Lecture 16

Limitations of the Fourier Transform: STFT

16.1 Learning Objectives

- Recognize the key limitation of the Fourier transform, ie: the lack of spatial resolution, or for time-domain signals, the lack of temporal resolution.

- Understand the logic behind the Short-Time Fourier Transform (STFT) in order to overcome this limitation.

- Recognize the trade-off between temporal and frequency resolution in STFT.

16.2 Introduction

The Fourier Transform is very good at identifying sinusoidal components of a time-domain signal $f(t)$[1]. However, the basic building blocks of the FT (complex exponentials: $e^{i2\pi tu}$) oscillate over all of time (between $-\infty$ and $+\infty$). For this reason, it is difficult for the FT to represent signals that are localized in time. As we have seen in several examples throughout this course, the FT of localized functions (such as the delta and rect) spans all of the frequency range. This informal observation in fact reflects a fundamental property (uncertainty principle) of the Fourier Transform. In this lecture, we will describe a Fourier-based approach to overcome this limitation of the Fourier Transform: the Short-Time Fourier Transform. Further, we will illustrate the uncertainty principle that describes the achievable time and frequency resolution that can be obtained via Fourier analysis.

[1] In this lecture, we will change our notation slightly, and describe our function f as a function of time t instead of as a function of space x as done in previous lectures. The reason for this is that the topic of the lecture, the Short Time Fourier Transform (STFT), is named after the time-domain case. However, note that this makes no difference in terms of the math and concepts.

16.3 The Short-Time Fourier Transform

In the STFT, we perform a series of windowing and FT operations: at each time shift t, we apply a window $w(t'-t)$ centered around t, and then calculate the Fourier Transform of $f(t')w(t'-t)$. For a function $f(t)$, the STFT is defined as follows:

$$\hat{f}(t,u) = \int_{-\infty}^{\infty} f(t')w(t'-t)e^{-i2\pi t'u}dt' \tag{16.1}$$

The choice of window presents important trade-offs, as will be described below.

16.4 How to Pick the STFT Window?

Intuitively, if we expect a certain signal of interest to have rapidly fluctuating properties, we should use a narrow window $w(t)$ (in order to enable fine temporal resolution). In contrast, if we expect our signal to have slowly fluctuating properties, we should use a wide window $w(t)$ (in order to enable fine frequency resolution). However, there is generally no optimal way to pick the STFT window. Indeed, this trade-off between temporal and frequency resolution is a fundamental feature of the STFT. The effect of window width on temporal and frequency resolution is illustrated in Figure 16.1. Further, the effect of the specific window shape has additional impact on the properties of the STFT, as demonstrated in Figure 16.2.

16.5 FT Uncertainty Principe: a Fundamental Limitation of Fourier Analysis

The following uncertainty principle holds for a function and its Fourier Transform. Consider a function $f(t)$ that has normalized energy ($\int_{-\infty}^{\infty}|f(t)|^2dt = 1$), for simplicity. Then, the following inequality holds:

$$\left(\int_{-\infty}^{\infty} t^2|f(t)|^2dt\right)\cdot\left(\int_{-\infty}^{\infty} u^2|\hat{f}(u)|^2du\right) \geq \frac{1}{16\pi^2} \tag{16.2}$$

Note that the first term in the left hand side of the equation above ($\int_{-\infty}^{\infty} t^2|f(t)|^2dt$) can be interpreted as a measure of how 'spread out' the signal is in time, whereas the second term ($\int_{-\infty}^{\infty} u^2|\hat{f}(u)|^2du$) is a measure of how spread out the signal is in frequency. What this inequality tells us is that a signal cannot possibly be sharply localized both in time and in frequency.

Note that this trade-off is directly at play in the STFT, and is controlled by the choice of STFT window $w(t)$. If we use a very narrow window $w(t)$ such that the STFT is highly localized in time, then the FT will be very spread out (not localized) in frequency. In contrast, if we use a broad window, we will obtain a highly localized frequency spectrum, at the cost of poor time localization. This fundamental trade-off is illustrated in

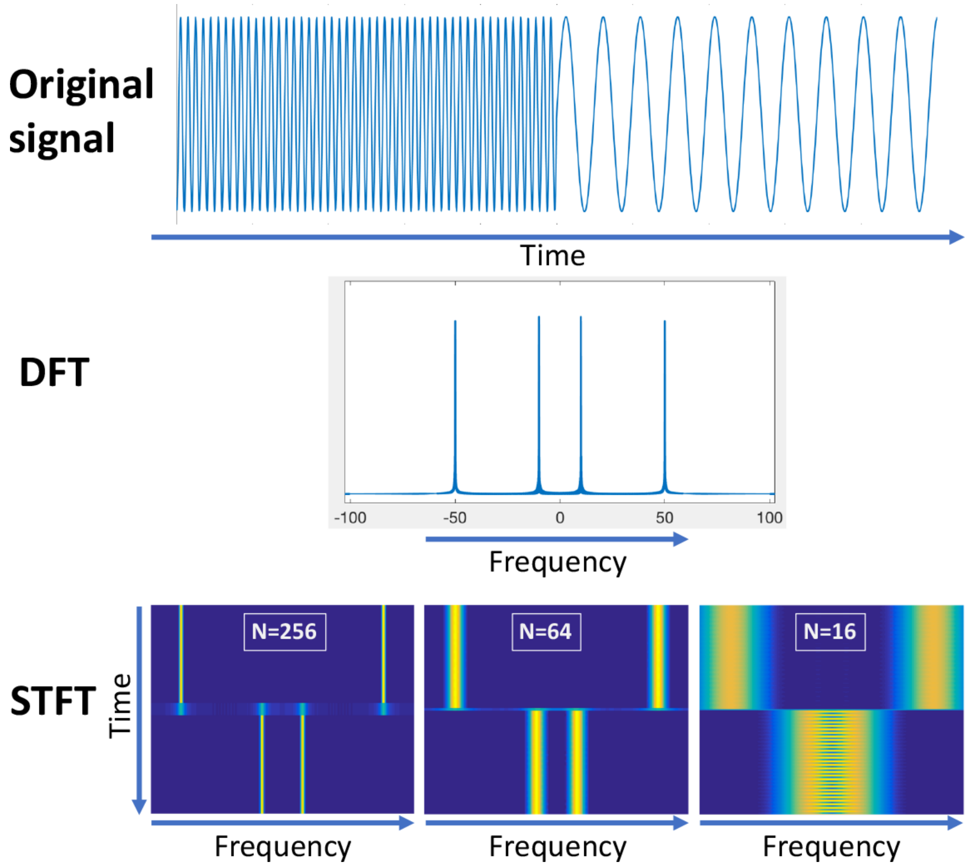

Figure 16.1: DFT vs STFT of a signal that has a high frequency for a while, then switches to a lower frequency. Note that the DFT has no temporal resolution (all of time is shown together in the frequency plot). In contrast, the STFT provides both temporal and frequency resolution: for a given time, we get a spectrum. This enables us to better represent signals with spectra that change over time. However, the STFT presents a fundamental trade-off between time resolution and frequency resolution. This trade-off is controlled by the choice of STFT window (note the improved time resolution but worsened frequency resolution as we make the STFT window narrower).

Figure 16.1, where the width of the window is varied from 256 samples to 16 samples leading to improved temporal localization (vertical axis) but worsened frequency localization (horizontal axis).

16.6 How to Overcome the Uncertainty Principle?

As we have observed in this lecture, the STFT enables us to trade-off time and frequency resolution as a way to analyze our signals. This trade-off is controlled by the choice of STFT window (eg: broad vs narrow), and limited by the uncertainty principle (Eq. 16.2). The main limitation of the STFT is that it has a fixed temporal resolution. But, can we

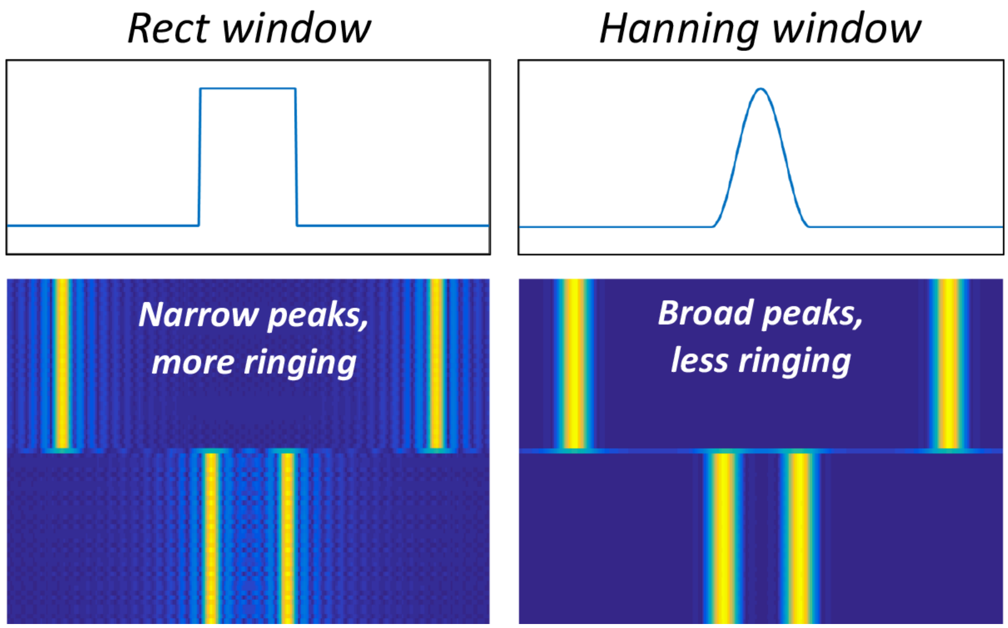

Figure 16.2: Besides the total width of the STFT window, the specific window shape determines the output of the STFT. This figure compares two choices of STFT windows with the same overall width, applied to the signal shown in Figure 16.1 above. A flat ('rect') window results in finer frequency resolution but more frequency ringing, whereas a more tapered ('hanning' in this case) window results in some loss of frequency resolution, but reduced ringing.

do better? One potential way to overcome this limitation would be to calculate a set of STFT transforms, each using a different window width. In this way, we would perform a kind of multi-resolution analysis, where we would analyze low and high frequency components, at different spatial locations, and with different spatial/frequency resolutions. However, this approach is quite redundant, as we would essentially need a 3D decomposition (with dimensions: time, frequency, window width) to analyze a 1D signal. Instead, it would be desirable to decompose our signal in a way that high frequency components are analyzed with high spatial resolution (since they vary rapidly in time), and low frequency components are analyzed with low temporal resolution. In other words, *there has to be a better way!* This better way is called wavelet analysis, and will be the topic of our next few lectures.