Effect of window size on resolution
How to deal with non-integral window frequencies?
Windowing functions

Interpreting Outputs from the Discrete Fourier Transform:

Since negative amplitudes and phase shifts by \( \pi \) have the same effect, and using complex numbers ignores the effect of phase, there are no negative amplitudes. There are negative frequencies, but there is always a “mirror image” among positive and negative frequencies.
Interpreting Outputs from the Discrete Fourier Transform:

By adding together the negative and positive frequencies, we obtain the spectrum that represents the actual components of the real signal.

No matter what the phase.....

---

Digital Audio Fundamentals: The Discrete Fourier Transform

We are going to start to look at spectra for windows of musical signals and trying to understand what they tell us about the musical performance.

There is a tradeoff between

**Temporal Resolution** – What is the shortest musical event we can observe?

**Spectral Resolution** – How many frequencies can we measure?

← Window of W Samples →
There is a tradeoff between

Temporal Resolution – What is the shortest musical event we can observe?
Spectral Resolution – How many frequencies can we measure?

The time period of the window is $W$/SampleRate, and the window rate is Sample Rate/$W$, e.g., 2000 samples lasts $2000/44100 = 0.045$ seconds, which is a window rate of 22.05 Hz.

Recall: There is a tradeoff between

Temporal Resolution – What is the shortest musical event we can observe?
Spectral Resolution – How many frequencies can we measure?

In a window of $W$ samples, we can measure a fundamental frequency $f$ whose period is the same as the window, and the harmonics $2f$, $3f$, ..., $(N/2)f$, thus, in a window of 2000 samples, we can measure 22.05 Hz, 44.1, 88.2, ..., 2205.

The measureable frequencies are multiples of the window rate $f$.

These are the ONLY frequencies we can measure.
To cover the range of human hearing (20 – 20,000 Hz), then, we would need a window of 44100/20 = 2205 samples, and this would give us the ability to measure frequencies

20, 40, 60, …., 22040

(Note that the upper bound is always the Nyquist Limit)

and the time resolution of this window is 0.05 = 1/20 sec.

To measure down to C2 (65.41 Hz, two octaves below Middle C) we would need a window of 44100/65.41 = 674.2 samples, with a resolution of 0.015 = 1/65.41 sec.

To measure down to E2 (82.41 Hz, the low string on a guitar) we would need a window of 44100/82.41 = 535.13 samples, with a resolution of 0.00186 = 1/82.41 sec.

Punchline: Probably for reasonable musical signals we have enough temporal resolution and the RANGE of measurable frequencies seems enough. . . .

BUT the resolution is ONLY at harmonics of the fundamental. Compare the list of measurable frequencies above with the frequencies of the lowest complete C-major scale on the piano:

32.7, 34.65, 36.71, 38.89, 41.2, 43.65, 46.25, 49, 51.91, 55, 58.27, 61.41, 65.41

20 40 60
For the C-major scale one octave above middle C we have the following:

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Window Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>523.25</td>
<td>554.37</td>
</tr>
<tr>
<td>587.33</td>
<td>622.25</td>
</tr>
<tr>
<td>659.26</td>
<td>698.46</td>
</tr>
</tbody>
</table>

PunchLine: We can’t even come close to measuring all frequencies in a musical signal with one window size: we don’t have enough frequencies and they don’t match precisely the pitches.

So we could use different window sizes.....
BUT, with one window, we can measure the multiples of a particular fundamental…. And this matches the way musical instruments work (in general):

A guitar tone of pitch 220 Hz (A below middle C, A on the G string) has its strongest components at the harmonics:

220, 440, 660, 880, ……

And we can find these (luckily enough!) with a window size of 2205:

44100 * 11 / 2205 = 220

So 220 Hz ought to show up as frequency 11, 440 Hz as 22, etc.

Let’s look…..

Hm…. Maybe…. Let’s compare with the Fourier Analysis produced by Electro-Acoustical Toolbox… Not so clear…..
What is a general solution to the problem of window sizes and frequencies?? The problem is in the “non-integral” components of the signal, which do not fit precisely in the window; these will be interpreted noisily by the DFT:

The problem is with the “incomplete” waveforms at the edges of the rectangular window:
The typical solution used is to de-emphasize the signal components at the edges, by tapering the amplitude of the signal using either a triangular function (which is used to modify the amplitude of the signal):

$$W(n,N) = 1 - \text{Abs}[\frac{n - (N/2)}{(N/2)}] \quad \text{for } 1 \leq n \leq N$$

Or some more complex function:

$$H(\tau, n, N) = \begin{cases} (1 - \alpha) \cos\left(\frac{2\pi \tau}{N} + \beta\right) + \alpha, & 0 \leq n < N, \\ 0, & \text{otherwise}, \end{cases}$$

where $\alpha = 1/2$. It is shown in figure 3.21a.

However you do it, these modifications will be interpreted as noise; the author of Musimatics called this "leakage, because energy that should be in one spectral harmonic spreads away (leaks) into adjacent harmonics" (p.139)
This is why the typical spectral analysis appears to be a continuous function (in addition to the fact that the signal may not just be harmonics!).

RECALL: The biggest problem with the DFT is in the “non-integral” components of the signal, which do not fit precisely in the rectangular window of N samples; these will be interpreted noisily by the DFT:

Rectangular Window Function (3.32)

\[
w(n, N) = \begin{cases} 
1 & 0 \leq n < N, \\
0 & \text{otherwise}, 
\end{cases}
\]
Let's make this precise with my DFT spreadsheet; when frequencies are integral (i.e., K complete periods within the window of N samples), we get precise measurements. Let's consider what happens with frequencies around 50 Hz:

Recall that the phase does not affect the measurement:
But now let's consider what happens when we try non-integral frequencies around 50 Hz, say 50.5 Hz; we get (nearly) symmetrical "leakage" around the frequency:

![Fourier Analysis on Input Signal Wave of 200 samples]

But let's see what happens as we change the frequency slowly from 50 to 51 Hz:

![Fourier Analysis on Input Signal Wave of 200 samples]
Let's see what happens as we change the frequency slowly from 50 to 51 Hz:

Fourier Analysis on Input Signal Wave of 200 samples

Fourier Analysis
Let's see what happens as we change the frequency slowly from 50 to 51 Hz:

![Fourier Analysis on Input Signal Wave of 200 samples](image1)

![Fourier Analysis on Input Signal Wave of 200 samples](image2)
Let’s see what happens as we change the frequency slowly from 50 to 51 Hz:

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Amplitude</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.5</td>
<td>2.3</td>
<td>2.12</td>
</tr>
<tr>
<td>59.1</td>
<td>3.142</td>
<td>3.125</td>
</tr>
<tr>
<td>78.8</td>
<td>3.2</td>
<td>3.4</td>
</tr>
<tr>
<td>97</td>
<td>4.3</td>
<td></td>
</tr>
</tbody>
</table>

Fourier Analysis on Input Signal Wave of 200 samples

Digital Audio Fundamentals: The Discrete Fourier Transform
Let's see what happens as we change the frequency slowly from 50 to 51 Hz:

Fourier Analysis on Input Signal Wave of 200 samples

Let's see what happens as we change the frequency slowly from 50 to 51 Hz:

Fourier Analysis on Input Signal Wave of 200 samples
Let's see what happens as we change the frequency slowly from 50 to 51 Hz:
Digital Audio Fundamentals: The Discrete Fourier Transform

RECALL: The typical solution used is to de-emphasize the signal components at the edges, by tapering the amplitude of the signal using either a triangular function (which is used to modify the amplitude of the signal)

\[
W(n,N) = 1 - \text{Abs}[(n - (N/2))/(N/2)] \quad \text{for } 1 \leq n \leq N
\]

Note: Looking forward, we can expect that this approach will change the amplitude measurement, since it reduces the overall sum of the samples!

Or some more complex function:

\[
H(x, n, N) = \begin{cases} 
(1 - \alpha) \cos \left( \frac{2\pi x^2}{N} + \beta \right) + \alpha, & 0 \leq n < N, \\
0, & \text{otherwise},
\end{cases}
\]

where \( \alpha = 1/2 \). It is shown in figure 3.21a.

A nice description of the “bewildering variety” (Musimathics’ nice phrase), with interesting graphics, is provided by [http://en.wikipedia.org/wiki/Window_function](http://en.wikipedia.org/wiki/Window_function)
Let’s try two of these in our experiment on a non-integral frequency of 50.4 Hz, using the Triangular and the Hann Windows:

![Windowing Functions](image)

Let’s try two of these in our experiment on a non-integral frequency of 50.4 Hz, using the Rectangular (as before), the Triangular, and the Hann Windows:

![Fourier Analysis on Input Signal Wave](image)
Let's try two of these in our experiment on a non-integral frequency of 50.4 Hz, using the Rectangular (as before), the Triangular, and the Hann Windows:

<table>
<thead>
<tr>
<th>Component Waves</th>
<th>Signal Wave</th>
<th>Probe Waves of Frequencies 1, 2, 3, ..., 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>Amplitude</td>
<td>Frequency</td>
</tr>
<tr>
<td>1</td>
<td>0.0515</td>
<td>0.0515</td>
</tr>
<tr>
<td>2</td>
<td>0.0437</td>
<td>0.0437</td>
</tr>
<tr>
<td>3</td>
<td>0.0353</td>
<td>0.0353</td>
</tr>
<tr>
<td>4</td>
<td>0.0274</td>
<td>0.0274</td>
</tr>
<tr>
<td>5</td>
<td>0.0202</td>
<td>0.0202</td>
</tr>
<tr>
<td>6</td>
<td>0.0137</td>
<td>0.0137</td>
</tr>
<tr>
<td>7</td>
<td>0.0076</td>
<td>0.0076</td>
</tr>
<tr>
<td>8</td>
<td>0.0026</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Fourier Analysis on Input Signal Wave of 200 samples

![Fourier Analysis using Triangular Window](image1)

![Signal Wave under Triangular Window](image2)

Let's try two of these in our experiment on a non-integral frequency of 50.4 Hz, using the Rectangular (as before), the Triangular, and the Hann Windows:

<table>
<thead>
<tr>
<th>Component Waves</th>
<th>Signal Wave</th>
<th>Probe Waves of Frequencies 1, 2, 3, ..., 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>Amplitude</td>
<td>Frequency</td>
</tr>
<tr>
<td>1</td>
<td>0.0515</td>
<td>0.0515</td>
</tr>
<tr>
<td>2</td>
<td>0.0437</td>
<td>0.0437</td>
</tr>
<tr>
<td>3</td>
<td>0.0353</td>
<td>0.0353</td>
</tr>
<tr>
<td>4</td>
<td>0.0274</td>
<td>0.0274</td>
</tr>
<tr>
<td>5</td>
<td>0.0202</td>
<td>0.0202</td>
</tr>
<tr>
<td>6</td>
<td>0.0137</td>
<td>0.0137</td>
</tr>
<tr>
<td>7</td>
<td>0.0076</td>
<td>0.0076</td>
</tr>
<tr>
<td>8</td>
<td>0.0026</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Fourier Analysis on Input Signal Wave of 200 samples

![Fourier Analysis using Hann Window](image3)

![Signal Wave under Hann Window](image4)
Clearly both of these reduce the "leakage," at the cost of reducing the amplitude; how does this play out in the operation of the DFT in a more complex signal? First let's try integral components and the rectangular window. We'll calculate the mean absolute error:

Next let's try integral components and the triangular window.
Next let's try integral components and the Hann window.

Punchline: Triangular and Hann windows, because they reduce the overall sum of the amplitudes by 0.5, get amplitude off by factor of 0.5, and after this correction, Hann Window does a perfect job on integral components, and this is true regardless of the phase!
Next we’ll try non-integral components and the rectangular window.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Amplitude</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.91</td>
<td>0.46</td>
</tr>
<tr>
<td>2</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

Next we’ll try non-integral components and the triangular window, with correction.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Amplitude</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.91</td>
<td>0.46</td>
</tr>
<tr>
<td>2</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>
Next we’ll try non-integral components and the Hann window, with correction.

Digital Audio Fundamentals: The Discrete Fourier Transform

Computer Science

Next we’ll try non-integral components and the Hann window, with correction, changing the phases.
Conclusions on windowing for the DFT:

(1) Window size determines frequency resolution: given a window size of $N$ samples, with a fundamental frequency of $F = \text{Sample Rate} / N$, we can only probe for the integral frequencies (the harmonics of $F$):

$$F, 2F, 3F, \ldots, kF, \ldots, \text{ceiling(Sample Rate} / 2) - 1$$

Any other frequencies will be subject to the “picket fence” problem and only approximated.

(2) Non-integral frequencies cause “leakage” to adjacent integral frequencies; good windowing functions (e.g., Hann) mitigate leakage effects and provide reasonably accurate measurements of amplitude of components, after correction.

Professional tools such as Electroacoustics Toolbox allow you to set these features, as well as window length, whether windows overlap, whether and how to average the successive measurements, whether and how to weight the measures to the psychoacoustical properties of human hearing, how to display the result, etc., etc., etc. and to output the analysis to a file.
Professional tools such as Electroacoustics Toolbox allow you to set these features, as well as window length, whether windows overlap, whether and how to average the successive measurements, whether and how to weight the measures to the psychoacoustical properties of human hearing, how to display the result, etc., etc., etc. and to output the analysis to a file.