Pitch Detection: Music, Physics, and the Brain

Tom Goodman

University of Birmingham

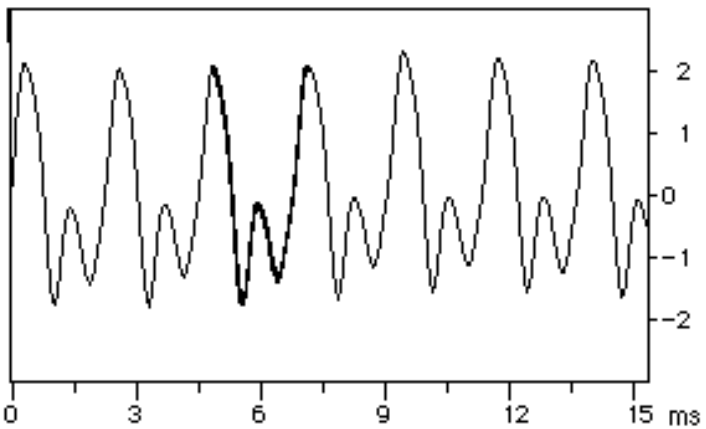
Wednesday 30th January 2018

Contents

Aim

Take in an acoustic musical signal, and output the notes present.

Aim



A4

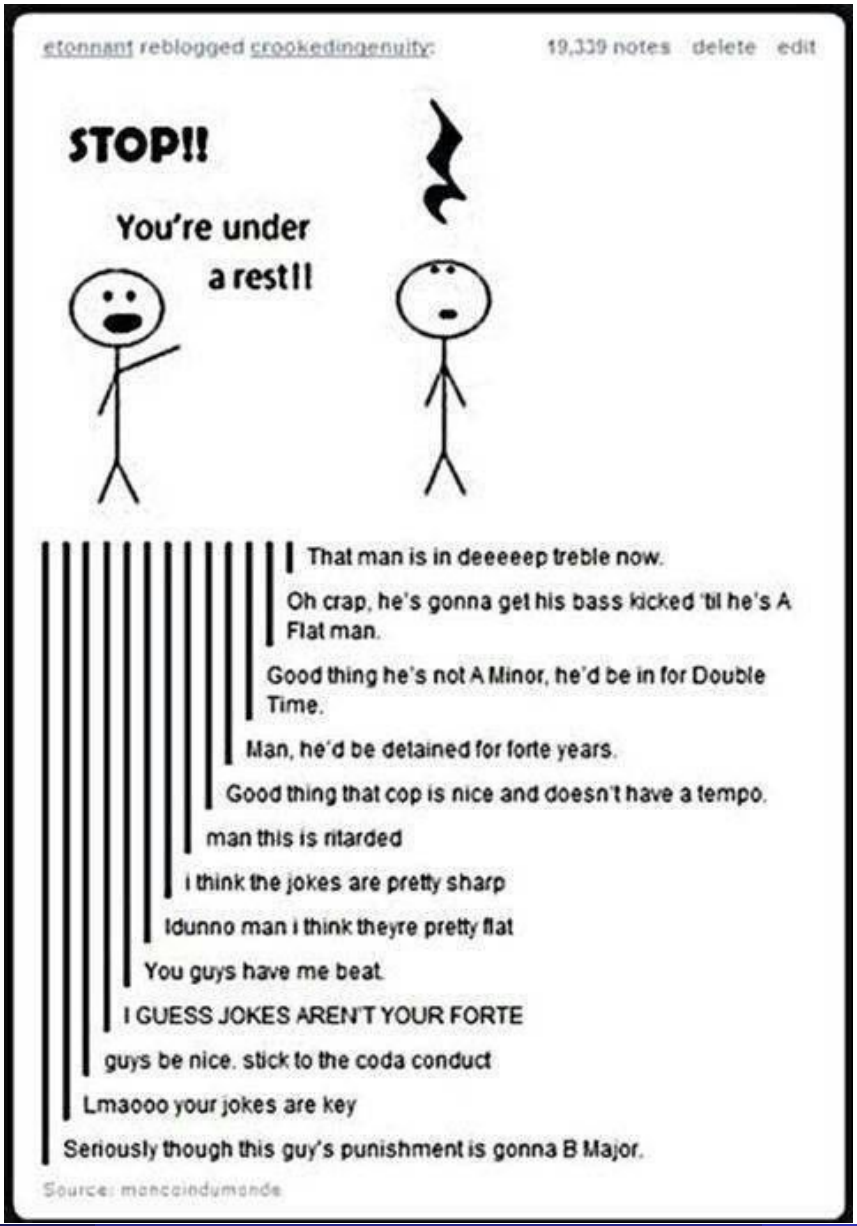
A Little History

history

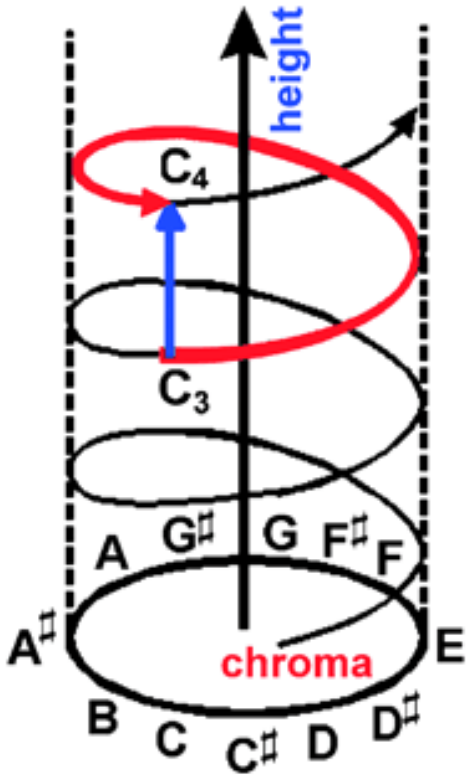
Pythagoras & Early Music (c. 550BC)

Helmholtz & the Rise of Psychoacoustics (1800s)

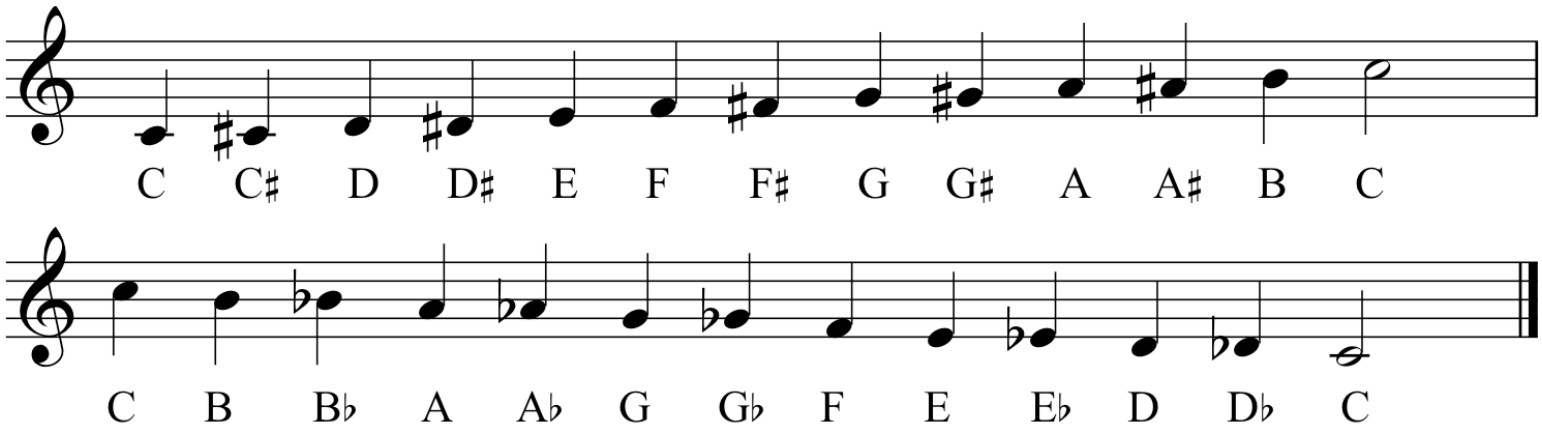
Some Basic Music Theory



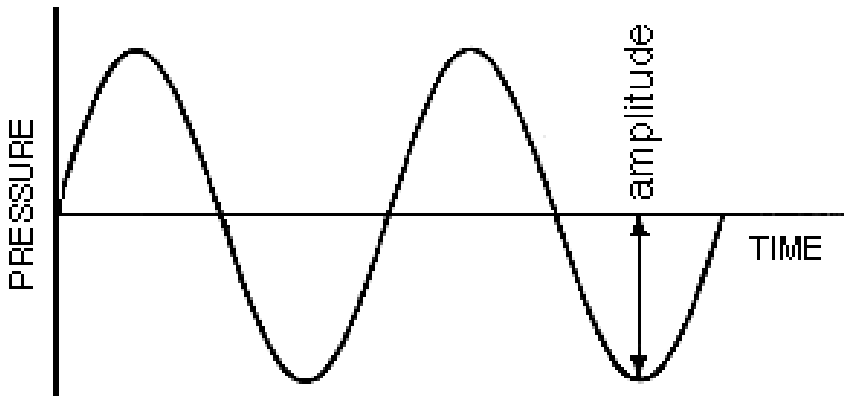
Categorising Notes



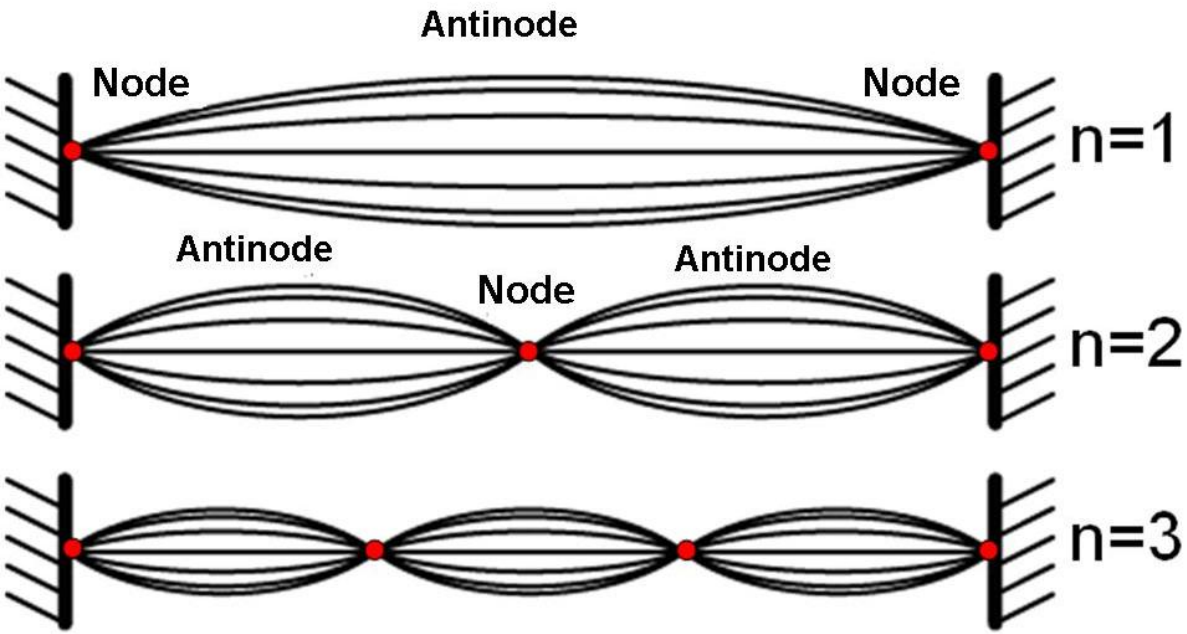
Categorising Notes



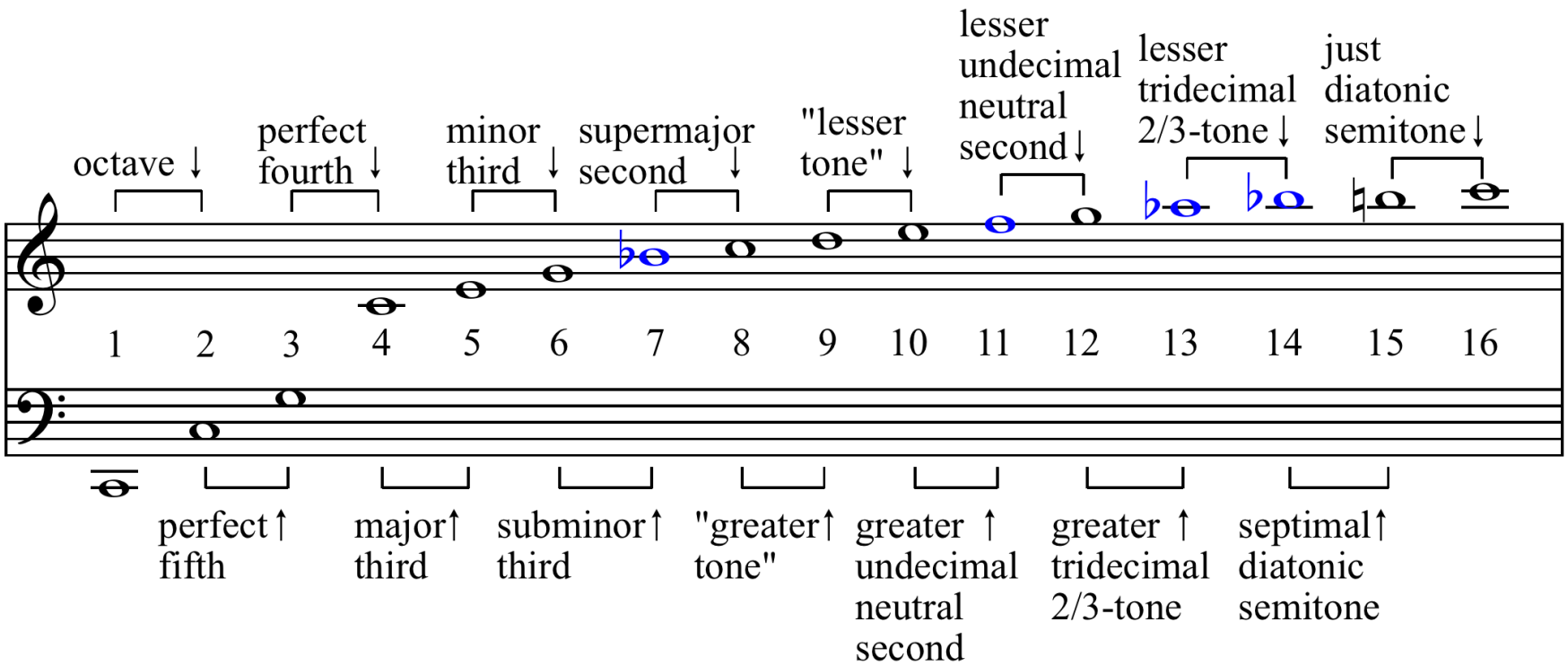
Musical Notes



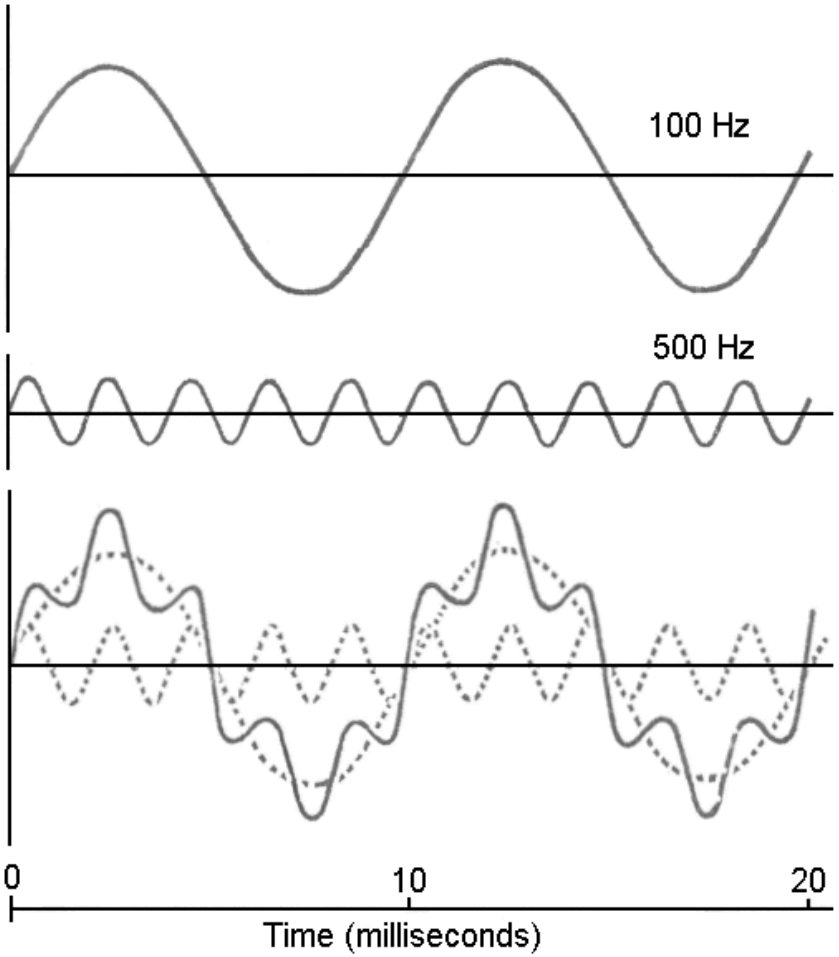
Standing Waves



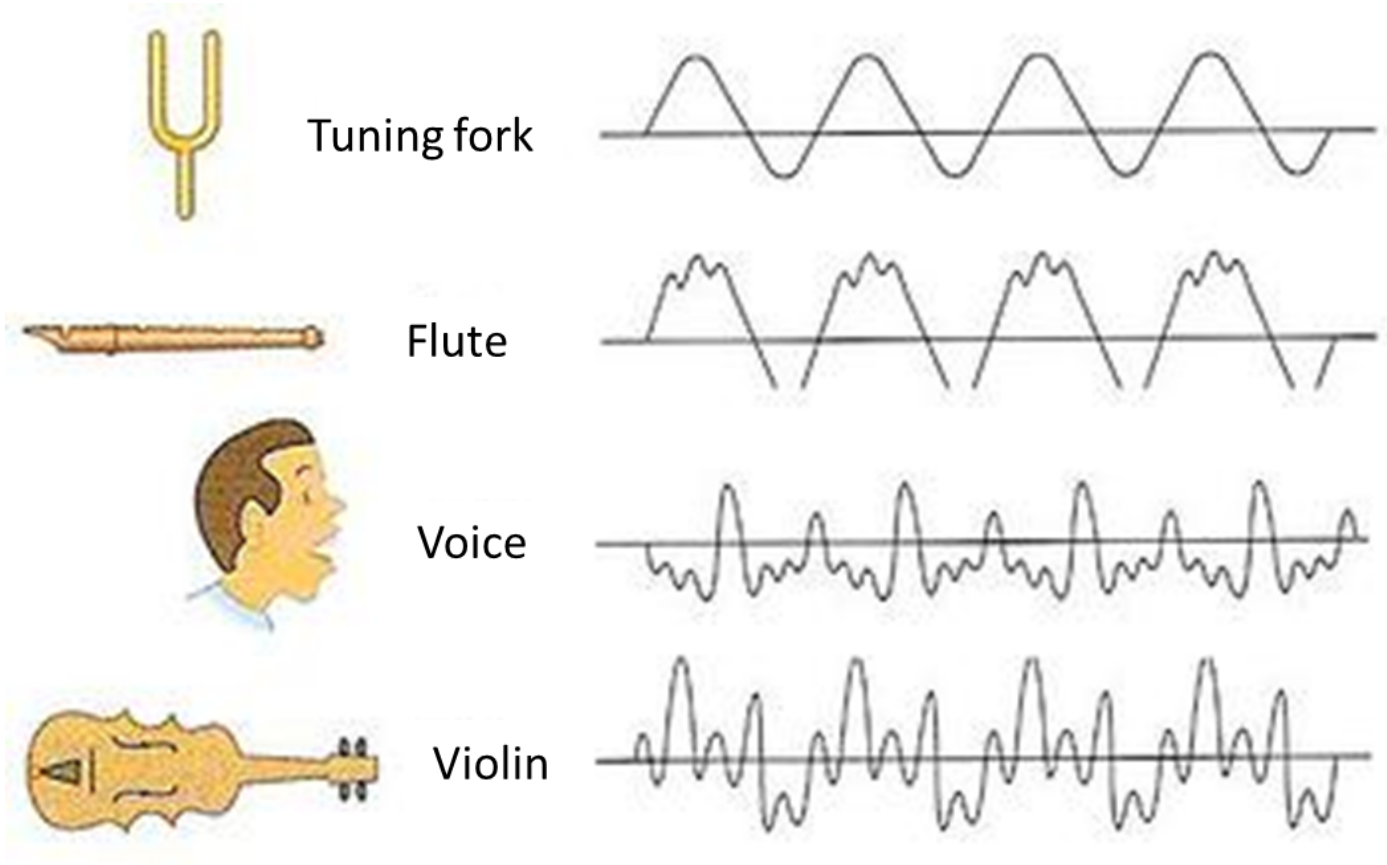
Harmonics


octave ↓
perfect fourth ↓
minor third ↓
supermajor second ↓
"lesser tone" ↓
lesser undecimal neutral second ↓
lesser tridecimal 2/3-tone ↓
just diatonic semitone ↓
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
perfect ↑ fifth
major ↑ third
subminor ↑ third
"greater ↑ tone"
greater ↑ undecimal neutral second
greater ↑ tridecimal 2/3-tone
septimal ↑ diatonic semitone

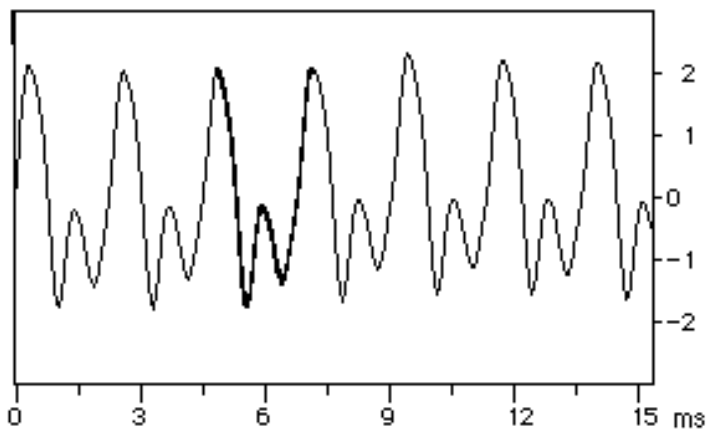
Implications

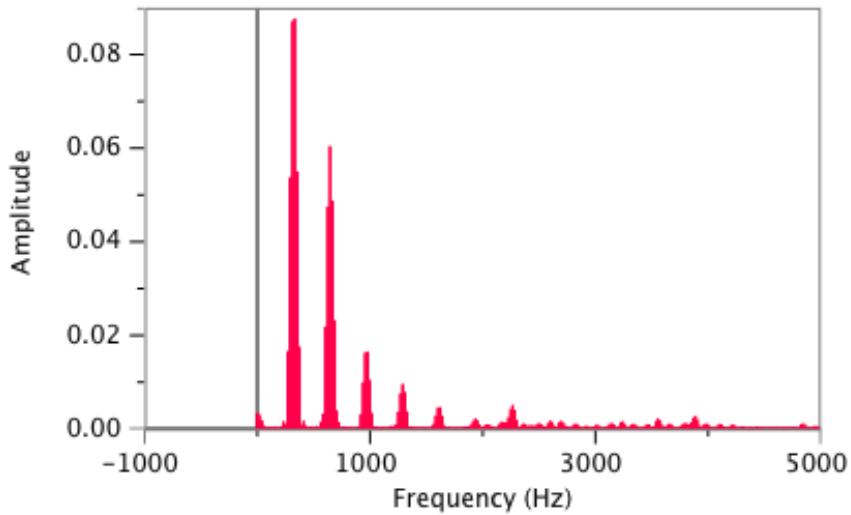


Timbre

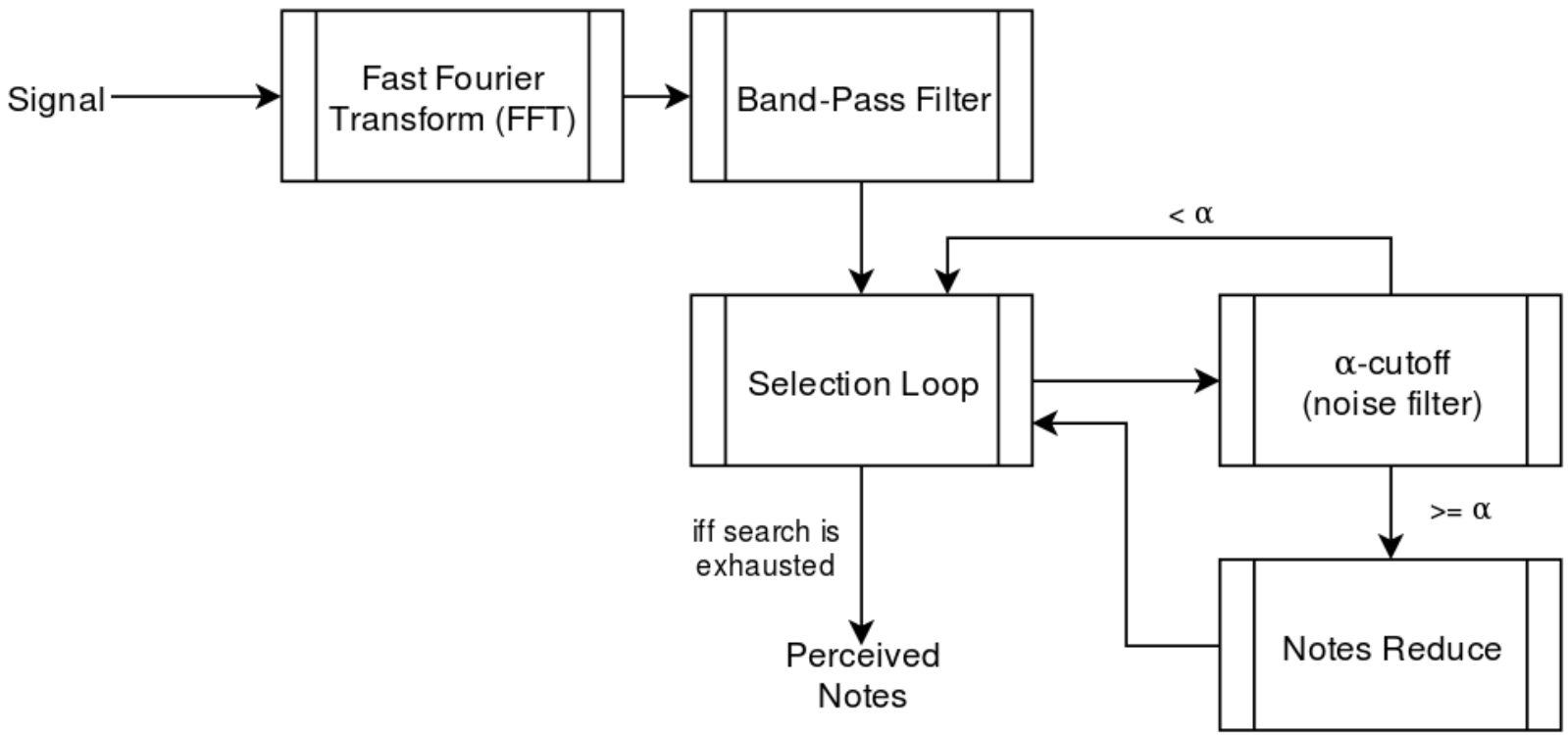


The Fourier Transform



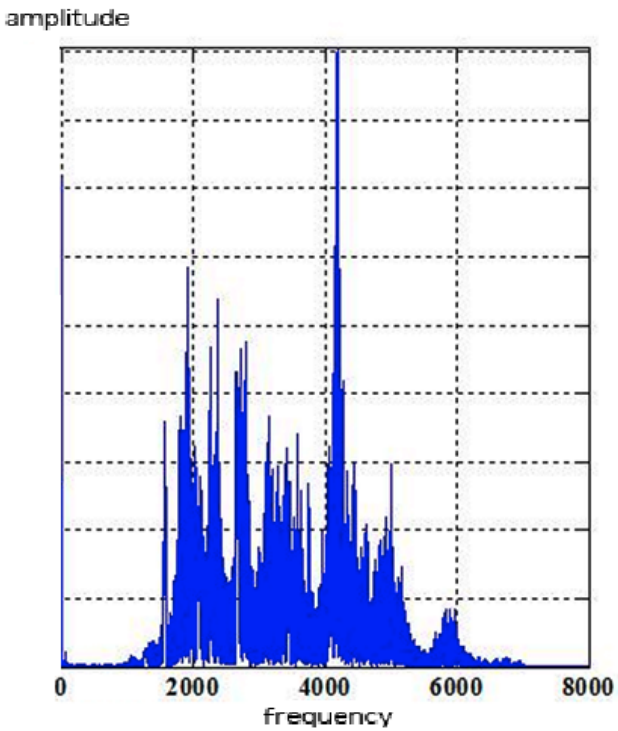


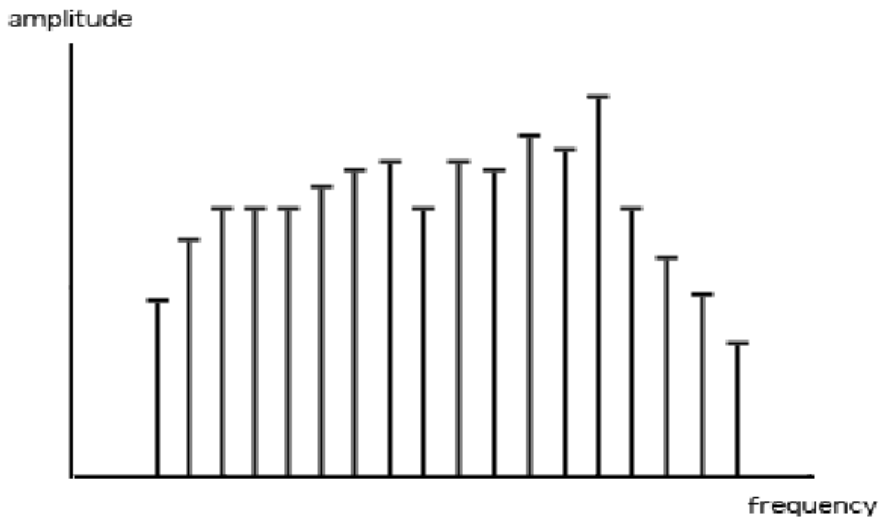
Overview


Signal
Fast Fourier Transform (FFT)
Band-Pass Filter
< α
Selection Loop
α-cutoff (noise filter)
>= α
iff search is exhausted
Perceived Notes
Notes Reduce

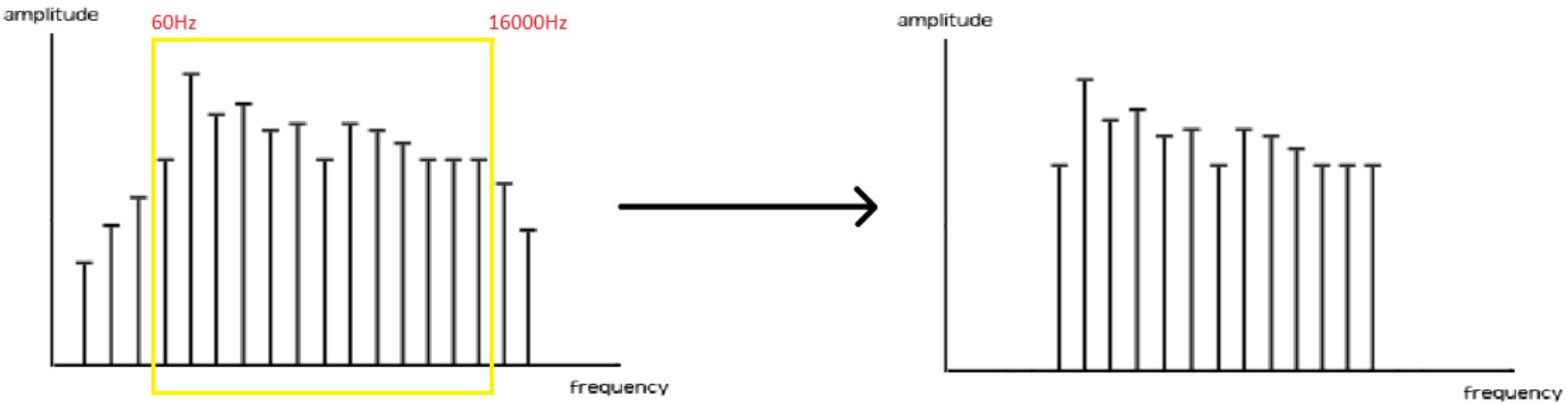
Assumption - No Undertones

Clustering

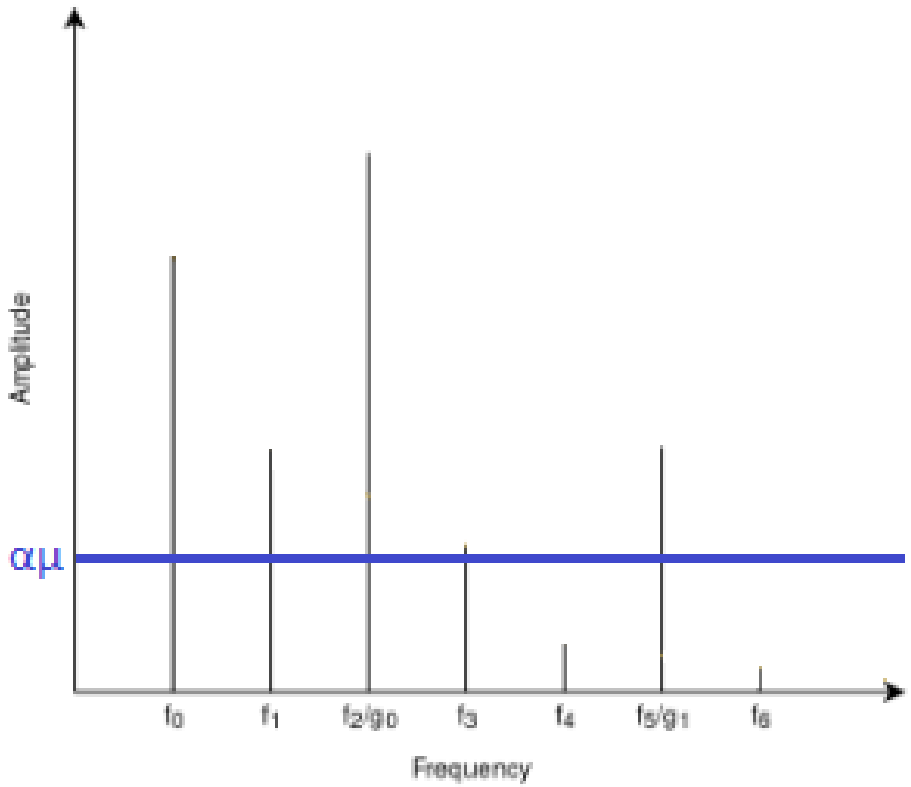

amplitude
0
2000
4000
6000
8000
frequency


amplitude
frequency

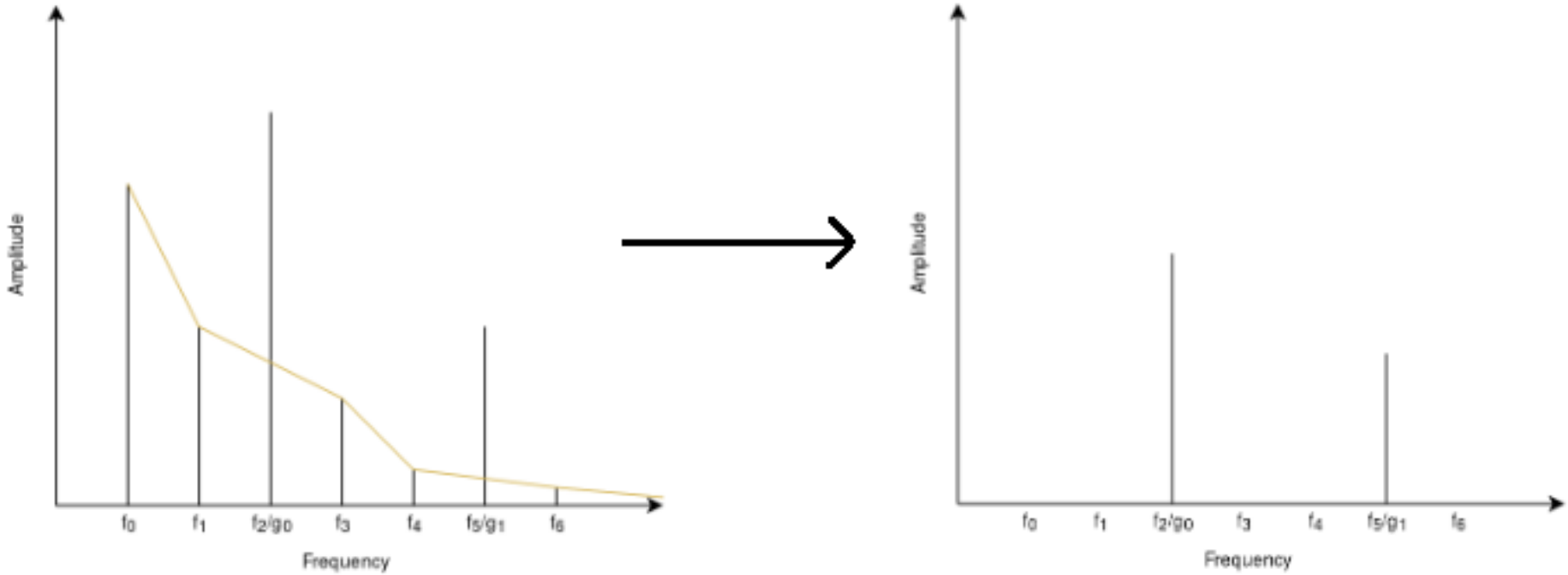
“Band-Pass” Filter



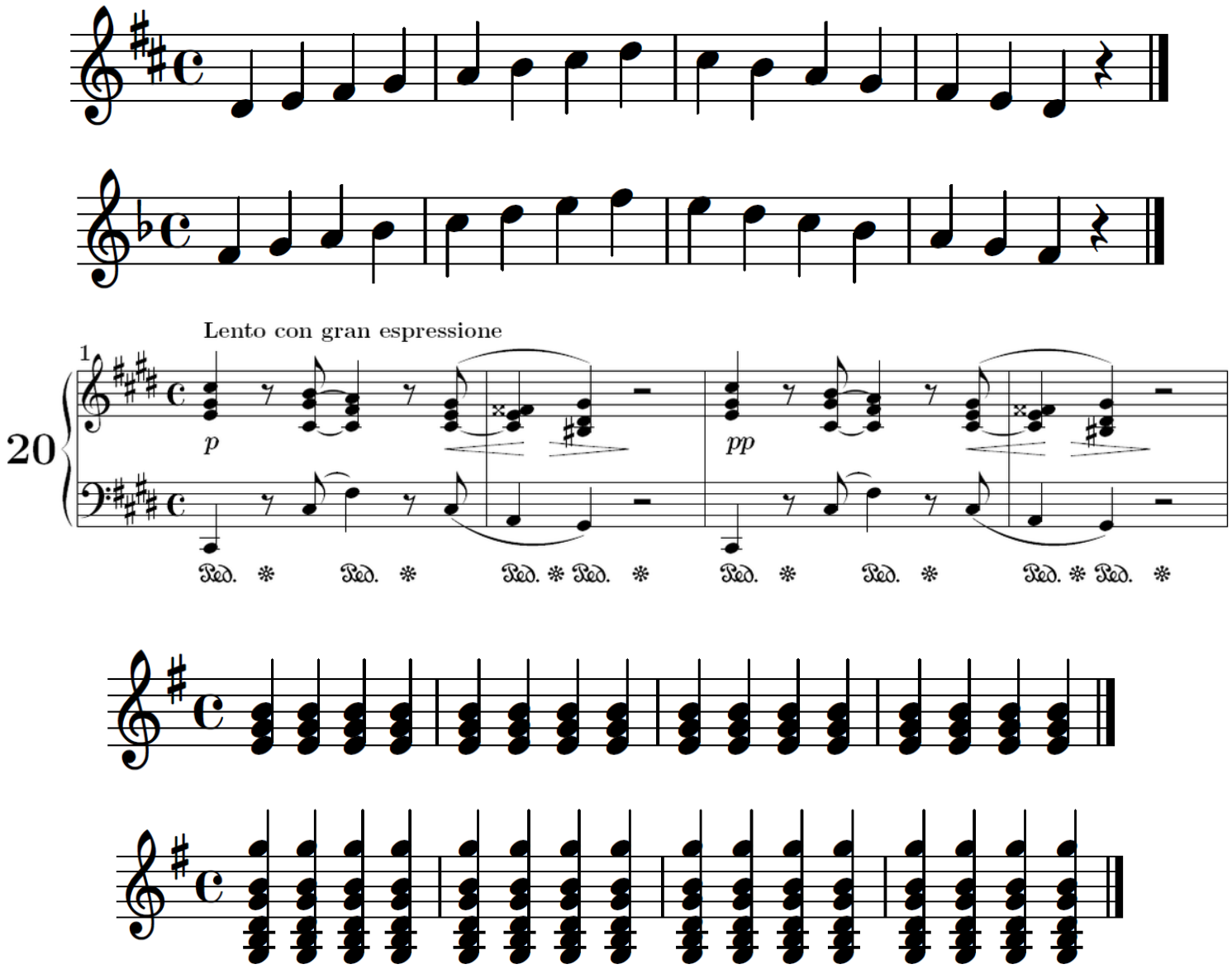
Selection Loop & α-cutoff



Notes Reduce



Testing



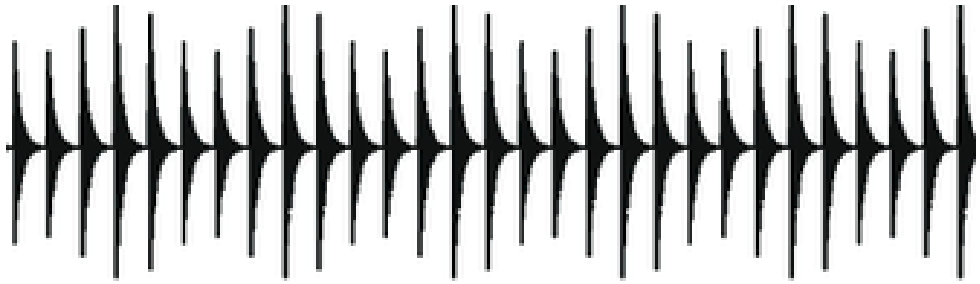
Results

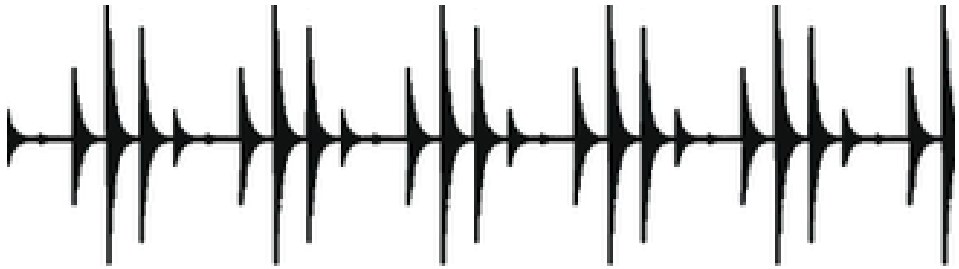
	Precision	**Recall**	**Specificity**	**Accuracy**	**F-Score**
Test 1	78.09%	93.89%	83.00%	86.89%	84.78%
Test 2	100.00%	97.22%	100.00%	97.85%	98.55%
Test 3	82.46%	75.67%	77.52%	76.36%	78.85%
Test 4	86.31%	82.50%	53.33%	75.83%	83.47%
Test 5	65.08%	100.00%	64.29%	79.07%	78.47%
μ	82.39%	89.86%	75.63%	83.20%	84.82%

Further Considerations

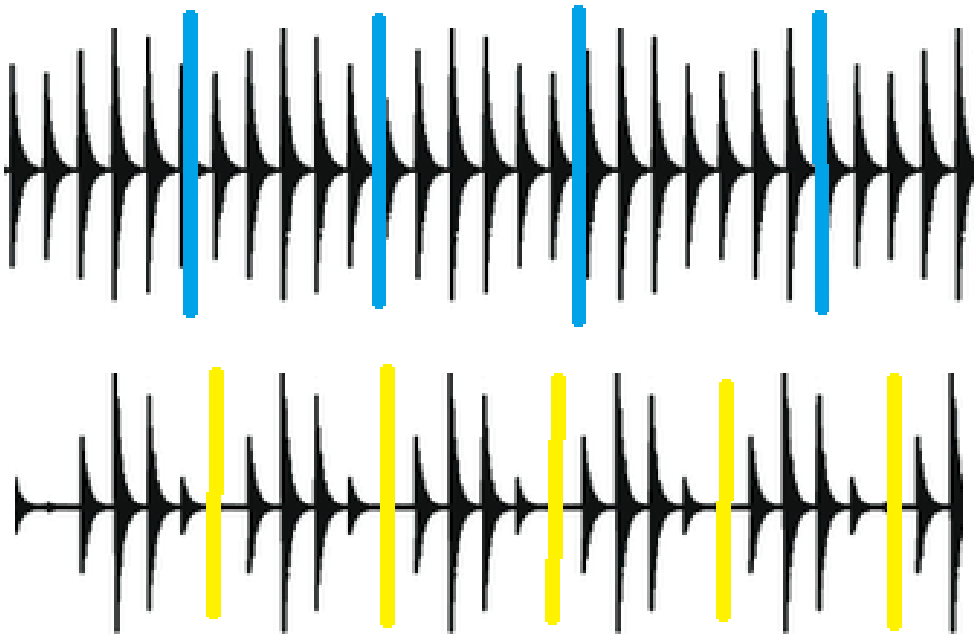
- Adapt evaluation to use a larger & more widely-adopted dataset (eg. MAPS)
- Adapt 'Notes Reduce' to model harmonics on an instrument-to-instrument basis
- Apply source separation to reduce polyphonic problems to multiple monophonic problems
- Take temporal aspects of sound into account
- Take inspiration from biological systems for pitch detection in humans (such as the ear/brain)

A Human Solution

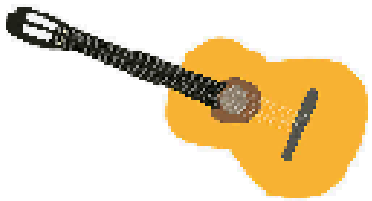


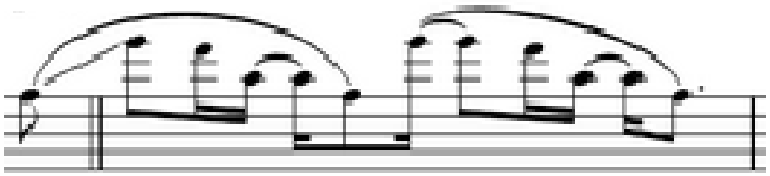


A Human Solution

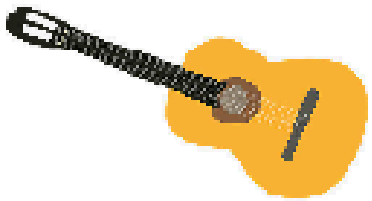


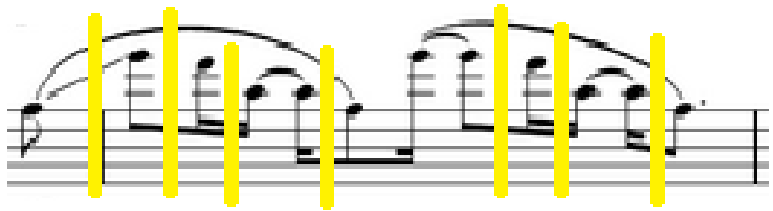
A Human Solution



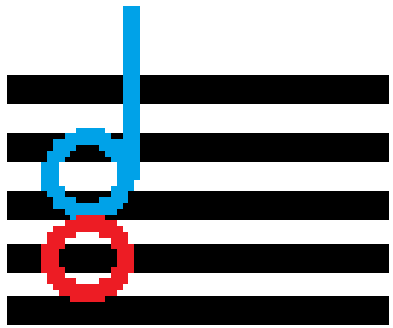


A Human Solution

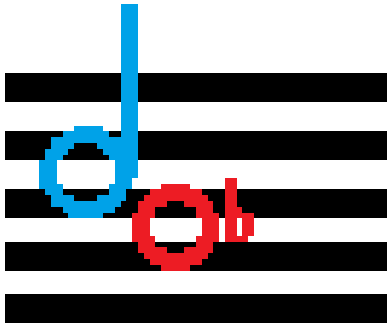




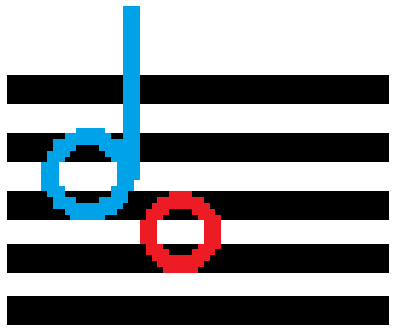
A Human Solution



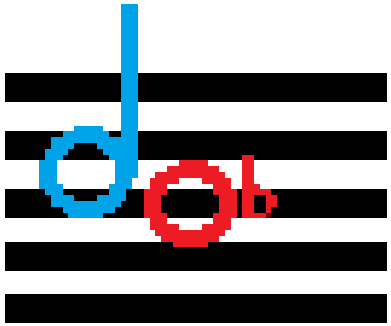
A Human Solution



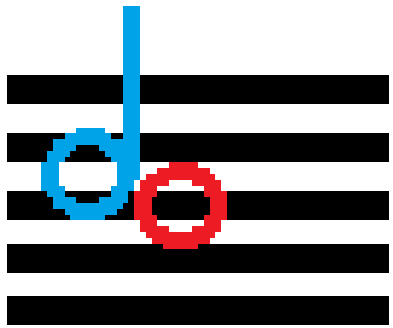
A Human Solution



A Human Solution



A Human Solution



A Human Solution

Some Thoughts for the Future

- Do the ear and brain process monophonic and polyphonic signals in different ways?
- How does the brain subconsciously abstract timbre?
- Can we build models and solutions in similar ways?

Questions?

Tom Goodman
t.a.goodman@cs.bham.ac.uk
http://tomg.io/
Office 218

School of Computer Science
University of Birmingham

Slides at http://tomg.io/research-skills-slides.pdf
Paper at http://tomg/io/isspit-1.pdf

Feel free to ask any questions!