Audio signal processing

Erik Pärlstrand

May 2018

1 Introduction to discrete forier transforms

Real sinusoid

A real sinusoid in a discrete time domain can be expressed by:

$$x[n] = Acos(2\pi fnT + \phi)$$

where x is the array of real values of the sinusoid, n is an integer value expressing the time index, A is the amplitude value of the sinusoid, f is the frequency value of the sinusoid in Hz, T is the sampling period equal to $1/f_s$, f_s is the sampling frequency in Hz, and ϕ is the initial phase of the sinusoid in radians.

Complex sinusoid

A complex sinusoid in a discrete time domain can be expressed by:

$$\bar{x}[n] = Ae^{j(\omega nT+\phi)} = Acos(\omega nT + \phi) + jAsin(\omega nT + \phi)$$

where $\bar{x}$ is the array of complex values of the sinusoid, n is an integer value expressing the time index, A is the amplitude value of the sinusoid, ω is the frequency of the sinusoid in radians per second (equal to $2\pi f$), T is the sampling period equal $1/f_s$, f_s is the sampling frequency in Hz and ϕ is the initial phase of the sinusoid in radians.

Discrete Fourier Transform (DFT)

The N point DFT of a sequence of real values x (e.g. a sound) can be expressed by:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}$$

where n is an integer value expressing the discrete time index, k is an integer value expressing the discrete frequency index, and N is the length of the DFT.

Inverse Discrete Fourier Transform (IDFT)

The IDFT of a spectrum X of length N can be expressed by:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k]e^{j2\pi kn/N}$$

where $n = 0, ..., N-1$. Here, n is an integer value expressing the discrete time index, k is an integer value expressing the discrete frequency index, and N is the length of the spectrum X.

Magnitude spectrum

The magnitude of a complex spectrum X is obtained by taking its absolute value: $|X[k]|$. Sometimes the magnitude is expressed in a decibel scale: $20 \cdot log_{10}|X[k]|)$.

2 Fourier Properties

Convolution and filtering

The convolution operation between two series x_1 and x_2 is defined as:

$$(x_1 * x_2)[n] = \sum_{m=-\infty}^{\infty} x_1[m]x_2[n-m]$$

In signal processing, convolution is a mathematical operation that can be used for filtering. Filtering involves selectively suppressing certain frequencies present in the signal. Filtering is often performed in the time domain by the convolution of the input signal with the impulse response of a filter. However, the DFT has the following property:

Convolution in the time domain results in multiplication in frequency domain, i.e. $x_1[n] * x_2[n] \implies X_1[k]X_2[2]$.

Therefore, the same operation can also be done in the DFT domain using this properties, by multiplying the DFT of the input signal by the DFT of the impulse response of the filter.

Let's consider an example of filtering, namely linear filtering:

An example of such a case is if a signal f contains some noise and one wants to generate an average of f over a certain period in time. One can then choose a signal g of an appropriate shape and let it slide over f. The resulting signal h then illustrates an average of f during a selected number of time units. This gives a clearer signal where insignificant deviations have been removed.

Zero-padding

Zero-padding a signal is done by adding zeros at the end of the signal. If we perform zero-padding to a signal before computing its DFT, the resulting spectrum will be an interpolated version of the spectrum of the original signal. In most implementations of the DFT (including the FFT algorithms) when the DFT size is larger than the length of the signal, zero-padding is implicitly done.

Zero phase windowing

Zero phase windowing of a frame of signal puts the centre of the signal at the "zero" time index for DFT computation. By moving the centre of the frame to zero index by a circular shift, the computed DFT will not have the phase offset which would have otherwise been introduced. This is caused by the following property of the DFT:

A shift causes the DFT to be multiplied by a complex exponential, i.e.

$$x[n - n_0] = e^{-j2\pi k n_0 / N} X[k]$$

When used in conjunction with zero-padding, zero phase windowing is also useful for the creation of a frame of length of power of 2 for FFT computation.

Real, even and odd signals

A signal is real when it does not have any imaginary component, and all sounds are real signals. A signal x is even if $x[n] = x[-n]$, and odd if $x[n] = -x[-n]$. The DFT has the following properties:

If $x[n]$ real $\implies |X[k]|$ even and $< X[k]$ odd ($<$ represents angle).

If $x[n]$ even and real $\implies |X[k]|$ even and $< X[k] = n\pi$.

Additional properties

It's linear, i.e. $ax_1[n] + bx_2[n] = aX[k] + bX[k]$.