Windowing

Digital Signal Processing

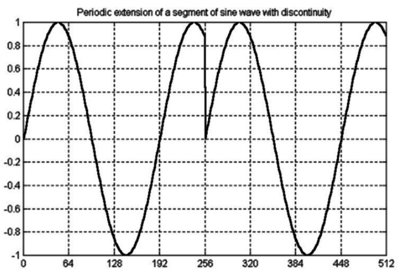
Notes

Contents

Notes

Finite Block DFT

- If signal is periodic with period of length of block of samples being processed then analysis with FFT causes no problems.
- However, if this is not the case then a discontinuity at the point where the original data ends and the extension starts occurs.



Notes

Finite Block DFT

For example:
Consider a single complex sinusoid with frequency f_0:

$$x(t) = e^{j2\pi f_0 t}$$

exists for all time and whose spectrum is $\delta(f - f_0)$.
Taking a finite block, length T_c, then resulting continuous-time waveform spectrum is, in general, non-zero at all frequencies:

$$X_{T_c}(f) = \int_{-T_c/2}^{+T_c/2} e^{j2\pi f_0 t} e^{-j2\pi f t} dt = \int_{-T_c/2}^{+T_c/2} e^{-j2\pi (f-f_0)t} dt$$

Notes

Finite Block DFT

Because of the symmetry of the limits of integration, only the even part of the integrand needs to be considered (integrating the odd part over symmetric limits will give a result of zero):

$$\begin{aligned}
X_{T_c}(f) &= \int_{-T_c/2}^{+T_c/2} cos(2\pi(f - f_0)t)dt \\
&= \left[\frac{\sin(2\pi(f-f_0)t)}{2\pi(f-f_0)}\right]_{-T_c/2}^{+T_c/2} \\
&= \frac{2\sin(\pi(f-f_0)T_c)}{2\pi(f-f_0)} \\
&= T_c \frac{\sin(\pi(f-f_0)T_c)}{\pi(f-f_0)T_c} = T_c \mathrm{sinc}((f-f_0)T_c)
\end{aligned}$$

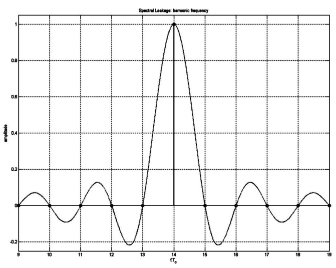
Sampling the spectrum at frequencies $f = k/T_c$, performing implicit time-domain periodic extension (with period T_c) then

$$X[k] = T_c.\mathrm{sinc}(k - f_0 T_c)$$

Notes

Harmonic Frequencies

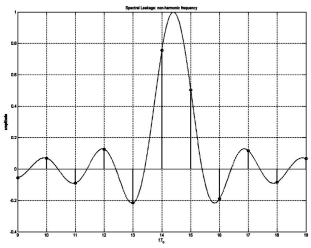
- First possibility: if $f_0 T_c$ happens to equal an integer, say n, then $X[k] = 0$ except for $k = n$ i.e. the spectrum we get is "right" because the periodic extension matches what the original sinusoid would have done: there is no discontinuity.



Notes

Non Harmonic Frequencies

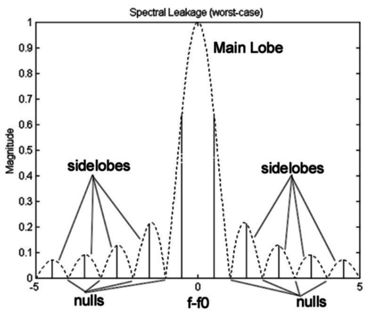
- Second possibility: if f_0T_c is not an integer (which is usually the case) then $X[k]$ is generally nonzero for all k. i.e. taking a block of data and periodically extending it has the effect of replacing an individual frequency by a sinc function which is defined over all frequencies (because the periodic extension is "wrong" and a discontinuity is created by it).



Notes

Non Harmonic Frequencies

The problem in the second case is called spectral leakage because the single frequency "leaks" into all the other frequencies (values of k), mostly into those nearby but some into the whole of the frequency spectrum!



Notes

Main Lobe

Height of largest frequency sample (main lobe) is

$$T_c \mathrm{sinc}(k^* - f_0 T_c),$$

where k^* is the integer closest to f_0T_c.

- For a harmonic frequency, the height is $T_c\mathrm{sinc}(0)$.
- For a non-harmonic frequency, the worst case (lowest possible) height is $T_c\mathrm{sinc}(0.5)$ (reduction of $\sim 4dB$).
  - Loss of sensitivity known as Scalloping Loss.
- The single frequency has become "*blurred*" into a region of frequencies.
  - Reduces the Resolution of a spectral analysis of the signal.
  - If two frequencies close together - they may appear as a single peak.
  - Resolution is ability to "*resolve*" - ability to see two peaks in spectrum when two frequencies of equal amplitude are present.

Notes

Side Lobes

- A peak observed in an analysis may not be a true signal, it may be an artifact of the interaction of the sidelobes of other signals
- It may be impossible to identify a small signal in the presence of a large one because it is swamped by the larger signal's sidelobes.
  - Problem of Dynamic Range: ratio of largest value to the smallest value.
  - i.e. Ratio of amplitudes of largest and smallest signals detected simultaneously.
  - Dynamic range is height of main lobe above the highest sidelobe.

Notes

Side Lobes

- Another closely related issue: "far-field dynamic range" is a qualitative property of the sidelobes far from the main lobe.
  - Depends on sidelobe fall-off rate as well as the height of the highest sidelobe.
  - If sidelobes fall off rapidly then very small signal can be detected in the presence of a large one, provided its frequency is far enough away from the frequency of the large signal.
- Interaction of sidelobes or the main lobe of one signal with the main lobe of another can result in a main lobe peak which no longer coincides with the signal's location, i.e. using DFT spectrum peaks to identify the frequencies present in a signal can yield biased estimates of the frequency locations.

Notes

Controlling Spectral Leakage

Spectral leakage cannot be avoided:

- It is an inevitable consequence of only having a finite set of samples to process with our DFT.

However, there are things that can be done to make its impact better-controlled:

- zero-padding which fills in the dotted-lines in the main-lobe+sidelobes diagram in the previous section, thereby allowing us to see the spectral leakage clearly
- windowing which can be used to make the sidelobes lower, at cost of making the main-lobe wider

Notes

Zero Padding

The DFT of a block of a signal, we assumed that the periodic extension is taking place immediately at the end of the available data.
Instead take a longer period but include null data (i.e. zeros) on to end of actual data.
Known as **zero-padding** with the effect of sampling the spectrum at a finer frequency spacing:

$$\frac{1}{\text{transform-length}} < \frac{1}{\text{data-length}}$$

We define bin or frequency bin as the reciprocal of the data length (not affected by zero-padding). Other definitions also popular.

Notes

Zero Padding

Zero padding gives us more points on the frequency axis,

- It does not add any new information:
- It makes use of spectral leakage to interpolate the spectrum
- Provides a better "picture" which enables us (or even a computer) to locate, for example, the peaks in a spectrum more accurately.
- Attempting to get closer to a continuous spectrum by periodically extending after some suitable multiple M of the data length
- i.e. taking frequencies which are multiples of $\frac{1}{M \times T_c}$ rather than the $\frac{1}{T_c}$ of the un-padded transform.

Notes

Zero Padding

It is possible to do zero-padding which is not an integer multiple of the original data length.

- Spectrum peaks generally won't coincide with those which would have arisen if a transform equal to the data length had been used.
- This is often not a problem with the DFT since the original signal was not actually periodic anyway (i.e. $f_0.T_c$ is not an integer)!
- This might occur when we are using an FFT (whose length is a power of 2) but the original data length isn't a power of 2.

Notes

Zero Padding

Note: The term "resolution" is sometimes used in two conflicting ways by different authors:

- the ability to distinguish two close frequencies of equal amplitude by detecting two peaks in the transform;
- the sample spacing along the frequency axis.

Zero padding does not improve resolution in sense 1 though it obviously improves it in sense 2! In ENG643D2 we will restrict the usage of the term "resolution" to sense 1.

Notes

Windowing

The problems associated with spectral leakage in a DFT can be quite severe.
So far we have assumed a sudden cut-off.
This is a discontinuity and is referred to as a rectangular window.
Step transition at the ends of the rectangular window give rise to the discontinuity which can be held responsible for the spectral leakage.

Notes

Windowing

Instead of a rectangular window we may use a more gradual change.

- By multiplying the data segment by some other carefully chosen function $w(t)$ which tapers to zero at its ends,
- thereby eliminating the discontinuity in value
  - (but still having discontinuity somewhere in the derivatives).

Notes

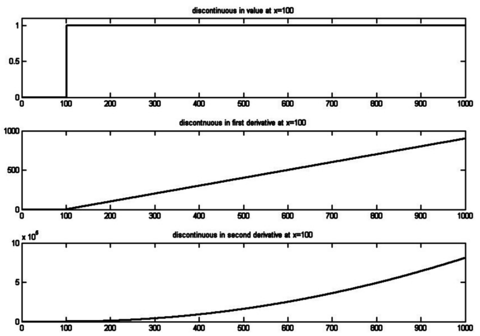
Windowing

Notes

- If data is discontinuous in value (like rectangular window), where ends do not join up, then the sidelobes fall off at 6dB per octave.
- If, instead the periodically extended data joins up at the ends (by forcing it to zero) whilst being discontinuous in its first derivative, then the sidelobes fall off as $1/f^2$ i.e. 12 dB per octave.

Windowing

Notes



Three types of discontinuity (all the discontinuities are at $x = 100$).

Windowing

Notes

Process of forcing the data to join is the essence of windowing.

- It aims to reduce sidelobes and hence improve dynamic range of the transform (i.e. ability to see small signals in the presence of large ones).

However:

- Improving the dynamic range (by reducing the sidelobe levels) degrades the resolution (by making the main lobe wider)

Trade-off between resolution and dynamic range depends on application.

The envelope of the spectrum of the window determines amplitudes of the sidelobes.

Rectangular Window

A rectangular window of length T has a (continuous) Fourier transform:

$$W_{rect}(f) = T\mathrm{sinc}(fT) = [T] \times [\sin(\pi fT)] \times \left[\frac{1}{\pi fT}\right]$$

Spectrum given by:

$$W(f) = [T] \times [g] \times [\sin(\pi fT)] \times [e(fT)]$$

where

- g is the "coherent gain" of the window ()
  - Average value of $w(t)$ - indicates scaling applied to amplitude of frequency
- $|e(fT)|$ is the "envelope" of the window's spectrum.
  - Analogous to time domain envelope of AM carrier
  - For rectangular window $g = 1$ and using bins (not Hz, using $x = fT$) then

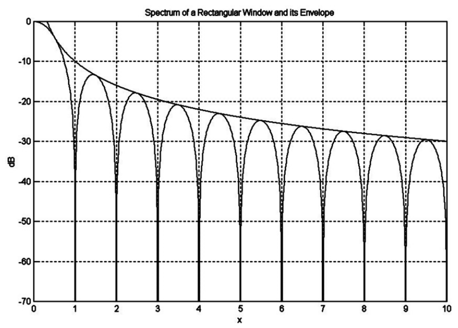
$$env_{rect}(x) = \frac{1}{\pi|x|}$$

and it is just this $1/x$ shape which leads to the 6dB/octave slope for the spectrum of this window.

Notes

Rectangular Window

The rectangular window spectrum and its envelope are shown below.



Notes

Some Common Windows

- Previously analysed the properties of the rectangular window.
- Has high sidelobes and a slow fall-off rate of the sidelobes against frequency.
- Other windows do better on these parameters, albeit at the expense of a wider main lobe.

Notes

Bartlett Window

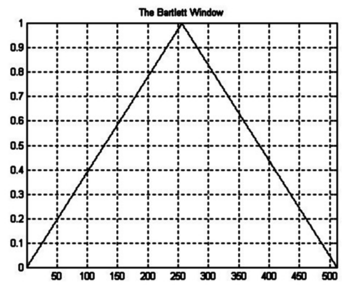
If the data being analysed using the DFT arrives sequentially and is to be processed as it arrives (e.g. a spectrum analyser or some forms of filtering) then we may want to be able to overlap our DFTs so as not to lose any information.
A window which allows this is the Bartlett (triangular) window
It overlaps exactly with a shift of 50% of the window length.

Notes

Bartlett Window

The Bartlett window has the advantage of being very simple to calculate:

$$w[n] = 1 - \frac{2n - N}{N} : n = 0, ..., N - 1$$



Notes

Bartlett Window

The *spectrum* of the Bartlett window is best viewed as a sine-squared wave multiplied by an envelope of the form:

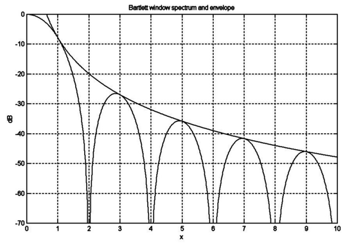
$$env_{Bart}(x) = \frac{4}{\pi^2 x^2}$$

where $x = fT$.

Notes

Bartlett Window

The $1/x^2$ asymptotic trend for the Bartlett window gives it a fall-off rate of 12 dB/octave.
Unlike the rectangular window, it is continuous in value at its ends but discontinuous in its first derivative. Its spectrum is illustrated below:



Notes

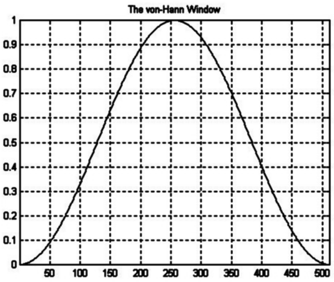
Von Hann Window

Another window which shares the 50% overlap property with the Bartlett but whose sidelobes fall off at 18dB per octave is the von-Hann (sin^2) window (also known as the "Hann" window and, confusingly, the "Hanning" window).
This is continuous in its first derivative but discontinuous in its second. Needless to say, its main lobe is wider than the Bartlett (though not by much). Its formula is:

$$w[n] = 0.5\left(1 - cos\left(\frac{2\pi n}{N}\right)\right) : n = 0, ..., N-1$$

Notes

Von Hann Window



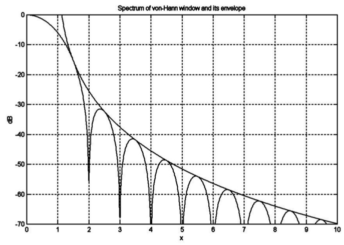
Notes

Von Hann Window

The envelope of the spectrum of the von-Hann window is:

$$env_{Von-Hann}(x) = \frac{1}{\pi|x(x^2-1)|}$$

which tends asymptotically to $\frac{1}{\pi x^3}$ for large x ($>> 1$), giving rise to the 18dB/octave slope:



Notes

Hamming Window

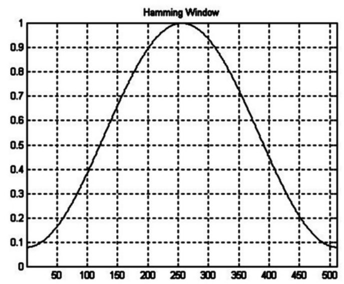
- Although the von-Hann window's rapid fall-off of sidelobes is desirable because it reduces the problems of long-distance spectral leakage,
- for some applications it may be more important to kill off the largest sidelobe,
- which is generally but not always the first, i.e. the one closest to the main lobe.
- The Hamming window does this.
- It adds a 8% of the rectangular window to 92% of the von-Hann window,
  which has the effect of cancelling out the first sidelobe.
- This makes it once more discontinuous in value
  (i.e. the sidelobes have only a 6dB/octave fall off) but its biggest sidelobe is lower than the von-Hann's.
- It is a different compromise.

Notes

Hamming Window

Formula:

$$w[n] = 0.54 - 0.46cos\left(\frac{2\pi n}{N}\right) : n = 0, ..., N-1$$



Notes

Hamming Window

The envelope of its spectrum is:

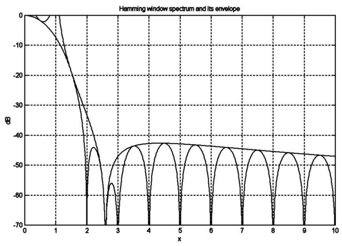
$$env_{Hamming}(x) = \frac{|\left(\frac{0.08}{0.54}\right)x^2 - 1|}{\pi|x(x^2-1)|}$$

The envelope for the Hamming window tends asymptotically to $\frac{0.148...}{\pi|x|}$ for large x ($>> 1$), which gives rise to only 6dB/octave sidelobe fall-off, as with the rectangular window, but the distant sidelobes are 16.6dB lower than for the rectangular because of the 0.148 factor.

Notes

Hamming Window

Also, the expression in its numerator gives rise to additional nulls at $x = \pm 2.598$ which is what suppresses the first sidelobe:



Notes

Windows' Asymptotes

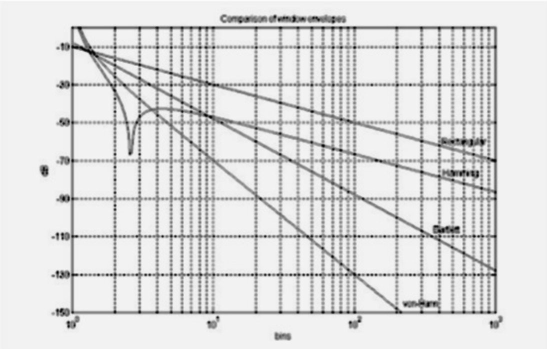
The envelopes of DFT/FFT windows' spectra (defined previously) control the behaviour of their sidelobes.

- The Rectangular & Hamming windows both fall off asymptotically at 20dB per decade, which is 6dB per octave, but the Hamming is 16.6dB lower and has an additional null at $x = 2.598$ which helps suppress the first sidelobe.
- The Bartlett falls off at 40dB per decade, which is 12dB per octave.
- The Von-Hann falls off at 60dB per decade, which is 18dB per octave.

Notes

Windows' Asymptotes

If we plot them on a log/log graph we can see their asymptotic behaviour.



Notes

Windows' Asymptotes

- Rectangular

$$e_R(x) = \frac{1}{\pi|x|} \approx \frac{0.3183}{x}$$

  for positive x
- Hamming

$$e_H(x) = \frac{\frac{0.08}{0.54}x^2 - 1}{\pi|x(x^2-1)|} \approx \frac{0.04716}{x}$$

  for large positive x
- Bartlett

$$e_B(x) = \frac{4}{\pi^2 x^2} \approx \frac{0.40528}{x^2}$$

  for large x
- Von-Hann

$$e_{V-H}(x) = \frac{1}{\pi|x(x^2-1)|} \approx \frac{0.3183}{x^3}$$

  for large positive x

Notes

Choice of Window

- Every DFT window is a compromise between main lobe width and side lobe level i.e. between resolution and dynamic range.
- The requirements in a particular situation (together with the actual number of data points available!) will determine the choice of window.
- If a finer frequency sampling is desired then a zero-padded transform can be adopted,
- in which case the time available to carry out the transform (in real-time applications) may also become a constraining factor.

Notes

Choice of Window

The following table of window parameters will help us summarise the situation:

Window	Highest side-lobe (dB)	side lobe fall off (dB/oct)	coherent gain	3dB B/W (bins)	6dB B/W (bins)	Scalloping loss (dB)
Rectangular	-13	-6	1	0.89	1.21	3.92
Bartlett	-27	-12	0.5	1.28	1.78	1.82
von Hann	-32	-18	0.5	1.44	2	1.42
Hamming	-43	-6	0.54	1.3	1.81	1.78
Dolph-Chebyshev ($\alpha = 2.5$)	-50	0	0.53	1.33	1.85	1.7

Notes

Chebyshev Window

- The Chebyshev (or Dolph-Tchebychev or Dolph-Chebyshev or ...) window is an asymmetrical window originally invented for use in radar as a way of minimising the sidelobes of an antenna's beam-shape.
- Its sidelobes don't fall off.
- It is actually defined by its shape in the frequency domain and its time-domain form has to be calculated using the inverse DFT.

Notes

Summary

- Windowing is an important consideration to help improve frequency estimation
- Many types of windows available: all with different combinations of features resulting in a trade off, often between side lobe height and main lobe width.

Notes