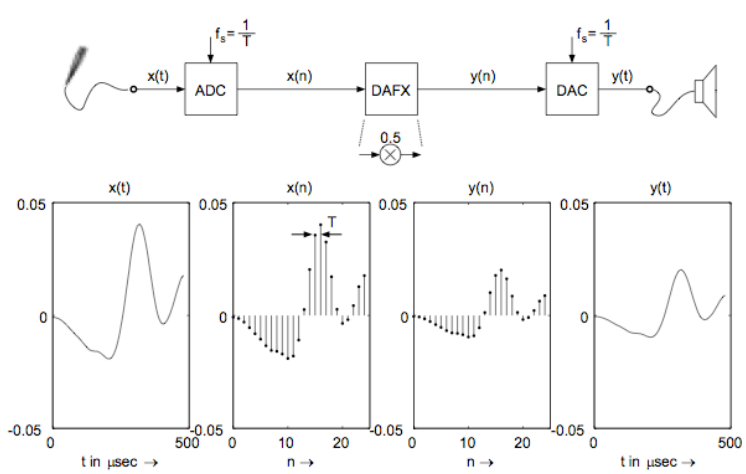
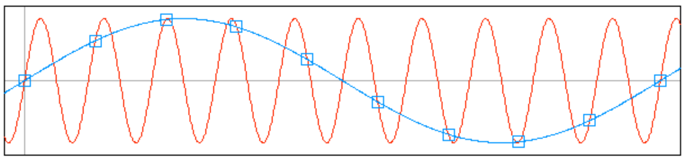
Fundamentals of DSP for music analysis

Digital Signals

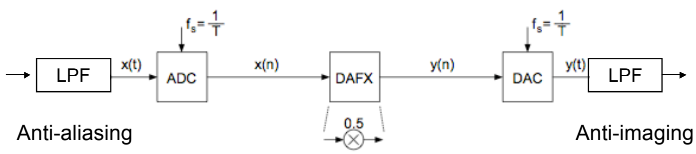


Digital Signals

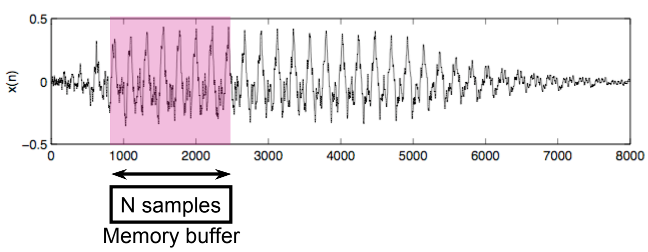
- According to the sampling theorem: $fs > 2fmax$
- Otherwise there is another, lower-frequency, signal that share samples with the original signal (aliasing).



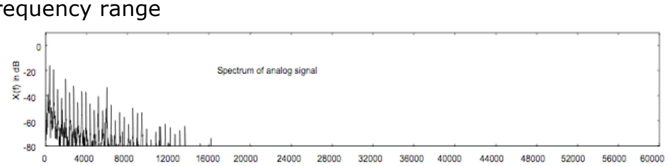
- Related to the wagon-wheel effect: http://www.michaelbach.de/ot/mot_strob/index.html



Block processing and spectrum



- For Block processing, signal data is sent to a buffer and processed as a block. The buffer is then filled with new data.
- A common example is spectral analysis using the DFT.
- The spectrum of a signal's segment shows the energy distribution across the frequency range



Discrete Fourier Transform

- The spectrum of a digital signal, *x(n)*, can be calculated as:

$$X(k) = DFT[x(n)] = \sum_{n=0}^{N-1} x(n) e^{-j2\pi nk/N}$$
$$k = 0,1,\ldots,N-1$$

- The resulting N samples *X(k)* are complex-valued:

$$X(k) = X_R(k) + jX_I(k)$$
$$\left|X(k)\right| = \sqrt{X_R^2(k) + X_I^2(k)}$$
$$\varphi(k) = \arctan\frac{X_I(k)}{X_R(k)}$$
$$k = 0,1,\ldots,N-1$$

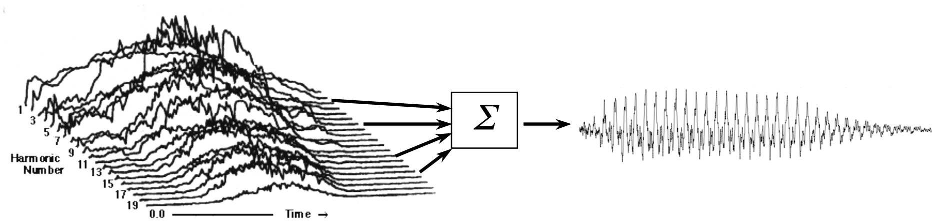
Inverse DFT (IDFT)

- The IDFT allows for the transformation of spectra in discrete frequency to signal in discrete time.
- It can be calculated as follows:

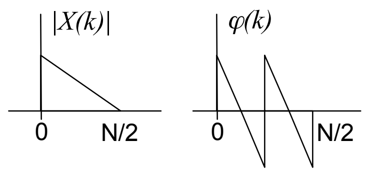
$$x(n) = IDFT[X(k)] = \frac{1}{N}\sum_{k=0}^{N-1} X(k) e^{-j2\pi nk/N}$$
$$n = 0,1,\ldots,N-1$$

- The fast version of the DFT is known as the Fast Fourier Transform (FFT) and its inverse as the IFFT. The FFT is an algorithm to compute the DFT, usually *O(N2)* operations long, in *O(NlogN)* operations
- Furthermore, there are a number of tricks to express the IDFT in terms of the FFT
- The FFT is so fast that even time-domain operations, like convolution, can be performed faster using FFT and IFFT instead.

What is Fourier saying?

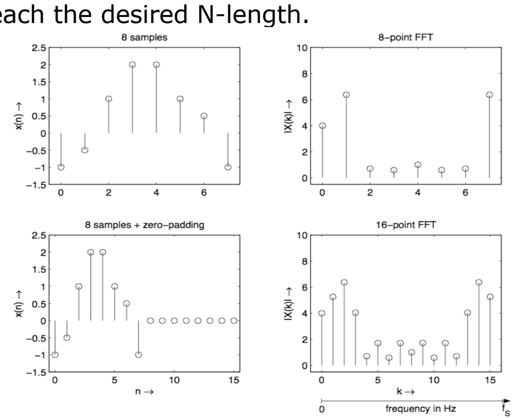


- Any periodic sound can be described by the summation of a number of sinusoids with time-varying amplitudes and phases
- Thus a complex spectrum is just a snapshot of those sinusoids' parameters



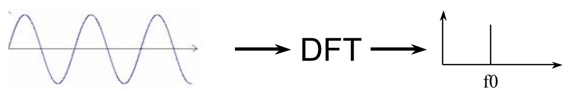
Frequency resolution

- As we now know, the frequency resolution is $\Delta f = fs/N$.
- It can be seen that to increase resolution we need to increase N
- However that implies a loss of temporal resolution
- A possible solution is to zero-pad, i.e. to add zero-valued samples until we reach the desired N-length.

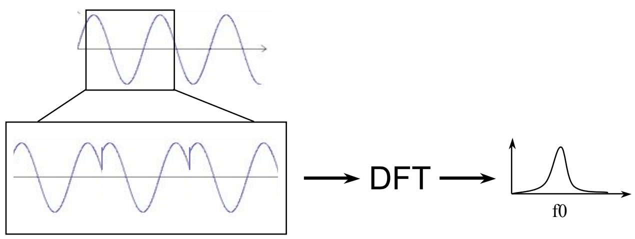


Spectral leaking

- In theory the DFT of a sinusoid shows one spectral line at f0



- In practice, unless we perform f0-synchronous analysis, there are discontinuities (sharp changes) at the segment boundaries that introduce some noise. Thus the spectral line around f0 is smeared.
- This is known as spectral leaking

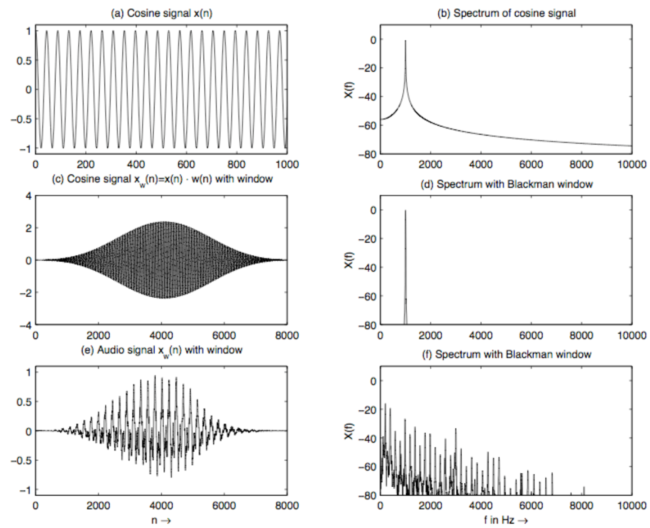


Windowing

- Segmenting is equivalent to multiplying the signal by a N-length rectangular window.
- Multiplication in time-domain is equivalent to convolution in the frequency domain
- The transform of a rectangular window is a Sinc function (*sin(x)/x*).
- We can have N = kT_0, where k is a positive integer, thus eliminating the discontinuities.
- Alternatively we can use a window that smoothly reduces the signal to zero at the boundaries
- Possible examples include Hamming (ω_H), Blackman (ω_B), Hanning, Triangular, Gaussian and Kaisser-Bessel windows.

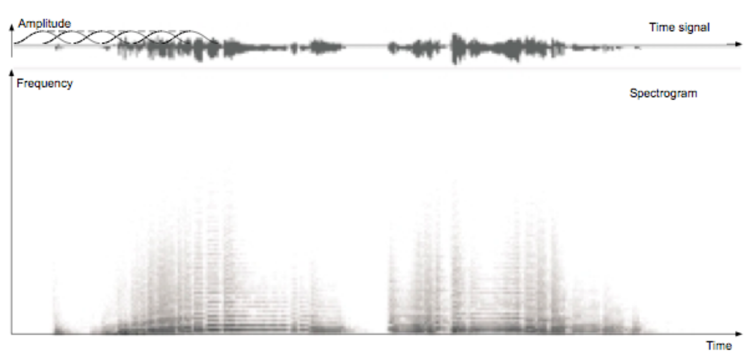
$$
\begin{aligned}
w_B(n) &= 0.42 - 0.5\cos(2\pi n/N) + 0.08\cos(4\pi n/N), \\
w_H(n) &= 0.54 - 0.46\cos(2\pi n/N) \\
&\quad n = 0, 1, \ldots N-1.
\end{aligned}
$$

Windowing



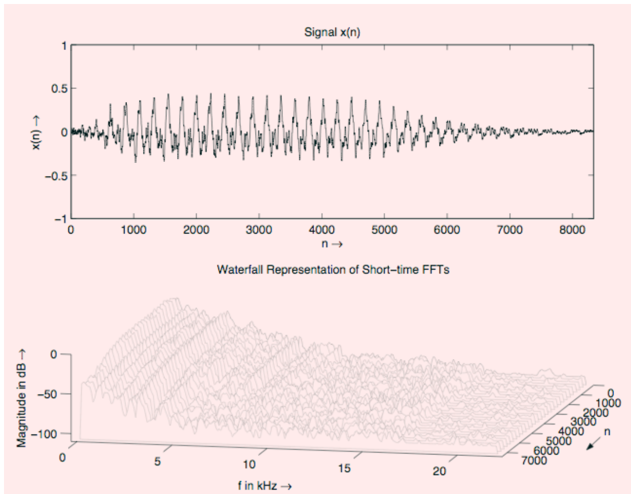
Time-frequency representation

- The Short-time Fourier Transform (STFT)
- Independent DFTs are calculated on windowed segments
- The segments usually overlap to compensate for the loss of temporal resolution
- Produces a 2-D spectrogram

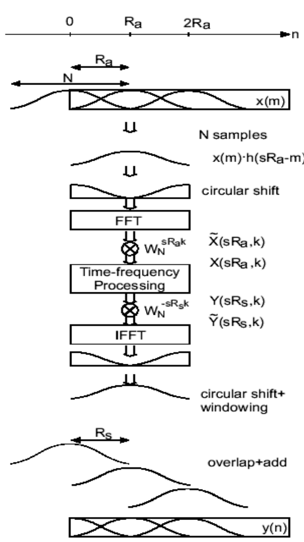


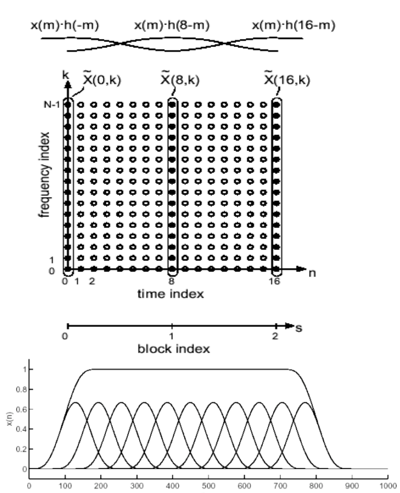
Time-frequency representation (2)

- A waterfall representation (just a different view)



FFT/IFFT Model





Useful References

- Smith, J.O., "Introduction to Digital Signal Processing",
- Zölzer, U. (Ed). "DAFX: Digital Audio Effects". John Wiley and Sons (2002)
  - Chapter 1: Zölzer, U. "Introduction".
  - Chapter 8: Arfib, D., Keiler, F. and Zölzer, U., "Time-frequency Processing".
  - Chapter 10: Amatriain, X., Bonada, J., Loscos, A. and Serra, X. "Spectral Processing".
  - Good read, Chapter 2: Dutilleux, P. and Zölzer, U. "Filters"
- Pohlmann, K. "Principles of Digital Audio". McGraw-Hill, Inc. (1995)