Complex Proper gaussian Random Variable – real and imaginary part of a complex random variable follow real gaussian distributions

Spectrogram – STFT magnitude visualization

Multiple Pitch Estimation (MPE) – The task of evaluating all the pitches in a short time frame (10ms)

Hidden Markov Model (HMM) – Markov model where the system being modeled has unobservable states.

Markov model/chain – stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. They satisfy the memoryless Markov property

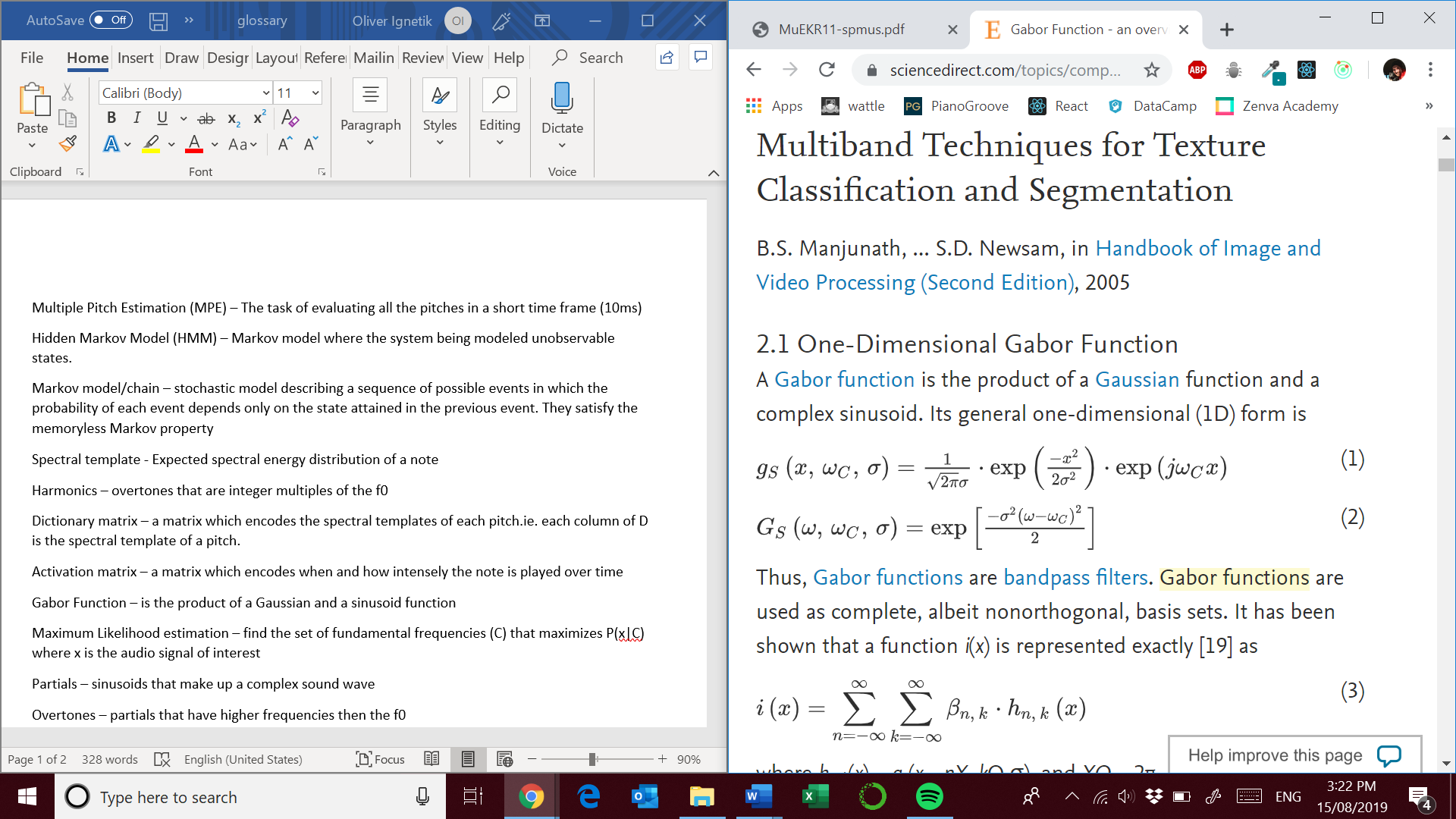
Spectral template - Expected spectral energy distribution of a note

Harmonics – overtones that are integer multiples of the f0

Dictionary matrix – a matrix which encodes the spectral templates of each pitch.ie. each column of D is the spectral template of a pitch.

Activation matrix – a matrix which encodes when and how intensely the note is played over time

Gabor Function – is the product of a Gaussian and a sinusoid function



Source: <https://www.sciencedirect.com/topics/computer-science/gabor-function>

Maximum Likelihood estimation – find the set of fundamental frequencies (C) that maximizes P(x|C) where x is the audio signal of interest

Partials – sinusoids that make up a complex sound wave

Overtones – partials that have higher frequencies then the f0

Harmonicity – the tendency of complex sound waves or instruments that produce perfect harmonics

Mean bandwidth – expectation of the bandwidth

Spectral centroid – a measure using to characterize a spectrum

Synchronicity – can be calculated using cross-correlation and finding the peak is a way to find synchronicity between 2 signals.

F0 – the partial with the lowest frequency

Markov Chain Monte Carlo – a method used to approximate the posterior distribution of a parameter of interest by random sampling in a probabilistic space. The posterior distribution is a combination of the prior distribution ( beliefs ) and the likelihood distribution ( data’s story )

Low order moving average model -an approach that is used to model univariate time series. The output variable will depend on the current and various past values of a stochastic term.

Spectral envelope of overtones – The envelope curve of the overtone’s amplitude in a frame

Gaussian interference sources – sources which introduce gaussian distributed noise

Narrowband spectra – spectra in which the bandwidth is narrow. For example, an harmonic overtone is considered a to have a narrowband spectra in isolation.

Nonnegative Matrix Factorization (NMF) – method for representing the audio signal as the matrix product of an activation matrix and a dictionary matrix

Probabilistic latent component analysis ( PCLA ) – a form of NMF that looks to maximize independence similarly to Independent Component Analysis ( ICA ). It is an Expectation-Maximization algorithm. It models a mixture of independent distributions latent within the data.

Source : <http://web.mit.edu/~punk/Public/AudioExtraction/PLCApage.html>

Bivariate probability distribution – probability that an event will occur when there are two independent random variables in your scenario

Eigeninstruments – linear combinations of basic instrument models represent fixed spectral templates

Log-frequency – expressing frequency in a logarithmic scale to make patterns more interpretable

Sparse coding – sources are non-active most of the time

Posteriogram – Represent the outputs based on some posteriors. ie. Sound classification events over time graph ( could be pitches or letters spoken )

Bayesian network – Probabilistic graphical model that uses Bayesian inference for probability computations.

Onset detection function – quantifies the amount of change in the signal properties from frame to frame

Bandwise magnitude – average calculation of the spectral magnitude within a bandwidth

Bandwise phase – average calculation of the spectral phase within a bandwidth

Spectral flux – how quickly is the power spectrum of the signal changing

Phase deviation – The peak difference between the instantaneous angle of the of a modulated wave and the angle of the carrier wave.

Complex domain detection functions – Generates detection functions that are sharp at the point of onsets and smooth elsewhere

Tempo – rate of the most salient pulse

Comb filter banks – bank of comb filters. A comb filter is a filter implemented by adding a delayed version of a signal to itself

Inter-onset interval histograms – histogram of frequency of different intervals between successive beat positions

Adaptive oscillators – oscillators that can learn the frequency of an input signal. The frequency of the oscillator will adapt to the frequency of any periodic input signal

Gain matrix – Matrix that represents the gain in a system

Chroma – octave degree of a symbol is known as the pitch’s chroma. For example, a particular pitch can be denoted by F#4 where F# is the chroma and 4 is the octave number.

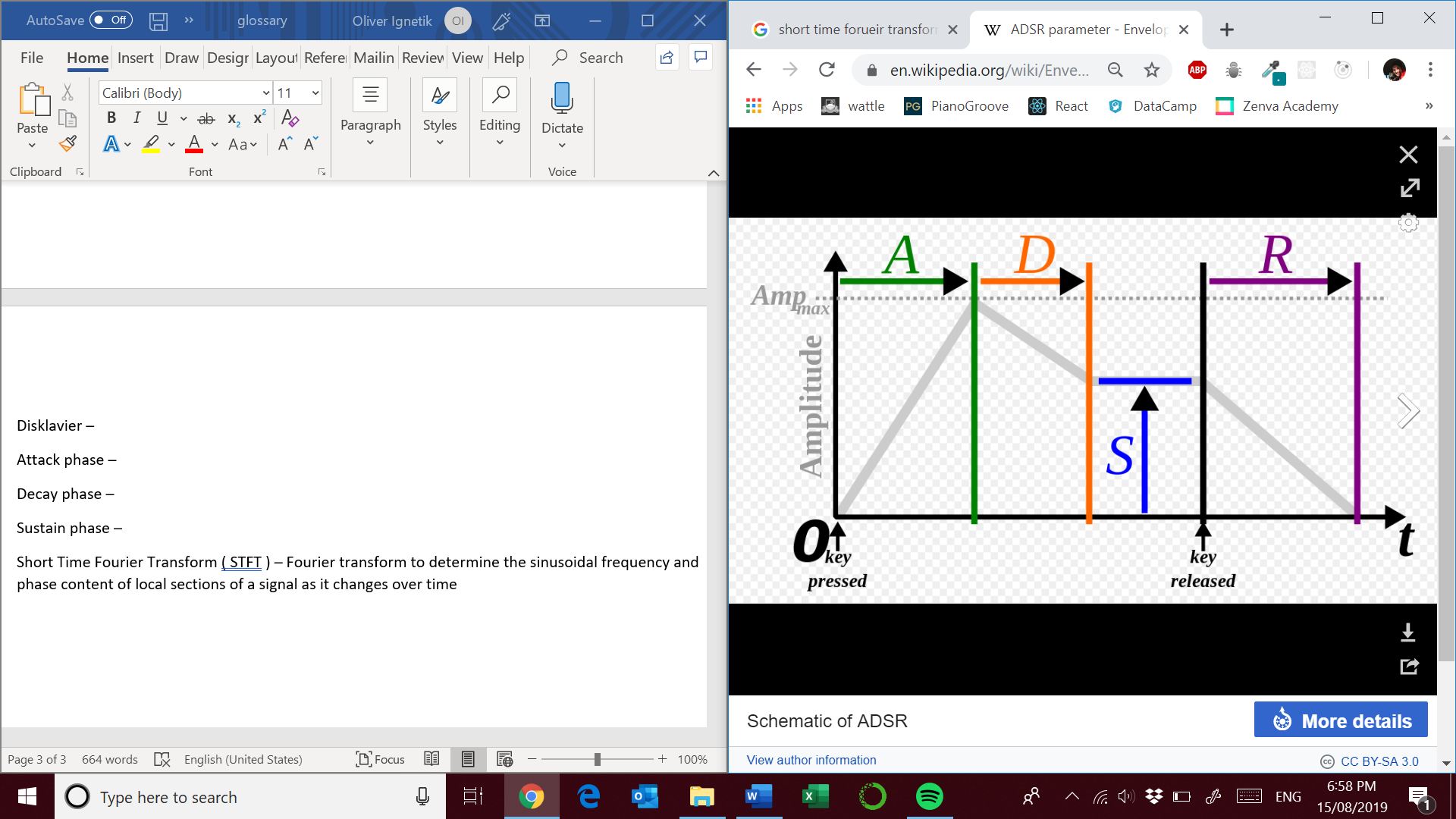
Gaussian Scaled mixture models – a probabilistic model which states that all generated data points are derived from a mixture of a finite Gaussian distributions that has no known parameters.

Disklavier – A piano/ keyboard that can play a library of music

Attack phase – time taken for the initial run-up

Decay phase – time taken from attack level peak to the sustain level

Sustain phase – main sequence of sound’s duration until release



Short Time Fourier Transform ( STFT ) – Fourier transform to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time

Mel spectrogram – mel-frequency cepstrum is a representation of the short term power spectrum of a sound based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. A spectrogram with a Mel Frequency Scale <https://www.mathworks.com/help/audio/ref/melspectrogram.html>. This scale is useful for representing the frequency resolution of the human auditory system.