Chromagram Representation for Music Signals

Yichi Zhang

ECE 472 Audio Signal Processing

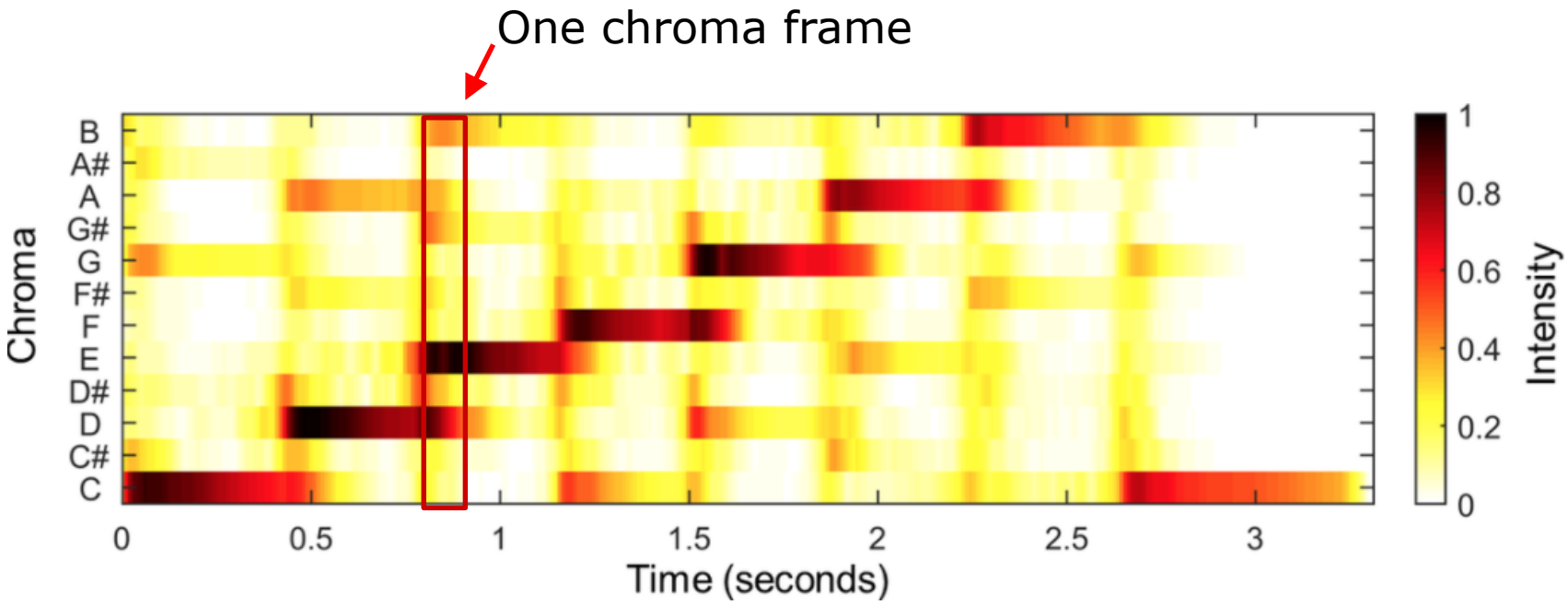
What is chromagram?

One type of feature representation widely used in music signal analysis.

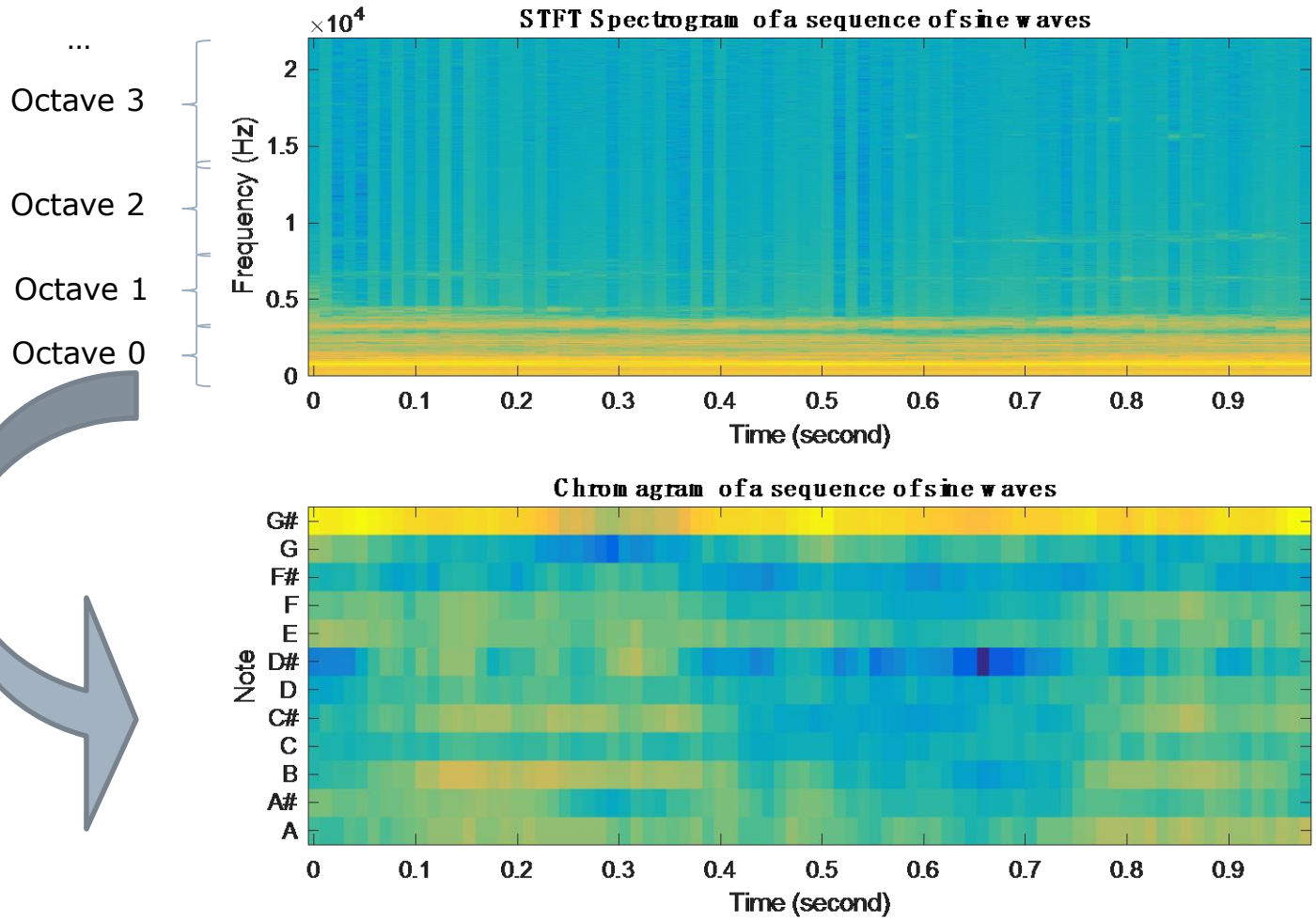
A matrix, similar to STFT spectrogram.

What is chromagram?

The entire spectrum is projected onto 12 bins representing 12 distinct semitones.



What is chromagram?



Equal temperament

The whole spectrum is divided into many octaves

For each octave, we have 12 distinct semitones/keys/notes:

Note	**C4**	**C#4**	**D4**	**D#4**	**E4**	**F4**	**F#4**	**G4**	**G#4**	**A4**	**A#4**	**B4**
Freq (Hz)	261.63	277.18	293.66	311.13	329.63	349.23	369.99	392.00	415.30	440.00	466.16	493.88
Note	**C5**	**C#5**	**D5**	**D#5**	**E5**	**F5**	**F#5**	**G5**	**G#5**	**A5**	**A#5**	**B5**
Freq (Hz)	523.25	554.37	587.33	622.25	659.25	698.46	739.99	783.99	830.61	880.00	932.33	987.77

Note to frequency conversion table:

http://www.phy.mtu.edu/~suits/notefreqs.html

Equal temperament

Frequency of reference pitch

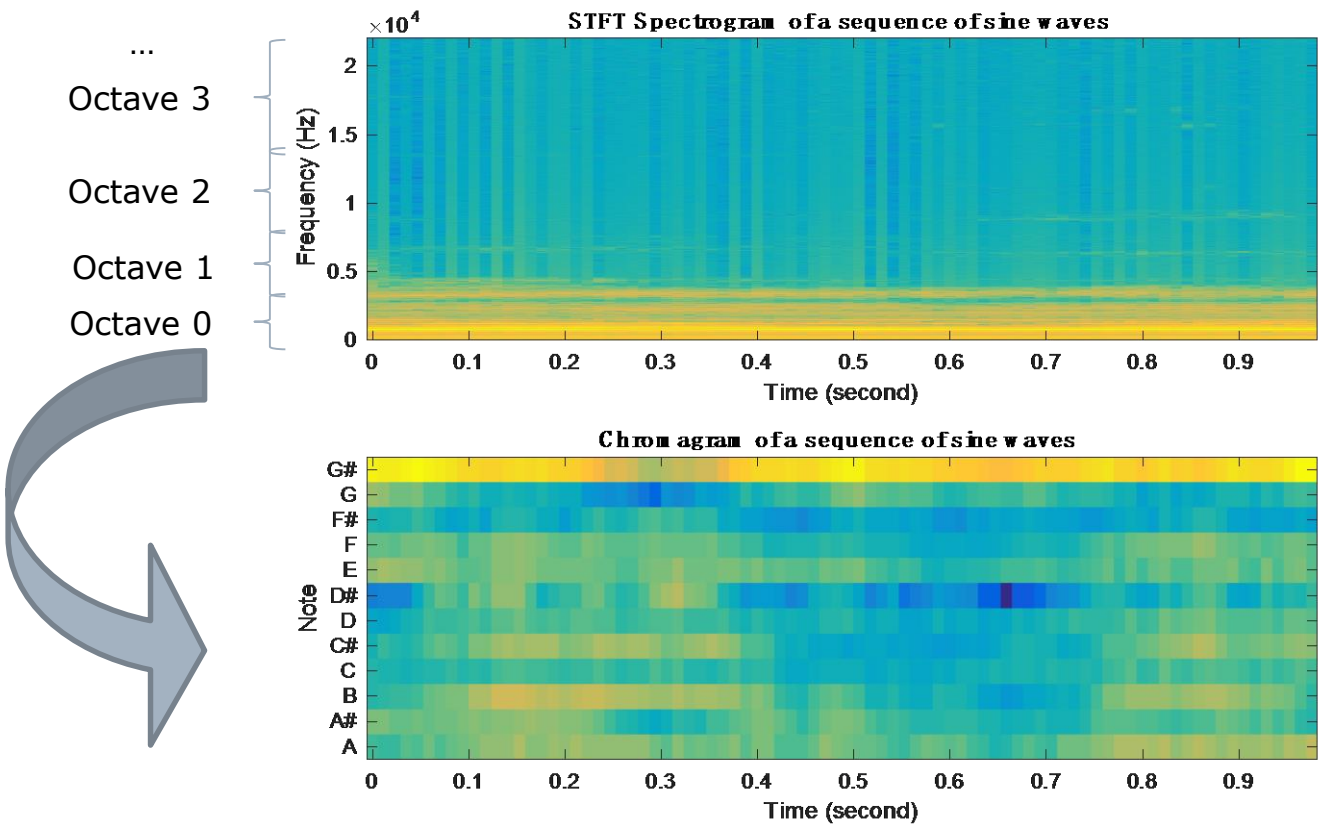
Number of half-steps from the reference pitch

$$f_m = f_{ref} \times 2^{m/12}$$

Frequency of desired pitch

Note	**C4**	**C#4 Db4**	**D4**	**D#4 Eb4**	**E4**	**F4**	**F#4 Gb4**	**G4**	**G#4 Ab4**	**A4**	**A#4 Bb4**	**B4**
Freq (Hz)	261.63	277.18	293.66	311.13	329.63	349.23	369.99	392.00	415.30	440.00	466.16	493.88
Note	**C5**	**C#5 Db5**	**D5**	**D#5 Eb5**	**E5**	**F5**	**F#5 Gb5**	**G5**	**G#5 Ab5**	**A5**	**A#5 Bb5**	**B5**
Freq (Hz)	523.25	554.37	587.33	622.25	659.25	698.46	739.99	783.99	830.61	880.00	932.33	987.77

How to get chromagram?



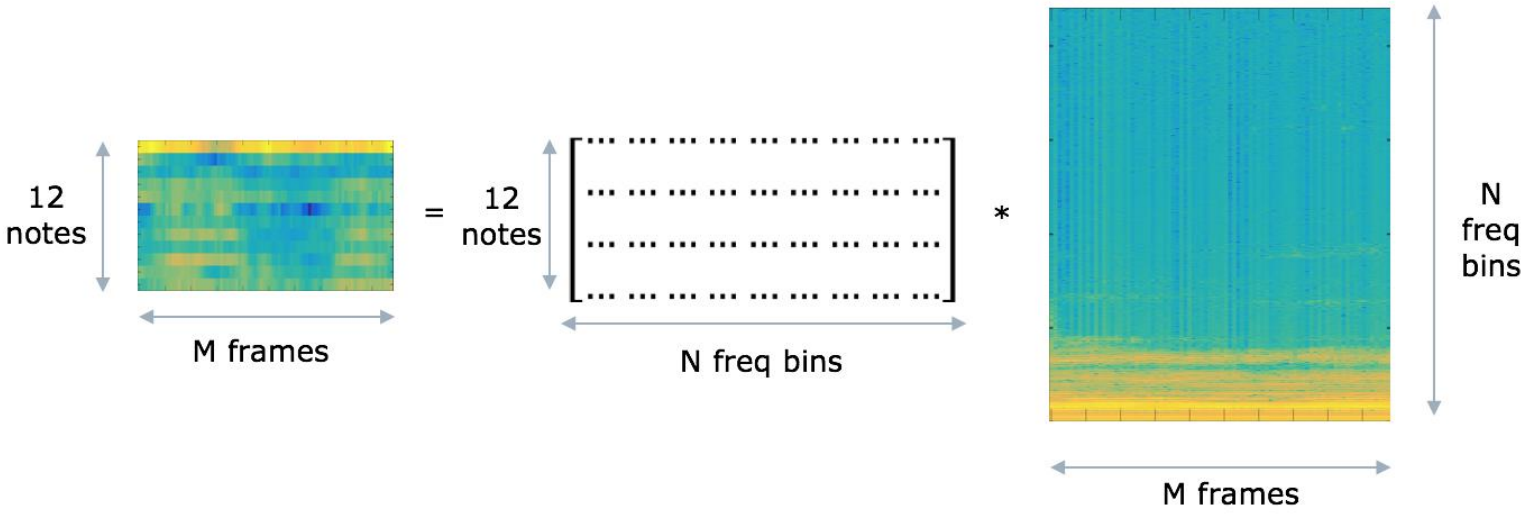
Solution: Find the projection of STFT spectral bins to chroma bins!

How to get chromagram?

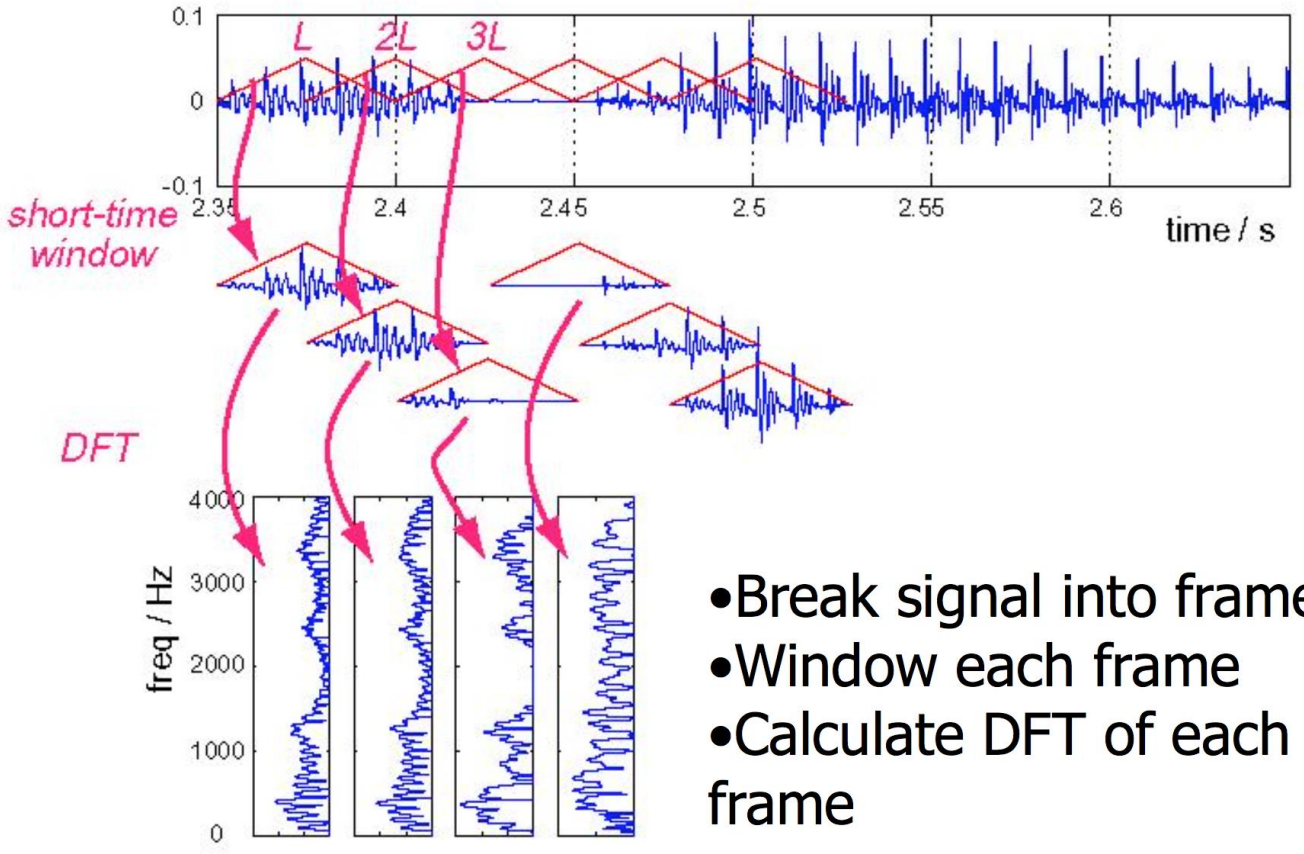
Step 1: Get a STFT spectrogram

Step 2: Get a conversion matrix

Step 3: Multiplication



S1: Get a STFT spectrogram



- Break signal into frames
- Window each frame
- Calculate DFT of each windowed frame

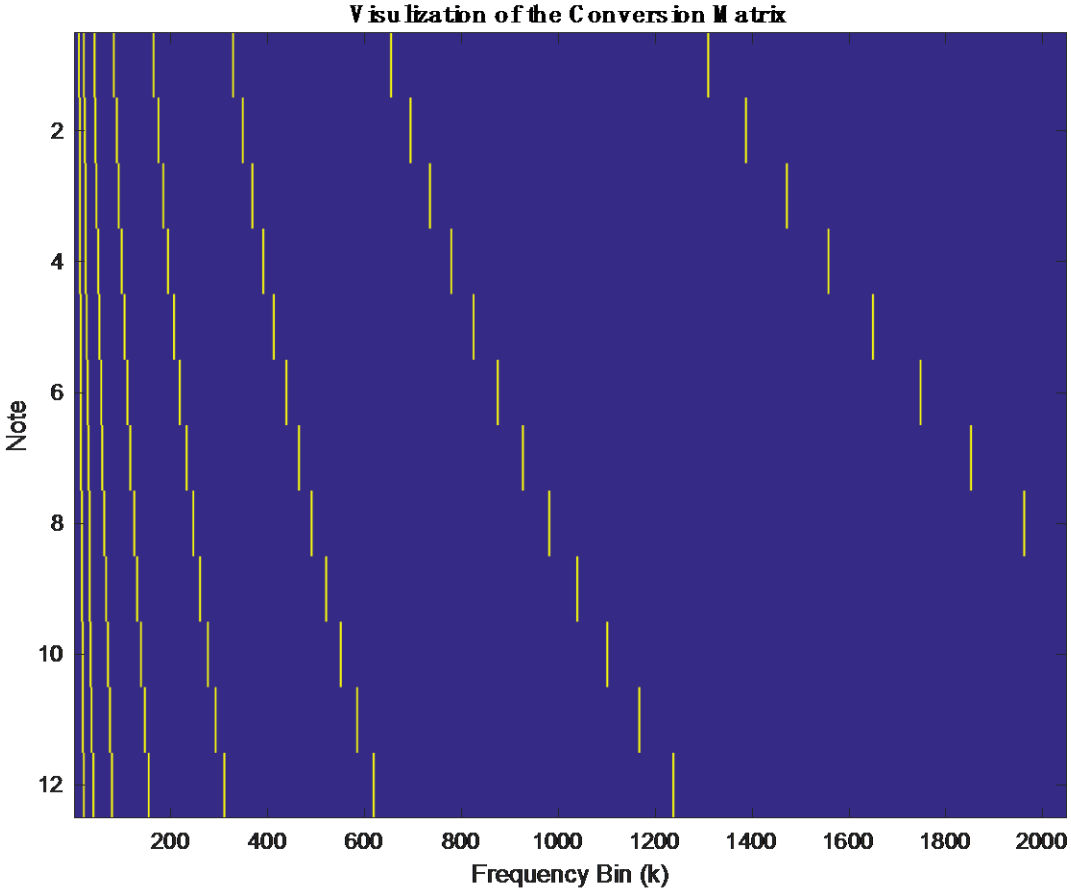
Excerpt from Prof. Duan's Computer Audition course slides

S2: Conversion matrix

Initialization: Choose the freq range: 1) a base note (f_base) for the lowest freq to be covered, 2) highest freq is Fs/2

for octave from 0 to max:
    for note from 1 to 12:
        fk = 2^((note-1)/12)*f_base
        problem: map FFT bin(s) in STFT correspond to fk
    end
end

S2: Conversion matrix



S2: Conversion matrix

Problem:

Higher octaves contain more freq bins (longer length in freq scale), multiple bins will be used to represent one freq value

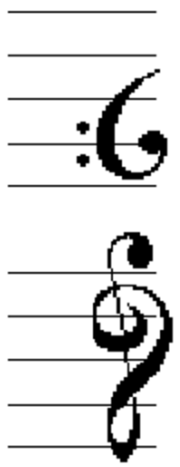
More robust solution:
Convert freq to MIDI number!

S2: Conversion matrix

$$m = 12*\log_2(f_m/440\ \text{Hz}) + 69$$

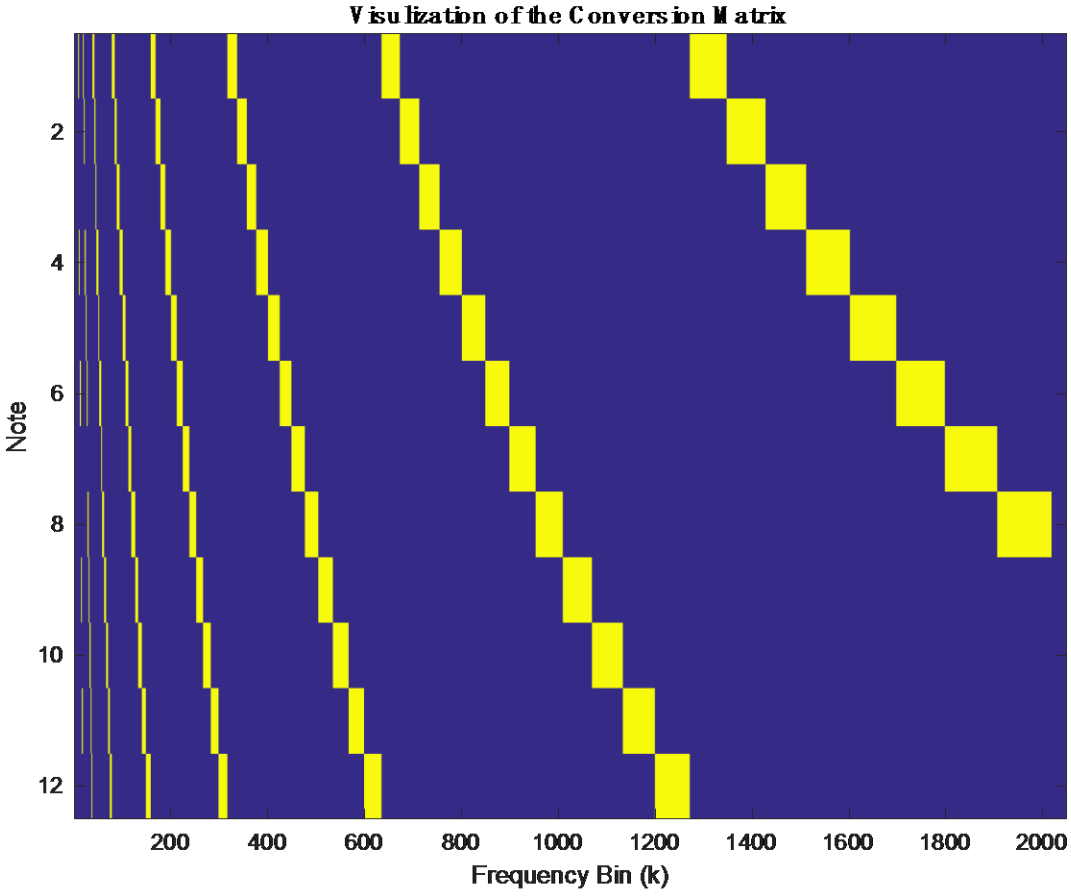
- Each octave has the same length in MIDI scale
- For note C4, MIDI = 60, MIDI = [59.5, 60.5] should also be considered note C4

MIDI number	Note name	Keyboard	Frequency Hz		Period ms	
21	A0		27.500		36.36	
22				29.135		34.32
23	B0		30.868		32.40	
24	C1		32.703		30.58	
25				34.648		28.86
26	D1		36.708		27.24	
27				38.891		25.71
28	E1		41.203		24.27	
29	F1		43.654		22.91	
30				46.249		21.62
31	G1		48.999		20.41	
32				51.913		19.26
33	A1		55.000		18.18	
34				58.270		17.16
35	B1		61.735		16.20	
36	C2		65.406		15.29	
37				69.296		14.29
38	D2		73.416		13.62	
39				77.782		12.86
40	E2		82.407		12.13	
41	F2		87.307		11.45	
42				92.499		10.81
43	G2		97.999		10.20	
44				103.83		9.631
45	A2		110.00		9.091	
46				116.54		8.581
47	B2		123.47		8.099	
48	C3		130.81		7.645	
49				138.59		7.216
50	D3		146.83		6.811	
51				155.56		6.428
52	E3		164.81		6.068	
53	F3		174.61		5.727	
54				185.00		5.405
55	G3		196.00		5.102	
56				207.65		4.816
57	A3		220.00		4.545	
58				233.08		4.290
59	B3		246.94		4.050	
60	**C4**		**261.63**		**3.822**	
61				277.18		3.608
62	D4		293.67		3.405	
63				311.13		3.214
64	E4		329.63		3.034	
65	F4		349.23		2.863	
66				369.99		2.703
67	G4		392.00		2.551	
68				415.30		2.408
69	**A4**		**440.00**		**2.273**	
70				466.16		2.145
71	B4		493.88		2.025	
72	C5		523.25		1.910	
73				554.37		1.804
74	D5		587.33		1.703	
75				622.25		1.607
76	E5		659.26		1.517	
77	F5		698.46		1.432	
78				739.99		1.351
79	G5		783.99		1.276	
80				830.61		1.204
81	A5		880.00		1.136	
82				932.33		1.073
83	B5		987.77		1.012	



http://newt.phys.unsw.edu.au/jw/notes.html

S2: Conversion matrix



Chromagram for chords

Chord: a group of notes (normally three or more) sounded in harmony together

G major chord: root note (G) + major third (B) + perfect fifth (D)

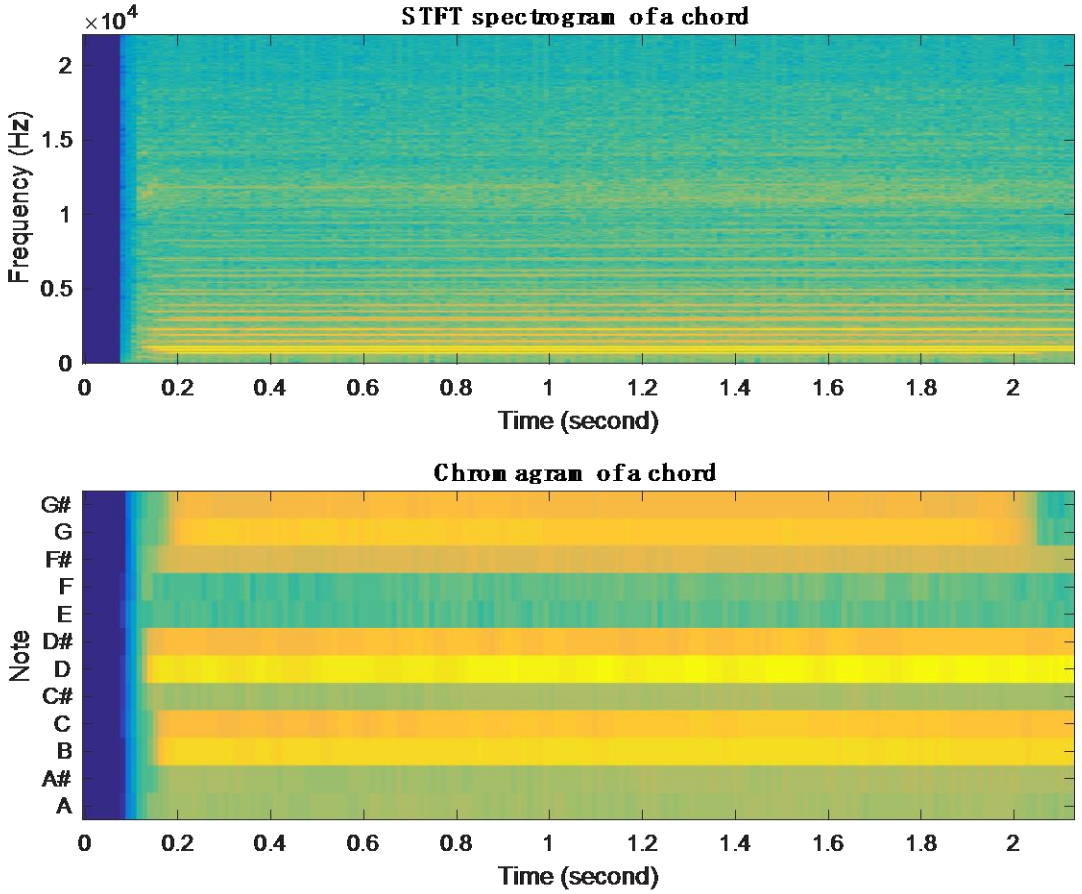
G minor chord: root note (G) + minor third (A#) + perfect fifth (D)

https://en.wikipedia.org/wiki/Major_chord

https://en.wikipedia.org/wiki/Minor_chord

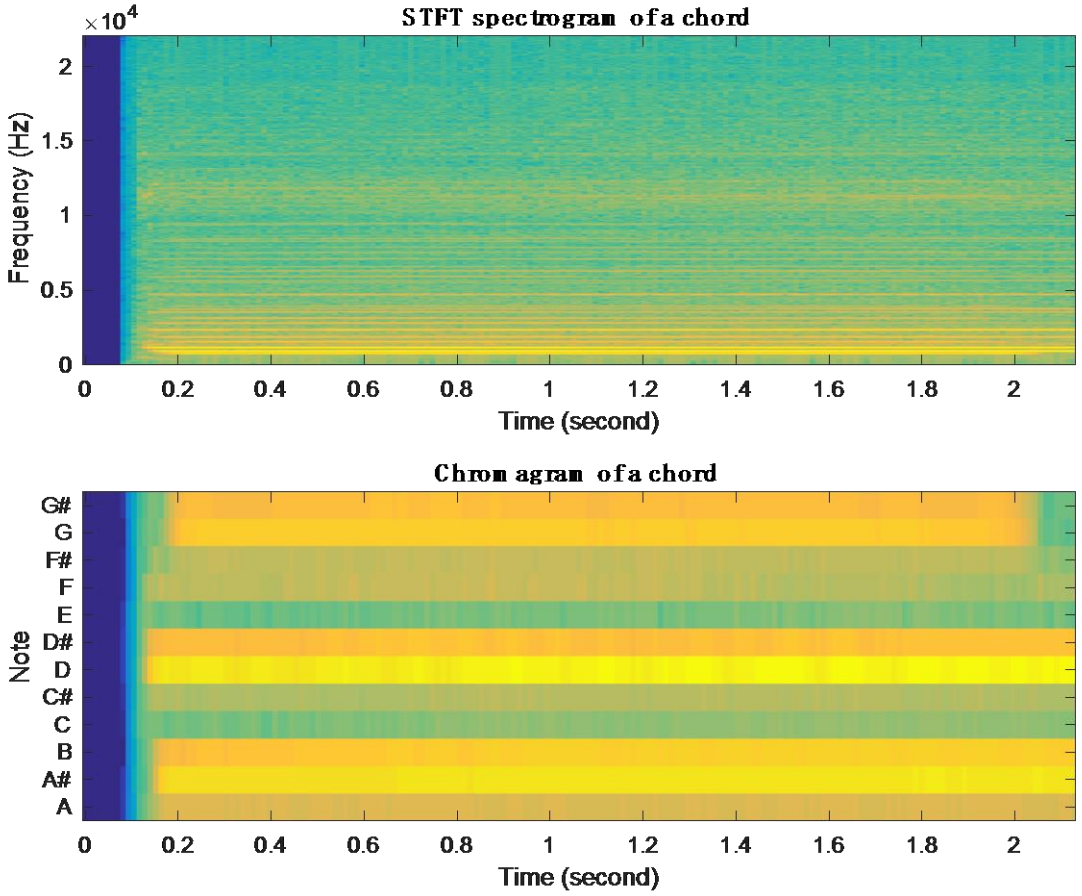
Chromagram for chords

G major = G + B + D



Chromagram for chords

G minor = G + A# + D



The End

Good luck with your HW5 !