Short Time Fourier Transform*

Ivan Selesnick

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 1.0†

Abstract

Introduction to the Short Time Fourier Transform, which includes it's definition and methods for its use.

1 Short Time Fourier Transform

The Fourier transforms (FT, DTFT, DFT, *etc.*) do not clearly indicate how the frequency content of a signal changes over time.

That information is hidden in the phase - it is not revealed by the plot of the magnitude of the spectrum.

> NOTE: To see how the frequency content of a signal changes over time, we can cut the signal into blocks and compute the spectrum of each block.

To improve the result,

1. blocks are overlapping
2. each block is multiplied by a window that is tapered at its endpoints.

Several parameters must be chosen:

- Block length, R.
- The type of window.
- Amount of overlap between blocks. (Figure 1 (STFT: Overlap Parameter))
- Amount of zero padding, if any.

*Version 2.4: Aug 9, 2005 2:16 pm -0500
†http://creativecommons.org/licenses/by/1.0

STFT: Overlap Parameter

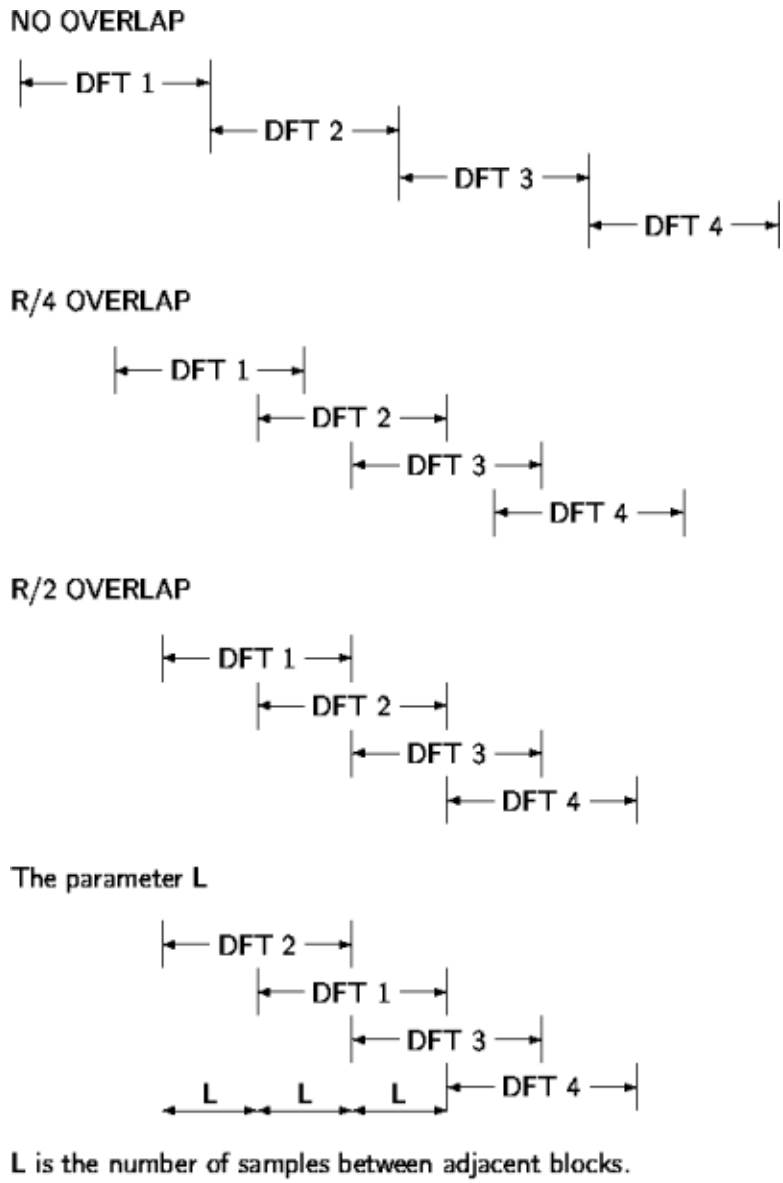


Figure 1

The short-time Fourier transform is defined as

$$\begin{aligned} X(\omega, m) &= \mathrm{STFT}(x(n)) := \mathrm{DTFT}(x(n-m)\, w(n)) \\ &= \sum_{n=-\infty}^{\infty} x(n-m)\, w(n)\, e^{-(i\omega n)} \\ &= \sum_{n=0}^{R-1} x(n-m)\, w(n)\, e^{-(i\omega n)} \end{aligned} \tag{1}$$

where $w(n)$ is the window function of length R.

1. The STFT of a signal $x(n)$ is a function of two variables: time and frequency.
2. The block length is determined by the support of the window function $w(n)$.
3. A graphical display of the magnitude of the STFT, $|X(\omega, m)|$, is called the **spectrogram** of the signal. It is often used in speech processing.
4. The STFT of a signal is invertible.
5. One can choose the block length. A long block length will provide higher frequency resolution (because the main-lobe of the window function will be narrow). A short block length will provide higher time resolution because less averaging across samples is performed for each STFT value.
6. A **narrow-band spectrogram** is one computed using a relatively long block length R, (long window function).
7. A **wide-band spectrogram** is one computed using a relatively short block length R, (short window function).

1.1 Sampled STFT

To numerically evaluate the STFT, we sample the frequency axis ω in N equally spaced samples from $\omega = 0$ to $\omega = 2\pi$.

$$\forall k, 0 \leq k \leq N-1 : \left(\omega_k = \frac{2\pi}{N} k\right) \tag{2}$$

We then have the discrete STFT,

$$\begin{aligned} X^d(k,m) := X\left(\tfrac{2\pi}{N}k, m\right) &= \textstyle\sum_{n=0}^{R-1} x(n-m)\, w(n)\, e^{-(i\omega n)} \\ &= \textstyle\sum_{n=0}^{R-1} x(n-m)\, w(n)\, {W_N}^{-(kn)} \\ &= \mathrm{DFT}_N\left(x(n-m)\, w(n)\,|_{n=0}^{R-1}, 0,\ldots 0\right) \end{aligned} \tag{3}$$

where 0,...0 is $N - R$.

In this definition, the overlap between adjacent blocks is $R - 1$. The signal is shifted along the window one sample at a time. That generates more points than is usually needed, so we also sample the STFT along the time direction. That means we usually evaluate

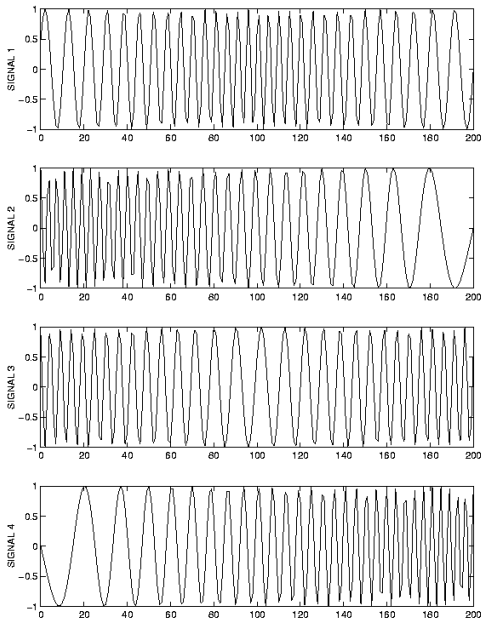
$$X^d(k, Lm)$$

where L is the time-skip. The relation between the time-skip, the number of overlapping samples, and the block length is

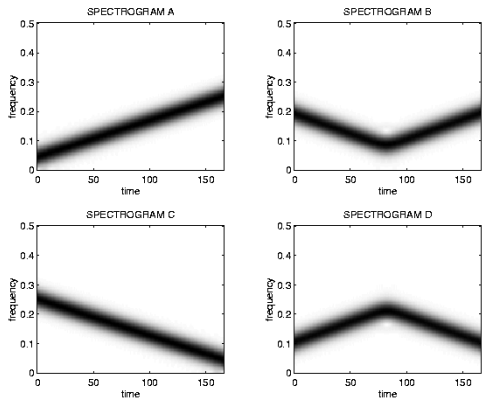
$$\text{Overlap} = R - L$$

Exercise 1 ***(Solution on p. 14.)***

Match each signal to its spectrogram in Figure 2.


SIGNAL 1
SIGNAL 2
SIGNAL 3
SIGNAL 4

(a)


SPECTROGRAM A
SPECTROGRAM B
SPECTROGRAM C
SPECTROGRAM D
frequency
time

(b)

Figure 2

1.2 Spectrogram Example

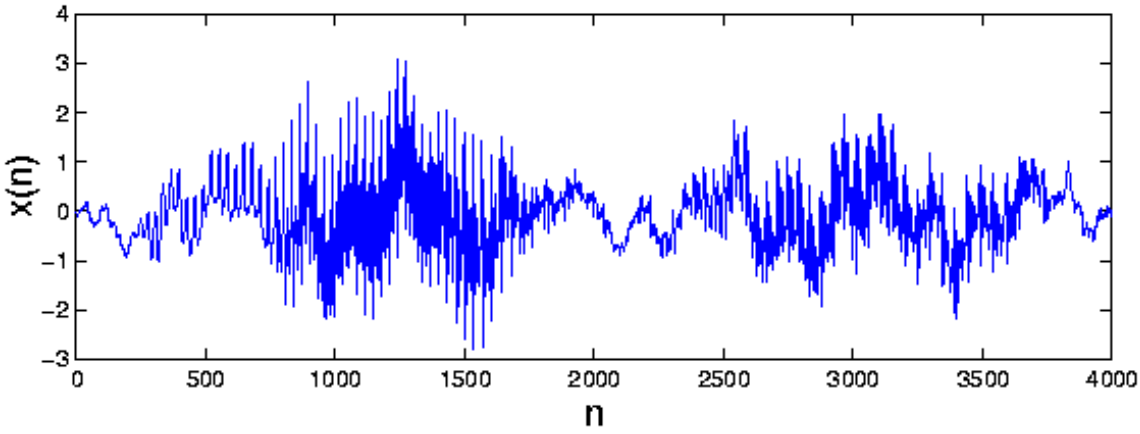


Figure 3

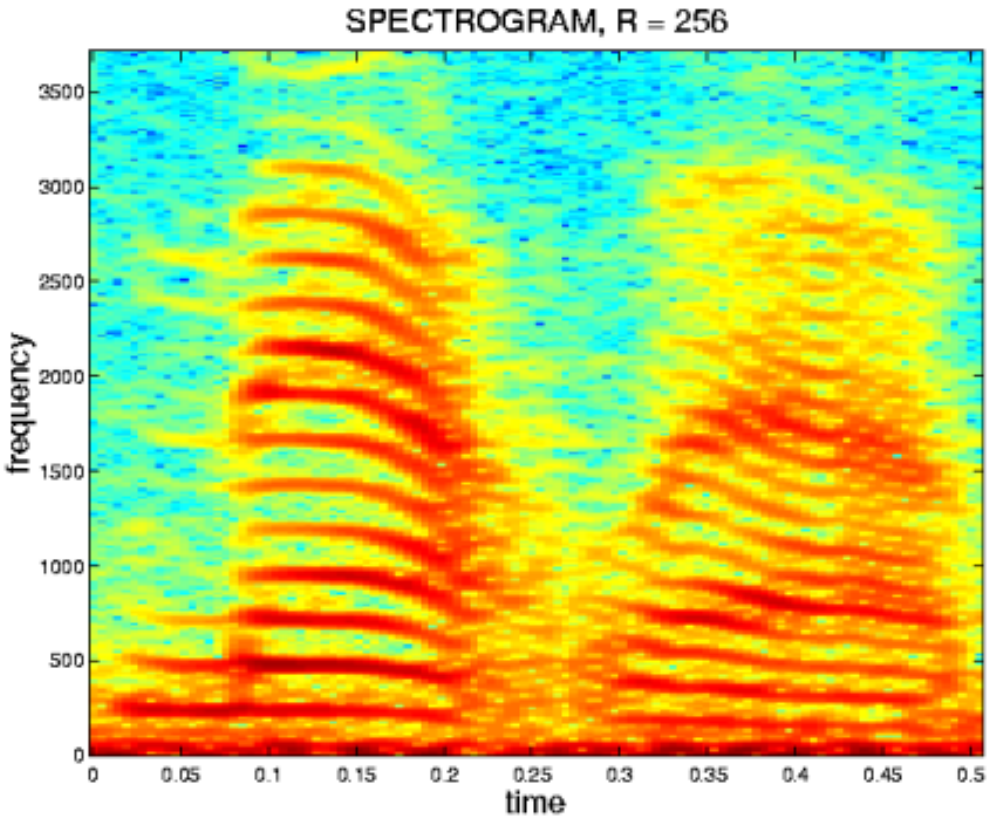


Figure 4

The matlab program for producing the figures above (Figure 3 and Figure 4).

```
% LOAD DATA
load mtlb;
x = mtlb;

figure(1), clf
plot(0:4000,x)
xlabel('n')
ylabel('x(n)')

% SET PARAMETERS
R = 256;                % R: block length
window = hamming(R);    % window function of length R
N = 512;                % N: frequency discretization
L = 35;                 % L: time lapse between blocks
```

```
fs = 7418;              % fs: sampling frequency
overlap = R - L;

% COMPUTE SPECTROGRAM
[B,f,t] = specgram(x,N,fs,window,overlap);

% MAKE PLOT
figure(2), clf
imagesc(t,f,log10(abs(B)));
colormap('jet')
axis xy
xlabel('time')
ylabel('frequency')
title('SPECTROGRAM, R = 256')
```

1.3 Effect of window length R

Narrow-band spectrogram: better frequency resolution

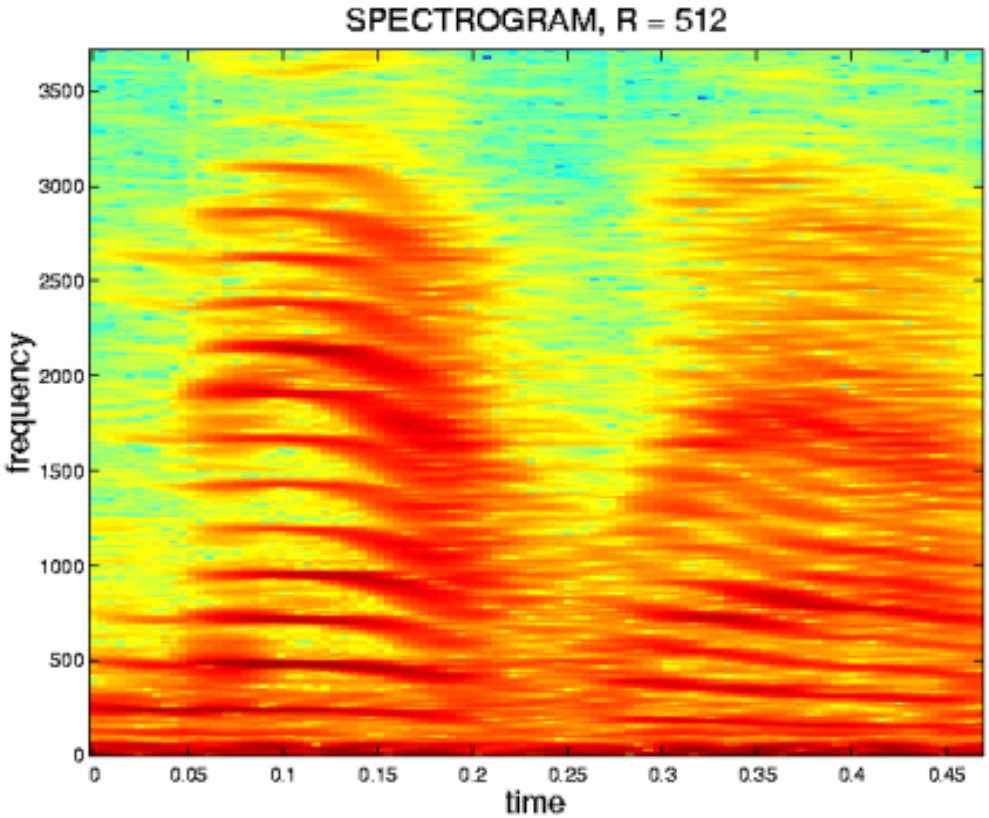


Figure 5

Wide-band spectrogram: better time resolution

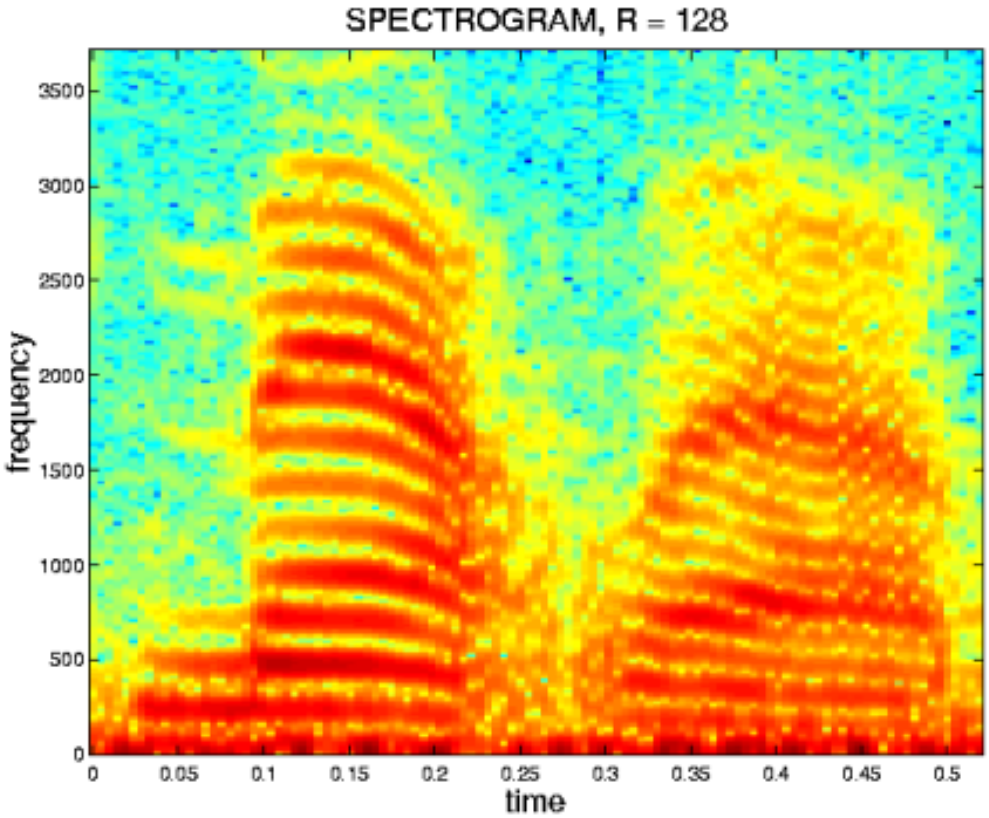
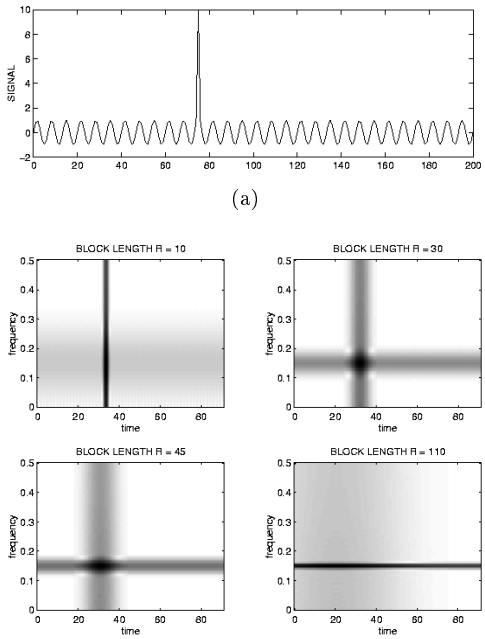


Figure 6

Here is another example to illustrate the frequency/time resolution trade-off (See figures - Figure 5 (Narrow-band spectrogram: better frequency resolution), Figure 6 (Wide-band spectrogram: better time resolution), and Figure 7 (Effect of Window Length R)).

Effect of Window Length R



(a)

(b)

Figure 7

1.4 Effect of L and N

A spectrogram is computed with different parameters:

$$L \in \{1, 10\}$$

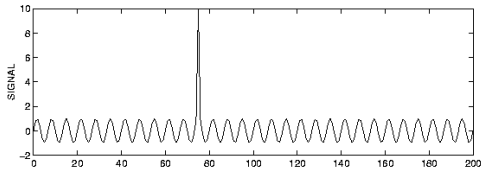
$$N \in \{32, 256\}$$

- L = time lapse between blocks.
- N = FFT length (Each block is zero-padded to length N.)

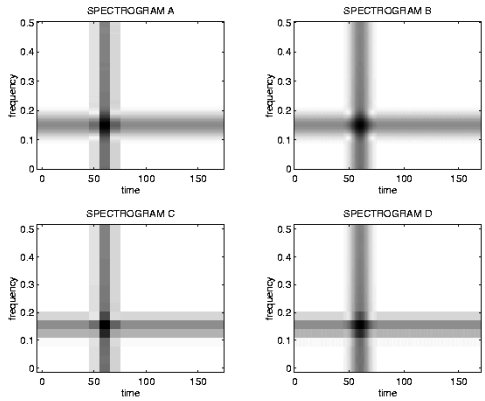
In each case, the block length is 30 samples.

Exercise 2 ***(Solution on p. 14.)***

For each of the four spectrograms in Figure 8 can you tell what L and N are?



(a)



(b)

Figure 8

L and N do not effect the time resolution or the frequency resolution. They only affect the 'pixelation'.

1.5 Effect of R and L

Shown below are four spectrograms of the same signal. Each spectrogram is computed using a different set of parameters.

$$R \in \{120, 256, 1024\}$$

$$L \in \{35, 250\}$$

where

- R = block length
- L = time lapse between blocks.

Exercise 3 ***(Solution on p. 14.)***

For each of the four spectrograms in Figure 9, match the above values of L and R.

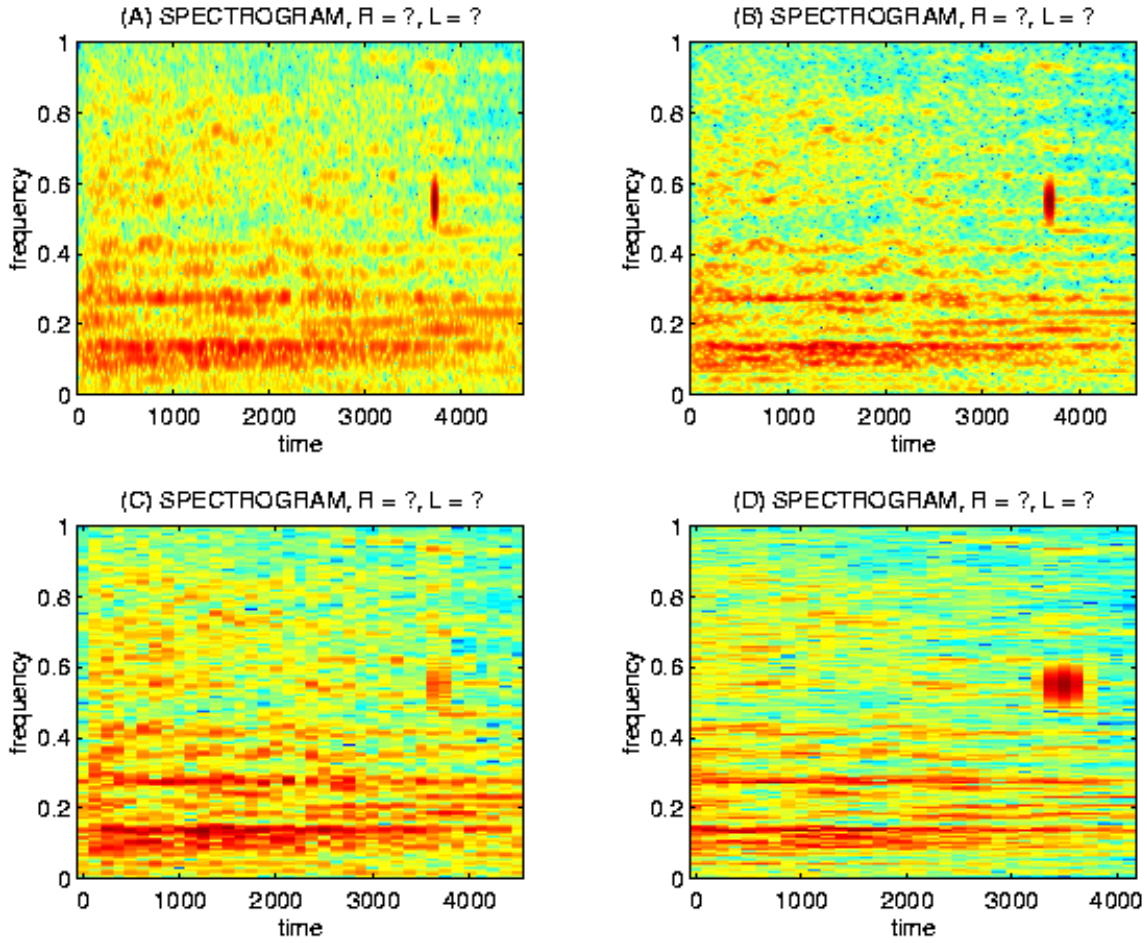


Figure 9

If you like, you may listen to this signal with the `soundsc` command; the data is in the file: `stft_data.m`. Here (Figure 10) is a figure of the signal.

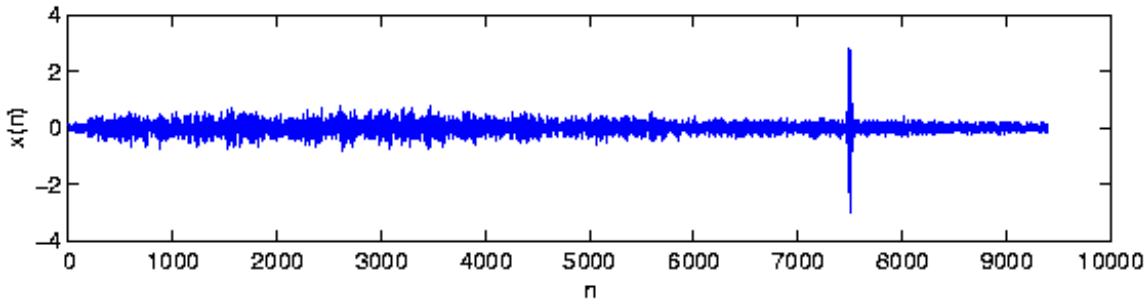

4
2
0
−2
−4
x(n)
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
n

Figure 10

Solutions to Exercises in this Module

Solution to Exercise (p. 3)

Solution to Exercise (p. 10)

Solution to Exercise (p. 11)