Lecture 18: Power Spectrum

Mark Hasegawa-Johnson
All content CC-SA 4.0 unless otherwise specified.

ECE 401: Signal and Image Analysis, Fall 2020
Motivation: Noisy Telephones

Auditory Filters

White Noise

Noise of Many Colors

Summary
1 Motivation: Noisy Telephones

2 Auditory Filters

3 White Noise

4 Noise of Many Colors

5 Summary
In the 1920s, Harvey Fletcher had a problem. Telephones were noisy (very noisy). Sometimes, people could hear the speech. Sometimes not. Fletcher needed to figure out why people could or couldn’t hear the speech, and what Western Electric could do about it.
Tone-in-Noise Masking Experiments

He began playing people pure tones mixed with noise, and asking people “do you hear a tone”? If 50% of samples actually contained a tone, and if the listener was right 75% of the time, he considered the tone “audible.”
People’s ears are astoundingly good. This tone is inaudible in this noise. But if the tone was only \(2\times\) greater amplitude, it would be audible.
Even more astounding: the same tone, in a very slightly different noise, is perfectly audible, to every listener.
What’s going on (why can listeners hear the tone?)

1kHz tone waveform

White noise waveform

Tone + white noise waveform

1kHz tone waveform

Bandstop noise waveform

Tone + bandstop noise waveform
Outline

1 Motivation: Noisy Telephones
2 Auditory Filters
3 White Noise
4 Noise of Many Colors
5 Summary
Remember the discrete time Fourier transform (DTFT):

\[ X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}, \quad x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)| e^{j\omega n} d\omega \]

If the signal is only \( N \) samples in the time domain, we can calculate samples of the DTFT using a discrete Fourier transform:

\[ X[k] = X\left(\omega_k = \frac{2\pi k}{N}\right) = \sum_{n=0}^{\infty} x[n] e^{-j\frac{2\pi kn}{N}} \]

We sometimes write this as \( X[k] = X(\omega_k) \), where, obviously, \( \omega_k = \frac{2\pi k}{N} \).
What’s going on (why can listeners hear the tone?)

1kHz tone waveform

White noise waveform

tone + white noise waveform

1kHz tone waveform

Bandstop noise waveform

tone + bandstop noise waveform
Here’s the DFT power spectrum ($|X[k]|^2$) of the tone, the white noise, and the combination.
Bandstop Noise

The “bandstop” noise is called “bandstop” because I arbitrarily set its power to zero in a small frequency band centered at 1kHz. Here is the power spectrum. Notice that, when the tone is added to the noise signal, the little bit of extra power makes a noticeable (audible) change, because there is no other power at that particular frequency.
Fletcher’s Model of Masking

Fletcher proposed the following model of hearing in noise:

1. The human ear pre-processes the audio using a bank of bandpass filters.
2. The power of the noise signal, in the $k^{\text{th}}$ bandpass filter, is $N_k$.
3. The power of the noise+tone is $N_k + T_k$.
4. If there is any band, $k$, in which $\frac{N_k + T_k}{N_k} > \text{threshold}$, then the tone is audible. Otherwise, not.
In 1928, Georg von Békésy found Fletcher’s auditory filters. Surprise: they are **mechanical**.

The inner ear contains a long (3cm), thin (1mm), tightly stretched membrane (the basilar membrane). Like a steel drum, it is tuned to different frequencies at different places: the outer end is tuned to high frequencies, the inner end to low frequencies.

About 30,000 nerve cells lead from the basilar membrane to the brain stem. Each one sends a signal if its part of the basilar membrane vibrates.
The Internal Ear


DOI:10.15347/wjm/2014.010. ISSN 2002-4436.
Frequency responses of the auditory filters

Here are the squared magnitude frequency responses ($|H(\omega)|^2$) of 26 of the 30000 auditory filters. I plotted these using the parametric model published by Patterson in 1974:
An acoustic white noise signal (top), filtered through a spot on the basilar membrane with a particular impulse response (middle), might result in narrowband-noise vibration of the basilar membrane (bottom).
An acoustic white noise signal (top), filtered through a spot on the basilar membrane with a particular impulse response (middle), might result in narrowband-noise vibration of the basilar membrane (bottom).

![White Noise Powerspectrum](image)

![Squared Frequency Response: Auditory Filter @ 1kHz](image)

![Auditory-Filtered White Noise Powerspectrum](image)
If there is a tone embedded in the noise, then even after filtering, it’s very hard to see that the tone is there...
But, Fourier comes to the rescue! In the power spectrum, it is almost possible, now, to see that the tone is present in the white noise masker.
If the masker is bandstop noise, instead of white noise, the spectrum after filtering looks very different...
Filtered tone + bandstop noise

... and the tone+noise looks very, very different from the noise by itself.

This is why the tone is audible!
What an excellent model! Why should I believe it?

Now that you’ve seen the pictures, it’s time to learn the math.

- What is white noise?
- What is a power spectrum?
- What is filtered noise?

Let’s find out.
Outline

1. Motivation: Noisy Telephones
2. Auditory Filters
3. White Noise
4. Noise of Many Colors
5. Summary
What is Noise?

By “noise,” we mean a signal $x[n]$ that is unpredictable. In other words, each sample of $x[n]$ is a random variable.
“White noise” is a noise signal where each sample, \( x[n] \), is uncorrelated with all the other samples. Using \( E[\cdot] \) to mean “expected value,” we can write:

\[
E [x[n]x[n + m]] = E [x[n]] E [x[n + m]] \quad \text{for } m \neq 0
\]

**Most noises are not white noise.** The equation above is only true for white noise. White noise is really useful, so we'll work with this equation a lot, but it’s important to remember: **Only white noise has uncorrelated samples.**
What is Zero-Mean, Unit-Variance White Noise?

Zero-mean, unit-variance white noise is noise with uncorrelated samples, each of which has zero mean:

$$\mu = E [x[n]] = 0$$

and unit variance:

$$\sigma^2 = E \left[ (x[n] - \mu)^2 \right] = 1$$

Putting the above together with the definition of white noise, we get

$$E [x[n]x[n + m]] = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0 \end{cases}$$
What is the Spectrum of White Noise?

Let's try taking the Fourier transform of zero-mean, unit-variance white noise:

\[ X(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \]

The right-hand side of the equation is random, so the left-hand side is random too. In other words, if \( x[n] \) is noise, then

- for any particular frequency, \( \omega \), that you want to investigate,
- \( X(\omega) \) is a random variable.
- It has a random real part (\( X_R(\omega) \))
- It has a random imaginary part (\( X_I(\omega) \)).
What is the Average Fourier Transform of White Noise?

Since \( X(\omega) \) is a random variable, let’s find its expected value.

\[
E \left[ X(\omega) \right] = E \left[ \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \right]
\]

Expectation is linear, so we can write

\[
E \left[ X(\omega) \right] = \sum_{n=-\infty}^{\infty} E [x[n]] e^{-j\omega n}
\]

But \( E [x[n]] = 0 \)! So

\[
E \left[ X(\omega) \right] = 0
\]

That’s kind of disappointing, really. Who knew noise could be so boring?
What is the Average Squared Magnitude Spectrum of White Noise?

Fortunately, the squared magnitude spectrum, \( |X(\omega)|^2 \), is a little more interesting:

\[
E \left[ |X(\omega)|^2 \right] = E \left[ X(\omega)X^*(\omega) \right]
\]

Goodness. What is that? Let’s start out by trying to figure out what is \( X^*(\omega) \), the complex conjugate of \( X(\omega) \).
What is the Complex Conjugate of a Fourier Transform?

First, let's try to figure out what $X^*(\omega)$ is:

$$X^*(\omega) = (\mathcal{F}\{x[m]\})^*$$

$$= \left( \sum_{m=-\infty}^{\infty} x[m]e^{-j\omega m} \right)^*$$

$$= \sum_{m=-\infty}^{\infty} x[m]e^{j\omega m}$$
Average Squared Magnitude Spectrum?

Now let’s plug back in here:

\[ E \left[ |X(\omega)|^2 \right] = E \left[ X(\omega)X^*(\omega) \right] \]

\[ = E \left[ \left( \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \right) \left( \sum_{m=-\infty}^{\infty} x[m]e^{j\omega m} \right) \right] \]

\[ = E \left[ \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} x[n]x[m]e^{j\omega(m-n)} \right] \]

\[ = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} E \left[ x[n]x[m] \right] e^{j\omega(m-n)} \]

But remember the definition of zero-mean white noise: \( E \left[ x[n]x[m] \right] = 0 \) unless \( n = m \). So

\[ E \left[ |X(\omega)|^2 \right] = \sum_{n=-\infty}^{\infty} E \left[ x^2[n] \right] \]
We have proven that the expected squared-magnitude Fourier transform of zero-mean, unit-variance white noise is a constant:

\[
E \left[ |X(\omega)|^2 \right] = \sum_{n=-\infty}^{\infty} E \left[ x^2[n] \right] = \sum_{n=-\infty}^{\infty} 1
\]

Unfortunately, the constant is infinity!
Wiener solved this problem by defining something called the Power Spectrum:

\[
R_{xx}(\omega) = \lim_{N \to \infty} \frac{1}{N} \left| \sum_{n=-(N-1)/2}^{(N-1)/2} x[n] e^{-j\omega n} \right|^2
\]
What is power?

- Power (Watts=Joules/second) is energy (in Watts) per unit time (in seconds).
- Example: electrical energy = volts×charge, power=volts×current (current = charge/time)
- Example: mechanical energy = force×distance, power=force×velocity (velocity = distance/time)
So the power spectrum of white noise is

\[ R_{xx}(\omega) = \lim_{N \to \infty} \frac{1}{N} |X(\omega)|^2 \]

where \( N \) is the number of samples over which we computed the Fourier transform.
And now here’s why white noise is called “white:”

\[
E [R_{xx}(\omega)] = \frac{1}{N} E \left[ |X(\omega)|^2 \right] \\
= \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} E [x^2[n]] \\
= \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} 1 \\
= 1
\]

The power spectrum of white noise is, itself, a random variable; but its expected value is the power of \(x[n]\),

\[
E [R_{xx}(\omega)] = E[x^2[n]] = 1
\]
Outline

1. Motivation: Noisy Telephones
2. Auditory Filters
3. White Noise
4. Noise of Many Colors
5. Summary
Suppose that we filter the white noise, like this:

\[ y[n] = h[n] \ast x[n] \leftrightarrow Y(\omega) = H(\omega)X(\omega) \]
The power spectrum of the filtered noise will be

\[ R_{yy}(\omega) = \lim_{N \to \infty} \frac{1}{N} |H(\omega)X(\omega)|^2 \]

\[ = |H(\omega)|^2 R_{xx}(\omega) \]

\[ = |H(\omega)|^2 \]

where \( N \) is the number of samples over which we computed the Fourier transform.
Noise with a flat power spectrum (uncorrelated samples) is called white noise.

Noise that has been filtered (correlated samples) is called colored noise.

- If it’s a low-pass filter, we call it pink noise (this is quite standard).
- If it’s a high-pass filter, we could call it blue noise (not so standard).
- If it’s a band-pass filter, we could call it green noise (not at all standard, but I like it!)
Motivation

Filters

White Noise

Colors

Summary

What is the Power of Filtered Noise?

Remember that, for white noise, we had

\[ E[R_{xx}(\omega)] = E[x^2[n]] = 1 \]

The same thing turns out to be true for filtered noise:

\[ E[y^2[n]] = \text{average} (E[R_{yy}(\omega)]) \]
\[ = \frac{1}{2\pi} \int_{-\pi}^{\pi} E[R_{yy}(\omega)] \, d\omega \]
For example, here is a white noise signal $x[n]$, and its power spectrum $R_{xx}(\omega) = \frac{1}{N}|X(\omega)|^2$:
Here’s the same signal, after filtering with a lowpass filter with cutoff $\pi/2$:
Power of Blue Noise: Example

Here’s the same signal, after filtering with a highpass filter with cutoff $\pi/2$: 

![Power Spectrum of a Blue Noise (Highpass Filtered)](image1)

![Time Domain Signal: Blue Noise (Highpass Filtered)](image2)
The relationship between energy in the time domain and energy in the frequency domain is summarized by Parseval’s theorem. There is a form of Parseval’s theorem for every type of Fourier transform. For the DTFT, it is

$$\sum_{n=-\infty}^{\infty} x^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega$$

For the DFT, it is:

$$\sum_{n=0}^{N-1} x^2[n] = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$$
For the power spectrum, Parseval’s theorem says that power is the same in both time domain and frequency domain:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} x^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_{xx}(\omega)\omega$$

The DFT version of power is:

$$\frac{1}{N} \sum_{n=0}^{N-1} x^2[n] = \frac{1}{N} \sum_{k=0}^{N-1} |R_{xx}(\omega_k)|^2$$
Outline

1. Motivation: Noisy Telephones
2. Auditory Filters
3. White Noise
4. Noise of Many Colors
5. Summary
Summary

- Masking: a pure tone can be heard, in noise, if there is at least one auditory filter through which $\frac{N_k + T_k}{N_k} > \text{threshold}$.
- Zero-mean, unit-variance white Noise: samples of $x[n]$ are uncorrelated, so the expected power spectrum is:

$$E[R_{xx}(\omega)] = 1$$

- Power spectrum in general: The relationship between power in the time domain, and power in the frequency domain, in general, is given by Parseval’s Theorem:

$$\sum_{n=-\infty}^{\infty} x^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega$$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=-(N-1)/2}^{(N-1)/2} x^2[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_{xx}(\omega) \omega$$