Data Transmission over Bandlimited Channels

In many cases the channel bandwidth limits the symbol or data rate. This lecture describes two different ways to estimate the symbol or data rate that can be transmitted over a band-limited channel.
After this lecture 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 an AWGN channel; 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

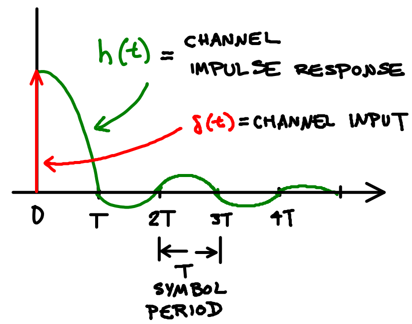
All practical channels are band-limited (either low-pass or band-pass) and the channel bandwidth is often what limits the maximum data rate. We will study two theorems, the Nyquist no-ISI criteria and Shannon's capacity theorem, that provide some guidance about maximum data rate that can be achieved over a bandlimited channel.

Inter-Symbol Interference

Bandwidth-limited channels attenuate the higher-frequency components of a signal. This increases the rise and fall times of waveforms and "smears" the signal in time. The channel thus extends the duration of each transmitted symbol in time so that each signal extends into and interferes with subsequently-transmitted symbols. This means there is a possibility that symbols will interfere with subsequently-transmitted symbols. This interference is called inter-symbol interference (ISI).

Nyquist no-ISI Criteria in Time

Consider a system that transmits symbols as (infinitely-)short pulses ("impulses"). A low-pass channel will limit the rise time of the pulses and cause the impulses to be smeared out in time. However, if the response of the channel to these impulses crosses zero at multiples of the symbol period then the impulses will not interfere with each other. This is called the Nyquist no-ISI criteria.



An example of an impulse response that meet this criteria is the $\text{sinc}()$ function:

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

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

Nyquist no-ISI Criteria in Frequency

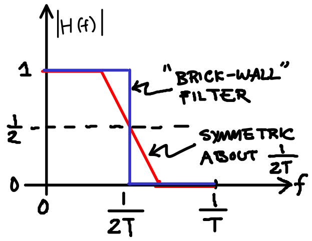
It is possible to derive the characteristics of the channel's frequency-domain transfer function that result in no ISI. This condition is that the channel frequency response have odd symmetry around half of the symbol frequency:

$$H(\frac{1}{2T} + f) + H(\frac{1}{2T} - f) = 1 \text{ for } 0 < |f| < \tfrac{1}{2T}$$

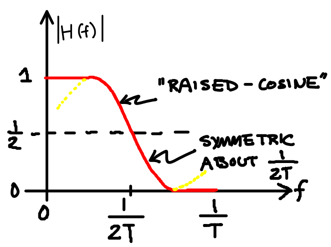
Note that this condition applies to the complex frequency response. Thus both the real and imaginary parts of $H(f)$ need to have this symmetry.

Just as there could be many impulse responses that are zero at multiples of the symbol period, there are many no-ISI transfer functions. For example, the following two straight-line transfer functions meet the no-ISI condition[1]:

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



The “brick-wall” filter (blue) has a response that is 1 below half of the symbol rate ($\frac{1}{2T}$) and zero above that. Although such a filter would have the minimum overall bandwidth required for a symbol period T, it is not physically realizable and has other problems as described below. The filter with the linear transfer function is more practical but still difficult to implement. A more practical transfer function is the so-called raised-cosine function which is a half-cycle of a cosine function offset to have a minimum value of zero and centered around half of the symbol rate:



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

Exercise 1: Draw the (real portion of) a raised-cosine transfer function that would allow transmission of impulses at a rate of 800 kHz with no interference between the impulses.

Pulse-Shaping Filter

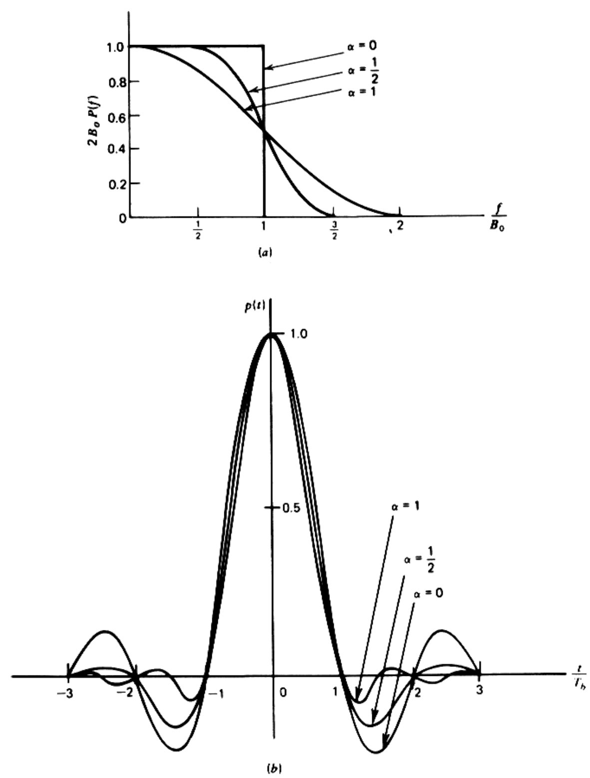
Note that the no-ISI criteria applies for a channel that produces no ISI for *impulses*, not the square pulses typically used. Since practical systems don’t transmit impulses, the Nyquist criteria cannot be applied directly to the physical channel itself. Instead, we consider that the transmitter includes a hypothetical filter that converts impulses to pulses before transmitting them over the channel. The response of this (im)pulse-shaping filter has to be included when evaluating the channel ISI. It is actually the combination of this impulse-shaping filter and the channel that has to meet the Nyquist criteria.

Exercise 2: What is the impulse response of a filter than converts input impulses to pulses of duration T? What is the shape of the frequency response of this filter? *Hint: the Fourier transform of a pulse of duration T is* $\frac{sin(\pi T)}{\pi T}$. What is the “bandwidth” of this filter (when is it first zero)? How does this compare to the “bandwidth” of the raised-cosine filter above?

Excess Bandwidth

Channels can have different transitions between passband and stopband of the transfer function while still meeting the no-ISI conditions. However, the width of this transition has an impact on the shape of the impulse response and on the sensitivity of the receiver to errors in the timing of the sampling point.

This parameter, α, is called the “excess bandwidth”. The following diagram[2] shows how the excess bandwidth parameter for a raised-cosine transfer function affects the impulse response.



[2]From Simon Haykin, “An Introduction to Analog and Digital Communication”, 1989.

Larger values of excess bandwidth (wider bandwidth channels) results in less "ringing" of the impulse response which in turn reduces the amount of ISI near the sampling point of the next symbol. This makes the receiver less sensitive to errors in its timing of the sampling point.

Exercise 3: What is the possible range of values of α?

Equalization

To avoid ISI, the total channel response including the pulse-shaping filters, transmit filters, the channel and the receiver filter(s) have to meet the Nyquist no-ISI condition.

When the channel by itself is unlikely to meet the no-ISI conditions, the transmitter and/or receiver use filters, known as "equalizers" that modify 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 adjacent channels. The receiver filter may filter out noise and interference from adjacent channels and thus improve the SIR and SNR. The communication system designer would design the transmitter and receiver filters to meet both the filtering and equalization requirements.

A common situation is a flat channel where interference is not an issue. 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.

Exercise 4: Could equalization be done at the receiver only? At the transmitter only? Why or why not?

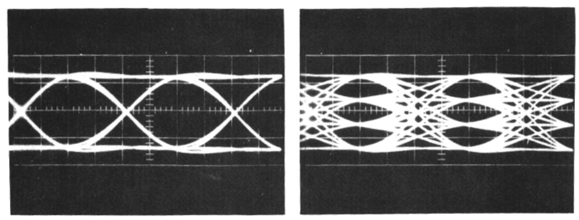
Adaptive Equalizers

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 "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 the receiver equalizer filter response using various algorithms.

Eye Diagrams

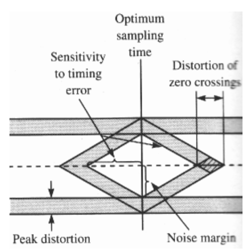
An eye diagram is superimposed plots of duration T (the symbol period) of the received waveform for random data. The eye diagram graphically demonstrates the effect of ISI. Some examples of eye diagrams produced by an appropriately-triggered oscilloscope[3]:



BINARY QUATERNARY

The vertical opening at the sampling time, called the "eye opening", represents the amount of ISI at the sampling point.

The horizontal opening indicates how sensitive the receiver would be to errors in sampling point timing[4]:



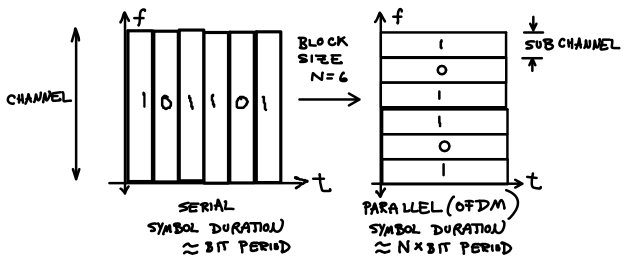
OFDM

An alternative to equalization is a technique called Orthogonal Frequency Division Multiplexing

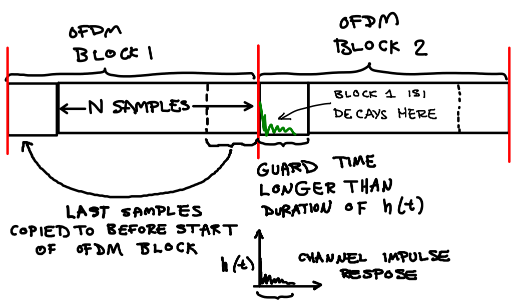
[3] From John G Proakis, "Digital Communications", 3rd Ed., 1983.

[4] Proakis, op. cit.

(OFDM). An OFDM transmitter gathers blocks of N consecutive symbols and uses them to modulate N "subcarriers" (modulation will be discussed later). The net effect is to reduce the symbol rate by a factor N. The value of N is typically a power of 2 to allow efficient implementation using Fast Fourier Transforms (FFTs).



The reduction in symbol rate (or increase in symbol period) reduces the impact of ISI because the impulse response of the channel is now a shorter fraction of the symbol period. Most OFDM systems also insert a "guard time" (or "guard interval") between symbols that is longer than the duration of the impulse response of the channel. This eliminates interference between symbols. To do this a small number of the final samples of each block of N samples are copied to the start of the symbol and transmitted during the guard time. Because the block of N samples is periodic this is called a "cyclic" or "periodic" extension.



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 modern ADSL, WLAN and 4G cellular standards.

Exercise 5: 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 many guard samples are used?

Shannon Capacity

The Shannon Capacity of a channel is the information rate above which it is not possible to transmit data with an arbitrarily low error rate. For the Additive White Gaussian Noise (AWGN) channel the capacity is:

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

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

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.

Shannon's work also does not specify how to achieve capacity, for example, what modulation and coding should be used. However, Shannon's work does hint that using error-correcting codes should allow us to achieve arbitrarily-low error rates as long as we limit the data (actually, information) rate to less than the channel capacity.

Exercise 6: What is the channel capacity of a 3 kHz channel with an SNR of 20dB?

Implementing systems that operate at close to channel capacity requires coding. Some systems using modern codes such as Low Density Parity Check (LDPC) codes can operate within a fraction of a dB of channel capacity.

Note that the symbol rate limitations defined by the Nyquist criteria do not limit the achievable bit rate or determine the capacity of the channel. For example, we can use arbitrarily large symbol sets to increase the bit rate without increasing the symbol rate.

The use of PR signalling or sequence estimation also allow us to transmit arbitrarily high symbol rates over channels that don't meet the Nyquist no-ISI criteria. Whether the symbols can be recovered without errors will depend on the SNR.

Also note that the Shannon capacity refers to *information* rate (the bit rate after maximum possible compression), not to the bit rate.

Exercise 7: What are some differences between the signalling rate limit imposed by the Nyquist no-ISI criteria and the Shannon Capacity Theorem?