Data Transmission over Bandlimited Channels

This chapter describes limits on the maximum symbol and information rate for band-limited channels.
After this chapter you should be able to: determine if a channel meets the Nyquist no-ISI criteria and, if so, the maximum signalling rate without ISI; determine the maximum error-free information rate over the BSC and AWGN channels; determine the specific conditions under which these two limits apply. You should be able to perform computations involving the OFDM symbol rate, sampling rate, block size and guard interval.

Introduction

All practical channels are band-limited – either low-pass or band-pass. There are two theorems, the Nyquist no-ISI criteria and Shannon's capacity theorem, that provide some guidance about maximum data rates that can be achieved over bandlimited channels.

Inter-Symbol Interference

Low-pass channels attenuate the higher-frequency components of a signal. This increases the rise and fall times of pulses and thus extends ("stretches") their durations.

The principle of superposition applies to linear time-invariant channels[1]. This means the output for a sum of inputs (pulses) is the sum of the corresponding outputs (the "stretched" pulses). We can thus compute the response of the channel to a sum of pulses by computing the sum of the "stretched" output pulses. These "stretched" output pulses will interfere with each other. This effect is called inter-symbol interference (ISI).

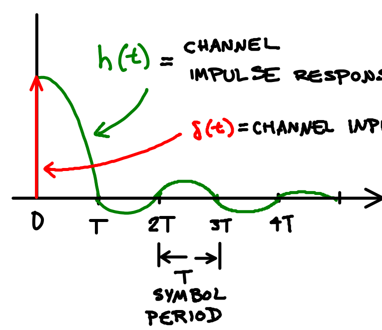
Exercise 1: Draw a square pulse of duration T and amplitude 1. Draw the pulse after it has passed through a linear low-pass channel that causes the signal to appear at the output delayed by $T/2$ and attenuated by $\frac{1}{2}$. Draw the output for an input pulse of the opposite polarity. Use the principle of superposition to draw the output of the channel for a positive input pulse followed by a negative input pulse. Have the pulses been distorted?

[1]This is a common assumption as we typically restrict voltage levels to those for which the channel is approximately linear and the channel typically does not change over intervals of many symbols.

Nyquist no-ISI Criteria in Time

Consider a system that transmits symbols every T seconds that are infinitely-short pulses ("impulses"). Low-pass channels that meet certain conditions can "stretch" the impulses into a "damped sinusoid" as shown below.

If the period of this "ringing" is one symbol period (T), then the ISI created by the impulse will add zero ISI to subsequent impulses. This is the Nyquist no-ISI condition.



Exercise 2: What is the impulse response of a channel that does not alter its input? Does this impulse response meet the Nyquist condition? Will it result in ISI?

An example of an impulse response that meet this criteria is[2]:

$$h(t) = \text{sinc}(\frac{t}{T}) = \frac{\sin(\pi t/T)}{\pi t/T}$$

which has value 1 at $t = 0$ and 0 at multiples of T.

Exercise 3: Draw the impulse response of a channel that meets the Nyquist condition but is composed of straight lines. Note that there are many such impulse responses.

Exercise 4: What causes the sinc() function to have periodic zero-crossings? What causes the amplitude to decay?

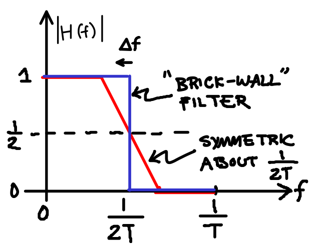
[2]The (normalized) sinc() function defined as $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$.

Nyquist no-ISI Criteria in Frequency

From the Nyquist no-ISI condition in the time domain it's possible to derive an equivalent conditions in the frequency domain. A common way to state this condition is that the channel's frequency response must have odd symmetry around half of the symbol frequency[3] ($\frac{1}{2T}$):

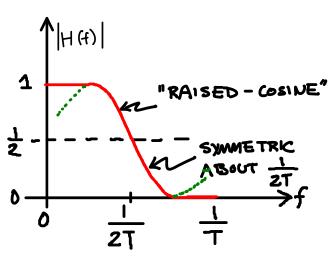
$$H(\frac{1}{2T} + \Delta f) + H^*(\frac{1}{2T} - \Delta f) = 1 \text{ for } |\Delta f| \leq \frac{\pm 1}{2T}$$

Just as there could be many impulse responses that are zero at multiples of the symbol period, there are many channel frequency responses that result in no ISI. For example, the following two straight-line transfer functions meet the no-ISI condition[4]:



The "brick-wall" filter (blue) has a response that is 1 below half of the symbol rate ($\frac{1}{2} \cdot \frac{1}{T} = \frac{1}{2T}$ where $\frac{1}{T}$ is the symbol rate) and zero above that. Although such a filter would have the minimum bandwidth required to meet the Nyquist condition for a symbol period T, it cannot be built because the delay through the channel would be infinite.

The most commonly-used transfer function is one that approximates the "raised-cosine" function which is a half-cycle of a cosine function offset vertically and scaled horizontally to have a minimum value of zero and centered around half of the symbol rate:



[3]The asterisk indicates complex conjugate. This can be ignored for real(izable) baseband channels.

[4]For simplicity we only show one component (the real or imaginary portion) of the transfer function.

Note that it is the symmetry around the frequency $\frac{1}{2T}$ that ensures there will be no ISI rather than the exact filter shape. Thus we are free to implement other transfer functions if they make the implementation easier.

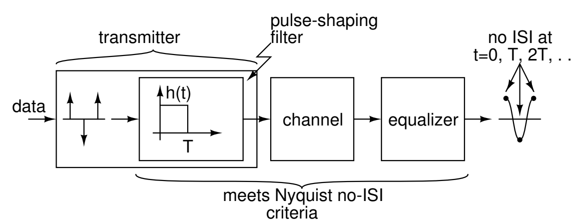
Exercise 5: Draw the magnitude of a raised-cosine transfer function that would allow transmission of impulses at a rate of 800 kHz with no interference between the impulses.

Typically we have no control over the transfer function of the channel. However, we can add filtering at the transmit or receive sides of the channel so that the overall transfer function meets the Nyquist criteria. This is called "equalization."

Pulse-Shaping Filter

Note that the no-ISI criteria ensures that a channel produces no ISI when transmitting *impulses*, not for the square pulses typically used by line codes.

However, we can imagine that the transmitter includes a filter that converts impulses to pulses. We can add an equalizer that corrects for both the channel and this imaginary impulse-to-pulse conversion filter:



Exercise 6: Draw the impulse response of a filter than converts input impulses to pulses of duration T? Draw the signal after the pulse-shaping filter in the diagram above.

Equalization

To avoid ISI, the overall total channel response must meet the Nyquist no-ISI condition.

As mentioned above, this is usually not the case and the transmitter and/or receiver must use a filter called an equalizer that modifies the overall transfer function to ensure the no-ISI condition is met.

Transmitter and receiver filters typically have other functions beside equalization. For example, the transmit filter may limit the bandwidth of the transmitted signal to limit interference to users on adjacent channels. The receiver filter may filter out noise and interference from adjacent channels and thus improve the SNR and SIR. The communication system designer would design the transmitter and receiver filters to meet both the filtering and equalization requirements.

A common situation is a channel whose impulse response is an impulse (a "frequency-flat" channel) such as a coaxial cable or satellite channel. In this case a reasonable approach is to put half of the filtering at the transmitter and half at the receiver. In order to achieve an overall raised cosine transfer function, each side has to use a "root raised cosine" (RRC) transfer function. The product of the two filters is thus the desired raised-cosine function which meets the no-ISI condition.

Equalizers also have to compensate for the (imaginary) pulse-shaping filter. Since the pulse-shaping filter has a low-pass (sinc(f)) shape, the equalizer response has more gain at higher frequencies that a true raised-cosine function [5].

Nyquist Criteria and Bit Rate

Note that the symbol rate limitations defined by the Nyquist criteria do not determine the maximum *bit* rate that can be transmitted over a channel – they only determine the maximum *symbol* rate that can be transmitted *without ISI*.

We can increase the bit rate while keeping the symbol rate fixed by increasing the number of bits per symbol (e.g. by using multiple levels). For example, using symbols each of which could be at one of 1024 levels we can transmit 10 bits per symbol.

Exercise 7: A channel has a 3 kHz bandwidth and meets the Nyquist non-ISI conditions with $\alpha = 1$. How many levels are required to transmit 24 kb/s over this channel using multi-level signalling?

We can also design receivers that recover the transmitted data even when the channel introduces ISI. For example, if we know the symbols that have been transmitted in the past and we know the channel impulse response then we can predict the ISI and subtract it from the current received symbols. This is called decision-feedback equalization (DFE).

[5] Sometimes called "sinc compensation."

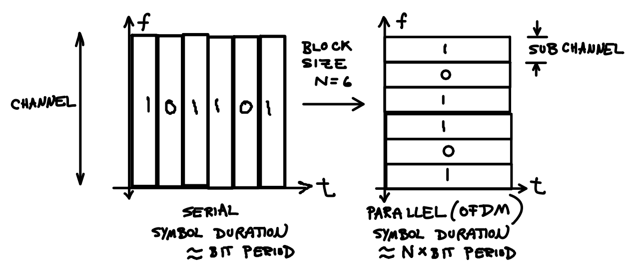
Adaptive Equalizers

In many communication systems the transfer function of the channel cannot be predicted ahead of time. One example is a modem used over the public switched telephone network (PSTN). Each phone call will result in a channel that includes different "local loops" and thus different frequency responses. Another example is multipath propagation in wireless networks. The channel impulse response changes as the receiver, transmitter or objects in the environment move around.

To compensate for the time-varying channel impulse response the receiver can be designed to adjust (adapt) the receiver equalizer filter response using various algorithms.

OFDM

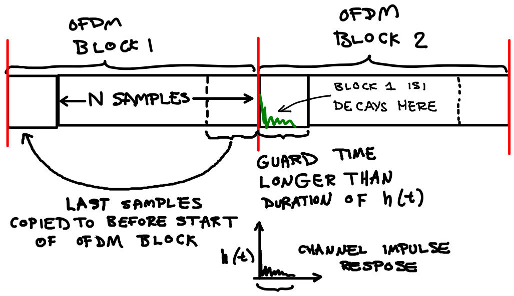
An alternative to equalization is a technique called Orthogonal Frequency Division Multiplexing (OFDM). An OFDM transmitter collects a group of N symbols at a time and uses them to modulate N "subcarriers" (the process of modulation is covered in another course). These subcarriers are transmitted in parallel over the same time duration that would have been required to transmit the N symbols serially. The net effect is to reduce the symbol rate by a factor N but with no impact on the overall bit rate.



We usually insert a "guard time" (or "guard interval") in-between symbols. Its duration is longer than the duration of the channel impulse response.

Since no data is transmitted during the guard time, this reduces the average data rate. However, the guard time is typically a small fraction of the OFDM symbol duration and so the impact on the overall throughput is relatively small.

Rather than transmitting nothing during the guard interval, a small part of the end of the block of N samples is copied to the start of the symbol and transmitted during the guard time. This is called a "cyclic" or "periodic" extension.



The value of N is typically a power of 2 to allow efficient implementation using the Fast Fourier Transform (FFT) algorithm.

OFDM has become more popular than adaptive equalization because it is simpler to implement and more robust. This is partly because it is not necessary to estimate the channel to correct for ISI. OFDM is used by most ADSL, WLAN and 4G cellular standards.

Exercise 8: The 802.11g WLAN standard uses OFDM with a sampling rate of 20 MHz, with $N = 64$ and guard interval of $0.8\mu s$. What is the total duration of each OFDM block, including the guard interval? How long is the guard time?

Shannon Capacity

The Shannon Capacity of a channel is the information (not bit) rate below which it is possible to transmit data with an arbitrarily low error rate.

An example of a channel is the Binary Symmetric Channel (BSC). This channel transmits discrete bits (0 or 1) with a bit error probability (BER) of p. The capacity of the BSC is measured in units of information bits per "channel use" (a transmitted bit) and is:

$$C = 1 - (-p \log_2 p - (1-p) \log_2 (1-p))$$

which is 0 for $p = 0.5$ (when each transmitted bit is equally likely to be received right or wrong) and 1 when $p = 0$ (the error-free channel) or when $p = 1$ (a perfectly inverting channel).

Exercise 9: What is capacity of a binary channel with a BER of $\frac{1}{8}$ (assuming the same BER for 0's and 1's)? *Hint:* $\log_2(\frac{7}{8}) \approx -0.2$.

For a continuous channel corrupted by Additive White Gaussian Noise (AWGN), the capacity can be shown to be:

$$C = B \log_2 \left(1 + \frac{S}{N}\right)$$

where C is the capacity (b/s), B is the bandwidth (Hz) and $\frac{S}{N}$ is the signal to noise (power) ratio.

Exercise 10: What is the channel capacity of a 4 kHz channel with an SNR of 30dB?

The Shannon limit does not say that you can't transmit data faster than this limit, only that if you do, you can't reduce the error rate to an arbitrarily low value. However, in practice, attempting to transmit at information rates above capacity results in high error rates.

Exercise 11: Can we use compression to transmit information faster than the (Shannon) capacity of a channel? To transmit data faster than capacity? Explain.

Shannon also showed that the error rate can be made arbitrarily low by increasing the codeword size (n) of error-correcting codes. But this is only possible if we limit the information rate to less than the channel capacity.

Systems using modern forward error-correcting (FEC) codes that use large codewords, such as Low Density Parity Check (LDPC) codes, can communicate at very low error rates over AWGN channels that have SNRs only a fraction of a dB more than the minimum required by the capacity theorem.

Exercise 12: What do the Nyquist no-ISI criteria and the Shannon Capacity Theorem limit? What channel parameters determine these limits?