Channel Characteristics and Impairments

This lecture describes some of the most common channel characteristics and impairments.
After this lecture you should be able to: classify channels as high-, low-, or band-pass; determine -dB channel bandwidth and percentage power signal bandwidth; convert between delay and phase shift; compute group delay from phase response; identify some causes of multipath propagation and their effects on the channel frequency response; distinguish between linear- and non-linear distortion; compute frequencies of third-order IMD products; compute THD; compute SNR; compute the probability that a Gaussian source will exceed a certain value; identify sources of near-end and far-end crosstalk; and distinguish between noise and interference.

Frequency Response

We can model a channel as a filter. The frequency response of this filter is called the frequency-domain transfer function, typically denoted as $H(f)$. It is the ratio of the voltage at the output of the channel to the voltage at its input at the frequency f:

$$H(f) = \frac{V_{\text{out}}}{V_{\text{in}}}$$

For linear systems the input and output are at the same frequency and we can measure both the amplitude ratio and the phase difference. The frequency response is thus a complex quantity and is a function of frequency. The ratio of the voltages is called the amplitude response and the difference of the phases is called the phase response:

$$|H(f)| = \frac{|V_{\text{out}}|}{|V_{\text{in}}|}$$

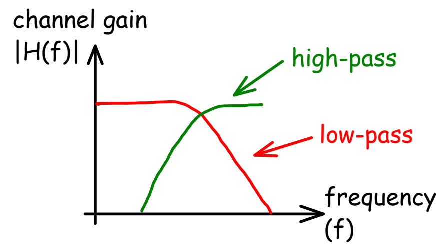
and

$$\angle H(f) = \angle V_{\text{out}} - \angle V_{\text{in}}$$

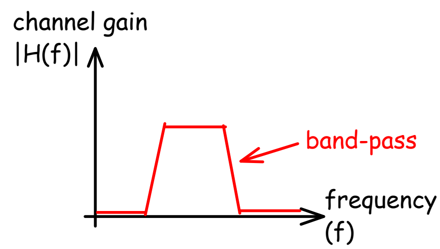
Exercise 1: A 10 dBm signal is applied to one end of a 50 ohm co-ax cable at frequencies of 1, 10 and 100 MHz. At the other end you measure voltages of 7, 4 and 0 dBm respectively. Plot the amplitude of the transfer function of the channel formed by this cable. Show dB on the vertical axis and log of frequency on the horizontal axis.

If the channel primarily passes signals below a certain frequency it is called a *low-pass* channel. A typical example is a twisted-pair cable because the distributed series inductance and parallel capacitance attenuate higher-frequency signals.

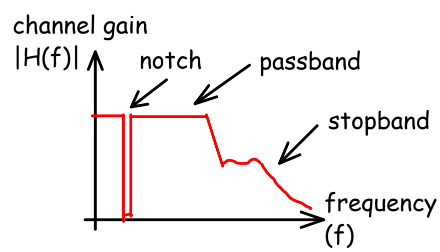
A *high-pass* channel often results from capacitive or transformer coupling which blocks DC or low-frequency signals.



A *band-pass* channel has both low- and high-pass characteristics. An example would be a transformer-coupled telephone line. Some band-pass channels result from attenuation by the channel but in other cases the band-pass nature of the channel intentional and is designed into the system to separate different users at different frequencies.



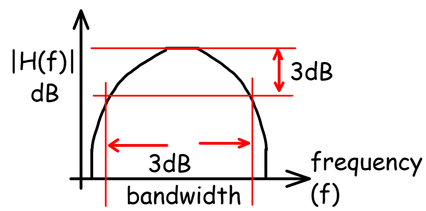
Channels often include several of the above characteristics. For example, a twisted pair loop may include a notch filter to remove 50Hz or 60Hz power line noise, it may have peaks and valleys in the frequency response due to reflections from taps or poor terminations and it will drop off with frequency due to higher losses at higher frequencies.



Bandwidth

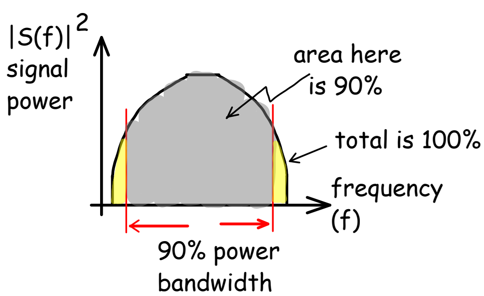
There are several definitions of bandwidth.

A common definition is the *3dB bandwidth*. This refers to the frequency range where the amplitude response is between 0 and -3dB relative to the maximum. Bandwidth definitions with values other than 3dB are also used (e.g. 6 dB or 10 dB bandwidth).



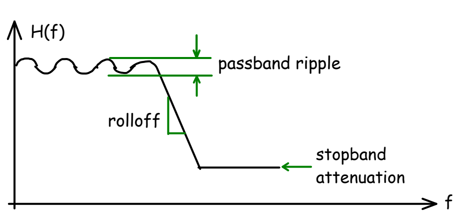
Exercise 2: How much *power* would a signal transmitted at the edge of the 3 dB bandwidth passband have compared to the the power it would have if transmitted at the frequency with the lowest loss? What would be the ratio of the *voltages*? What if the bandwidth was defined as the 6 dB bandwidth?

A definition of bandwidth that is useful for signals rather than channels is the *90% power bandwidth*. This is the frequency range that contains 90% of the signal power. Other values than 90% can be used (e.g. 99% or 99.9%).



Other definitions of bandwidth are used for specialized applications.

Bandwidth is a single number and cannot describe all aspects of the transfer function. Other specifications such as the steepness of the gain roll-off or the gain ripple in the passband are often important.



Phase Response

For a sinusoid:

$$x(t) = \sin(2\pi f t)$$

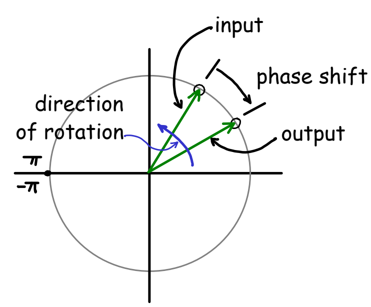
the phase ($2\pi f t$) increases linearly with time. The angular frequency ($2\pi f$) is the rate of change of phase with time.

The phase response of the channel is the difference in phase between the output and the input[1]. These phase shifts are often a result of delays through the channel.

Exercise 3: If t is a time, which happens first, t or $t-\delta$? If $f(t)$ and $f(t-\delta)$ are the input and output of a channel with delay, which is the input and which the output? What is the delay?

As you can see from the diagram below, delaying a sinusoid is equivalent to changing its phase. A time delay of τ introduces a phase shift of:

$$\Delta\theta = -2\pi f \tau \quad \text{radians}$$



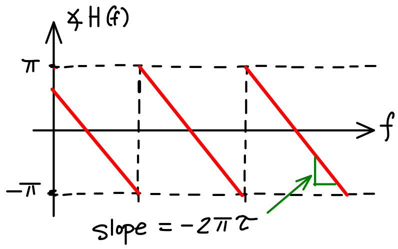
Note that a delay results in a smaller phase at the output and thus the phase shift is negative (output phase minus input phase).

Exercise 4: If the phase changes by $\frac{\pi}{2}$ and the frequency is 1 kHz, what is the delay? If the phase at the input is $\frac{\pi}{4}$, what is the phase at the output?

The phase shift is a linear function of the delay with -2π times the delay being the slope of a phase versus frequency curve.

Since phase "wraps" every 2π We cannot measure phase shifts outside the range of 2π. This causes the phase response appears to have discontinuities. We often plot the "wrapped" phase in the range $-\pi$ to π or 0 to 2π. The phase response for a channel with delay τ would be:

[1] To divide two complex values we divide the magnitudes and subtract the phases.



Exercise 5: A 100m transmission line has a velocity factor of 0.66. Plot the phase response of the cable over the frequency range 0 to 6 MHz.

Linear Phase Channels

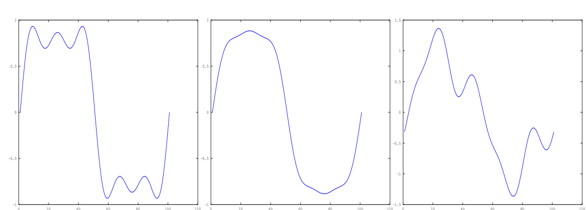
A signal that contains multiple frequency components[2] will be distorted if these frequency components are shifted in time with respect to each other. Therefore the delay introduced by the channel must be the same at all frequencies in order to avoid distortion.

From the equation relating phase shift and delay ($\Delta\theta = -2\pi f\tau$) we can see that a delay that does not vary with frequency will result in a linear phase response. Thus the following statements are equivalent:

- the phase response is linear
- all frequency components will be delayed the same amount

Channels must therefore have linear phase as well as constant amplitude response to avoid distortion.

The example below shows a pulse approximated by three sinusoids; the effect of attenuating the amplitudes of the higher-frequency components and the effect of delaying each of the frequency components relative to each other.



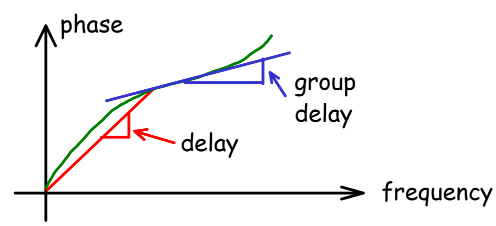
[2]Any signal other than a sinusoid has multiple frequency components.

Group Delay

When a channel is not linear phase we cannot define the delay through the channel since it is not the same at all frequencies. In other words, the phase versus frequency curve is not a straight line so we cannot measure its slope.

However, we can measure the phase versus frequency curve and calculate the slope at specific frequencies. We use the term "group delay" for this slope[3].

The following diagram demonstrates the difference between delay and group delay:



The differences in group delay for the various frequencies present in a signal determines the time dispersion of the signal and how much it will be distorted. If the group delay at all frequencies is the same (i.e. the phase versus frequency curve is a straight line over those frequencies) then all of the frequency components will see the same delay and the passband signal will not be distorted.

However, if the group delay (slope) changes over the frequency range of interest, then different frequency components will delayed by different amounts and the signal will be distorted.

Note that if the slope of the phase response curve has units of radians and the frequency axis has units of radians/second then the slope (the group delay) will have units of seconds.

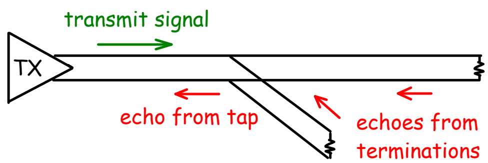
Exercise 6: A telephone line is being used to transmit symbols at a rate of 300 symbols/second using frequencies between 800 and 1200 Hz. If the group delay must be less than 10% of the symbol period, what is the maximum allowable difference in group delays over this frequency range?

Echo and Multipath

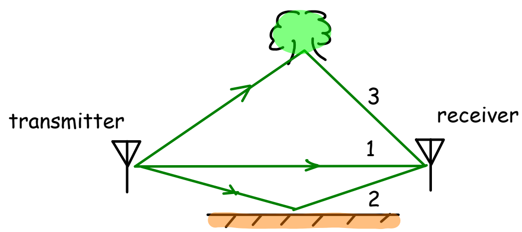
Another source of linear distortion is echoes and multipath propagation. Echoes can be due to trans-

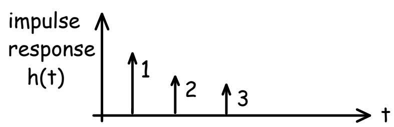
[3]The term derives from the relative delay for a group of frequencies.

mission lines that are tapped or not properly terminated.



Multipath propagation, a similar phenomenon, typically happens on wireless links with non-line of sight (NLOS) paths. Objects will reflect, diffract or scatter the radio signals.





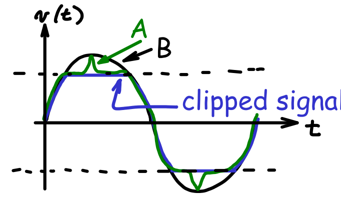
Since the each path length can be different, the delays for each path can be different. The different paths can add up constructively or destructively depending on the frequency and the delay. The frequency response can thus have peaks and nulls.

Non-Linear Distortion

Distortion caused by the amplitude and phase response of the channel are called linear distortions because they can be produced by linear combinations of (delayed versions of) the input signal. Linear distortions can be corrected, in principle, by applying a filter whose response, at each frequency, is the (complex) multiplicative inverse of the frequency response of the channel. It is important to understand that linear distortion can't change the frequencies present.

In practice there are limitations to how much linear distortion we can correct because the gain required to compensate for the attenuation will also amplify any noise that is present. Thus, in practice, large attenuations cannot be reversed.

There are some common distortions that are not linear. These cannot, in general, be undone. A typical example is clipping (the peaks of the signal are cut off) due to the limited dynamic range of an amplifier. For example, if the output of the amplifier is the clipped signal, we cannot determine if the input was waveform A or B:



Most amplifiers have some degree of non-linearity which causes non-linear distortion.

Unlike linear distortion, non-linear distortion creates new frequency components. The exact frequencies and levels of these components depends on the type of distortion and the frequencies present at the input.

THD

Non-linear distortion caused by baseband (low-pass) channels is usually measured as Total Harmonic Distortion (THD). THD is measured using a single sine wave as the input and measuring the powers of the fundamental frequency and harmonics at the output (e.g. using a 'scope with FFT function). THD is computed as:

$$\mathrm{THD} = \sqrt{\frac{P_1 + P_2 + P3 + \ldots}{P_0}}$$

which is the square root of the ratio of the sum of the output powers at the harmonic frequencies (the P_i's) to the output power at the input frequency (P_0)[4].

IMD

Non-linear distortion of narrow-band channels is usually measured as "two-tone intermodulation distortion" (IMD). IMD is measured using two closely-spaced signals as the input and measuring the power of the "intermodulation products" which appear at frequencies that are the sums and differences of multiples of the input frequencies.

For example, if the two frequencies are f_1 and f_2, there will be components at frequencies $\pm nf_1 \pm mf_2$.

[4] Unfortunately, there are other definitions of THD. Some include the powers of all frequency components in the denominator. Some omit the square root.

The component most often measured is the third-order distortion product because the two frequencies, at $2f_1 - f_2$ and $-f_1 + 2f_2$, fall near the frequencies of the original signals (f_1 and f_2). On the other hand, harmonics appearing at multiples of the original frequencies ($n \times f_1$ and $n \times f_2$) are relatively easy to filter out.

Exercise 7: The input to a non-ideal amplifier is the sum of two sine waves at frequencies of 1 and 1.1 MHz. What are the frequencies of the harmonics of these frequencies? What are the frequencies of the (positive) third-order IMD products?

Noise and SNR

Noise is a unpredictable ("random") signal that is added to the desired signal. Noise can be added by the channel or by the receiver.

Some sources of noise include the thermal noise that is present in any resistor at temperatures above 0 K, "shot" noise generated by semiconductor devices, electrical equipment such as motors and some lights, lightning and the sun.

Noise is the phenomenon that ultimately limits the performance of any communication system. Noise may cause errors in digital communication system or degrade the quality of an analog signal.

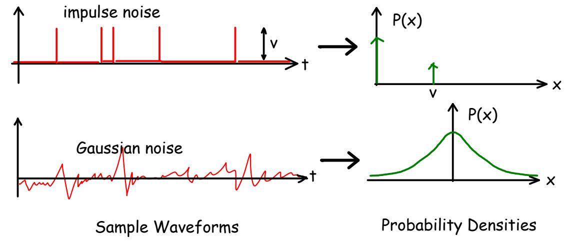
One important metric is the signal-to-noise ratio (SNR) which is the ratio of signal power to noise power.

Exercise 8: A sinusoidal signal is being transmitted over a noisy telephone channel. The voltage of the signal is measured with an oscilloscope and is found to have a peak voltage of 1V.

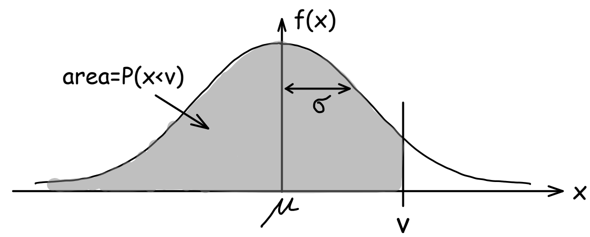
Nearby machinery is adding noise onto the line. The voltage of this noise signal is measured with an RMS voltmeter as 100mVrms. The characteristic impedance of the line is 600Ω and it is terminated with that impedance. Why was an RMS voltmeter used? What is the signal power? What is the noise power? What is the SNR?

Gaussian Probability Distribution

In addition to the distribution of the noise power in frequency, we also need to know how the voltage is distributed. For example, impulse noise has only two voltage levels (zero and the peak value of the impulse).



Signals that result from the sum of many small independent events have a probability distribution known as a Gaussian distribution. In communication systems this usually happens due to the sum of of voltages produced by the actions of very many individual photons, electrons or molecules.



The Gaussian distribution is the familiar "bell curve" or "normal" distribution. The probability is a maximum at the average value and drops off to smaller probabilities at larger or smaller voltages.

The Gaussian distribution is determined by two values: the mean (μ) and the variance (σ^2).

It's often useful to know the probability that the voltage of a Gaussian noise signal will exceed a certain voltage. If the noise signal x has a DC (mean) value μ and an RMS AC (zero-mean) voltage σ then the probability that the noise voltage is less than v is given by the Gaussian (Normal) cumulative distribution function (CDF). This is the area under the Gaussian distribution curve to the left of (less than) the value v.

Exercise 9: Would you use AC or DC coupling to measure: (a) σ, (b) μ, and (c) the RMS power? Would you measure the average or RMS power in each case? What is the RMS power of the signal x if it has a mean (DC) value of $\mu = 2$ V and $\sigma = 3$ V?

The plot in Figure 1 shows the shape of the Gaussian density function and also gives the cumulative probabilities along a second x-axis.

To find the probability that the voltage is greater than v we can use the fact that the sum of all probabilities is 1. Thus $P(x > v) = 1 - P(x \leq v)$.

To compute $P(v)$ we first compute a normalized value, t by subtracting the mean, μ, and dividing by the standard deviation, σ, of the distribution:

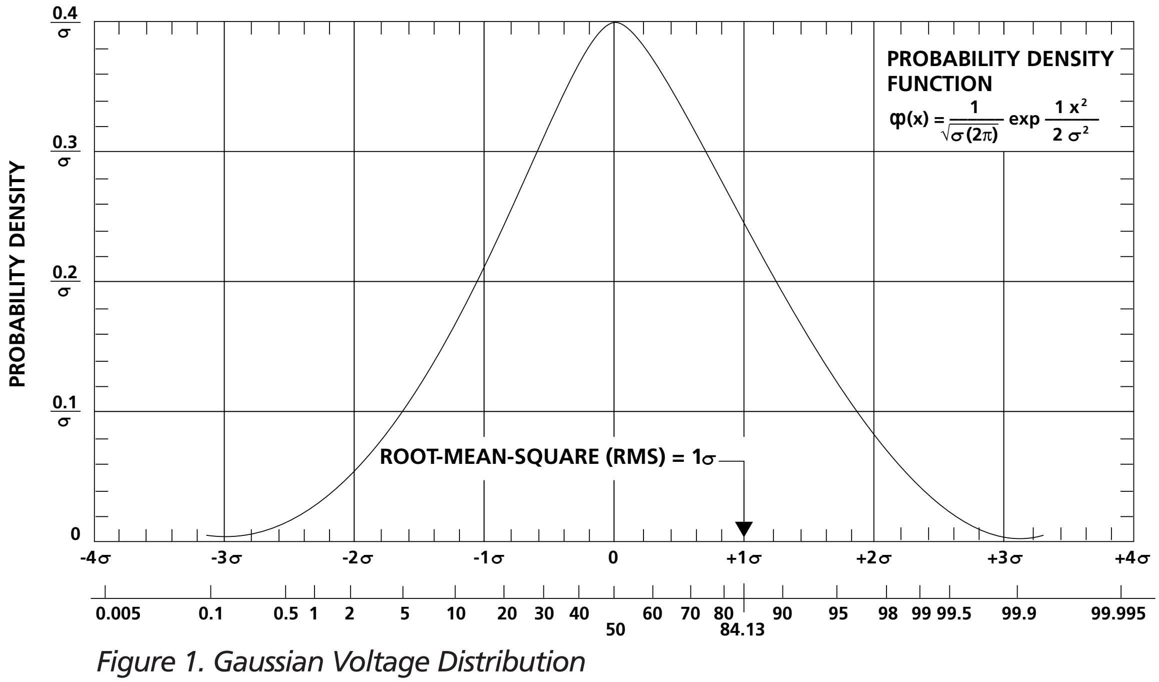


Figure 1: Gaussian density function and values of the cumulative distribution.

$$t = \frac{v - \mu}{\sigma}$$

Exercise 10: What are the units of t?

Exercise 11: The output of a noise source has a Gaussian (normally) distributed output voltage. The (rms) output power is 20mW and the output impedance is 100Ω. What fraction of the time does the output voltage exceed 300mV? Hint: the variance (σ^2) of a signal is the same as the square of its RMS voltage.

Some calculators will compute this ($P()$ function) or you can use the figure below..

Interference

Interference is a random man-made signal that has a similar effect as additive noise.

Interference may be caused by communication systems or other electrical devices. An interfering communication system may be unrelated or it may be part of the same system.

Examples of interference include:

- the signal from a cell phone may couple onto a audio cable and cause a "buzzing" noise (known as RFI - Radio Frequency Interference)
- at night the signal from a remote AM broadcast station may reach further than usual and cause interference to a local station.
- two WLAN cards may decide to transmit a packet at the same time resulting in neither one being received correctly (a "collision")
- the commutator on an electric motor may cause interference to a radio receiver
- the clock signals within a PC may cause interference to a TV receiver (known as EMI, Electro-Magnetic Interference)

Wireless systems are particularly vulnerable to interference because of the wide difference in the levels of the transmitted and received signals.

Power ratios similar to SNR that include the effect of interference include:

- SIR - **S**ignal to **I**nterference **R**atio
- SINR - **S**ignal to **I**nterference plus **N**oise **R**atio

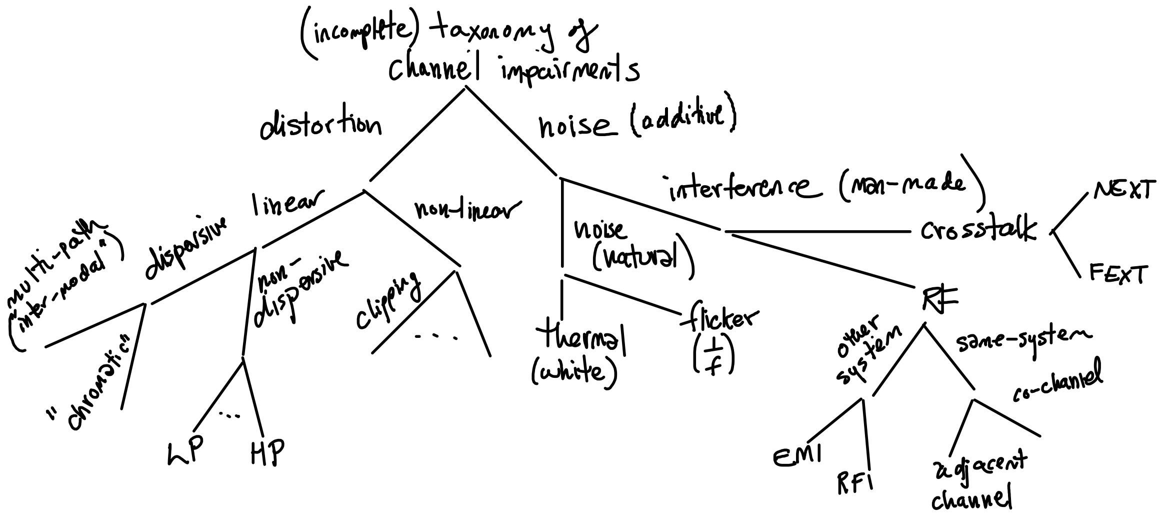


Figure 2: Taxonomy of Channel Impairments.

- SINAD - **S**ignal to **I**nterference plus **N**oise **A**nd **D**istortion **R**atio

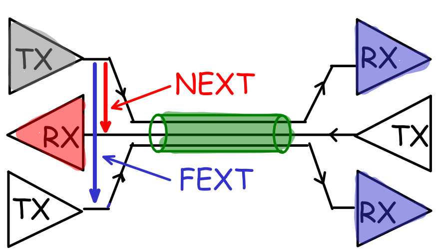
Crosstalk

Crosstalk is interference due to coupling between conductors that are run in the same cable. For example, telephone loops are often carried in cables with 25 or more pairs. There will be some coupling between the pairs. Signals on one pair can thus leak into another pair in the same cable. Another example is the coupling between pairs used in LAN UTP cables. Crosstalk will be affected by shielding, twist patterns and other design details. Datasheets will often specify crosstalk at different frequencies.

We can distinguish between:

- Near-end crosstalk (NEXT) is the leakage of the signal being transmitted onto the signal being received at the same location.

- Far-end crosstalk (FEXT) is the leakage of the signal being transmitted onto the signal being received at the other end of the link.



Note that the coupling between pairs in a cable happens throughout the length of the cable. The difference between NEXT and FEXT is in which receiver is affected (the near one or the far one), not where the coupling takes place.

Some DSL systems measure the coupling between pairs and subtract out the crosstalk. "Alien crosstalk" refers to crosstalk from other cables. It is usually not possible to cancel out this crosstalk.