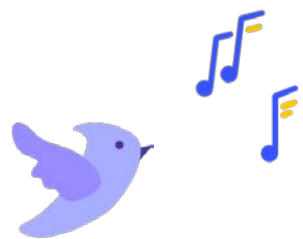
- Connect each single note sound together

### Demo

- MATLAB APP Designer
- Let's play the sound together!







# Q & A





**Thanks For Your Attention!**

