Université Concordia University

Department of Electrical & Computer Engineering

ELEC 691X/498X – Broadcast Signal Transmission Fall 2015

Instructor: Dr. Reza Soleymani, Office: EV-5.125, Telephone: 848-2424 ext.: 4103.

Office Hours: Wednesday, Thursday, 14:00 – 15:00

Time: Wednesday, 5:45 to 8:15

Room: H 521

Lecture 5:
Digital Modulation

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

In this lecture we cover the following topics:

- Digital Modulation,
- Different Modulation Schemes: BPSK, QPSK, MPSK, QAM, APSK.
- Pulse shaping,
- Performance of Modulation Different Modulation Schemes:
- -Bandwidth Requirement; Bandwidth Efficiency,
- - Bit Error Rate (BER),
- - Bandwidth/BER Trade-off.
- Multiplexing and Multiple Access: FDM/FDMA. TDM/TDMA, CDMA, OFDM/OFDMA.

Lecture 5: Digital Modulation

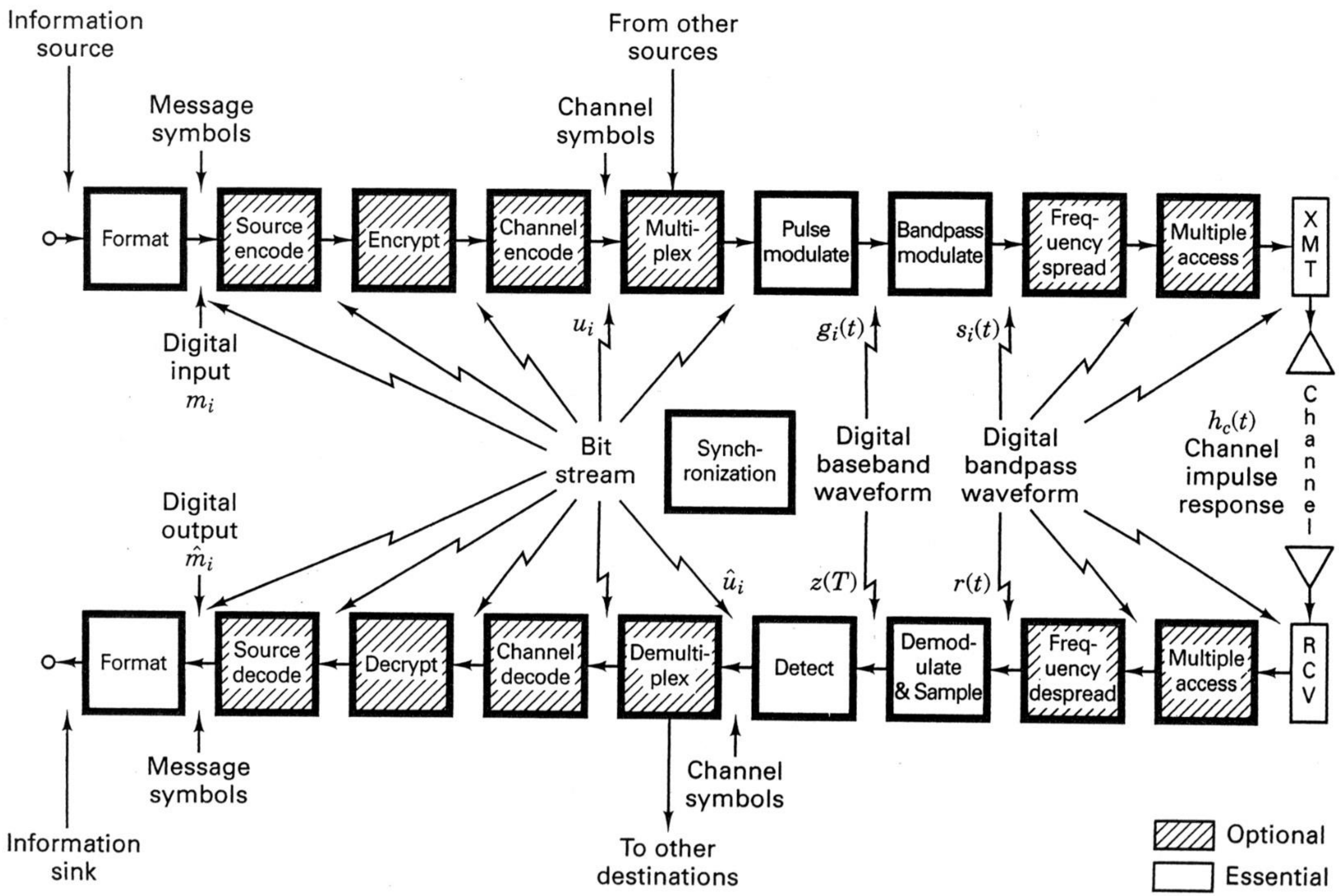
Department of Electrical & Computer Engineering

Let's first review some of the things we have learned so far. A broadcasting link is basically a one-way (or highly asymmetric) communications link. This means that the goal of a broadcast system is mainly to send content from the center to the users and the return channel is either non-existent or very narrowband used for feedback.

The block diagram on the next page shows all the components that may be used in a communication link and, as a consequence in a broadcasting link. It si important to note that not all of these functions are mandatory. Even the modulator and demodulator may not be used. This in the case of broadcasting in time (storing the video, say on a blue-ray disk).

Block Diagram of a Communications Link
(source: Bernard Sklar, Digital Communications: Fundamentals and Applications, Prentice Hall, 2001)

Concordia University
Department of Electrical & Computer Engineering



Block Diagram of a Communications Link

The first block (formatter) is basically responsible for turning the source into (raw) digital format. The case of digital video, it performs the functions such as sampling, quantization and adding some synchronization signals. The format in which the data will be at the output of the formatter (video capture) is usually in SDI format.

The next block compresses the raw video using a scheme such as MPEG. The out put at this point will consist of Transport Stream (TS) packets each 188 (204 if the RS code is used) bytes long. These packets are either formatted as an ASI or IP stream.

Encryption may be used either for personal purposes or Digital Rights Management (DRM). In any case, the goal is to prevent an unauthorized person from using the video.

Concordia University
Department of Electrical & Computer Engineering

The Channel Encoding block is responsible for helping the receiver in detecting and possibly correcting the error caused by the channel noise or storage media defects (for example damaged video disk). A communication/broadcasting system may have several levels of error control coding. These may appear at different points in the chain. We have already talked about the use of Reed Solomon (RS) codes. The16 parity bytes added to each 188 byte TS packet helps the decoder to correct up to 8 bytes of error. The channel may be noisier and create more errors. Then either the power of the transmitter has to be increased or the stream be further encoded, i.e., more parities be added prior to transmission. A class of codes used is LDPC codes. These are very performant, easy to decode with manageable encoding complexity.

Block Diagram of a Communications Link

The role of the multiplexer is to put together two or more streams. This is needed if one broadcasts more than one program. Nowadays, the multiplexing function is more and more performed through IP Encapsulation. This means that several video streams are put together as a single IP stream. Each IP packet may then contain TS packets belonging to different video sources. A header added to each TS packet allows the demultiplexer at the receiver side to separate the different programs.

Note that the multiplexer at this point, i.e., before transmission although doing similar function should not be confused by multiplexing done at different points of the broadcast chain, e.g., the one used for adding video, audio and metadata files to form a complete video stream.

Concordia University
Department of Electrical & Computer Engineering

The last thing we will talk about is the modulator. A digital modulator consists of two parts: the Pulse Shaping part (one you may associate with a digital electronic designer) and a Radio Frequency section (the RF section that is designed by an RF engineer).

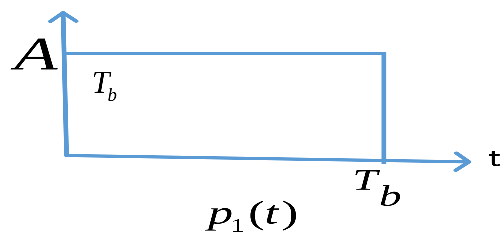
The function of the pulse shaping circuit is to apply a filter to the outgoing information so that the adjacent bits do not overlap, i.e., there is no Inter-Symbol Interference (ISI) and a channel does not go beyond its allocated frequency slot causing interference to other channels (ICI).

Before continuing with the details of modulation, I should point out that all the blocks at the receiver are the counterparts of the ones at receiver, inverting their function.

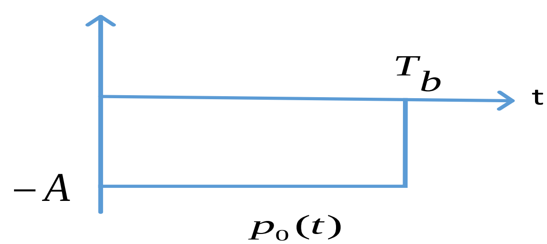
Concordia University
Department of Electrical & Computer Engineering

To see what pulse shaping means, let's start with a simple modulation scheme: a binary modulation format where 0's and 1's are transmitted using pulses of duration T_b seconds.

Say a 1 is represented by:

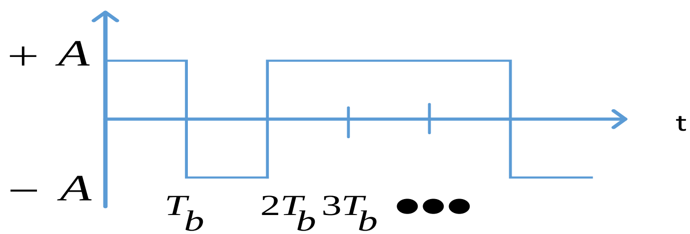


and a 0 is represented as:



UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

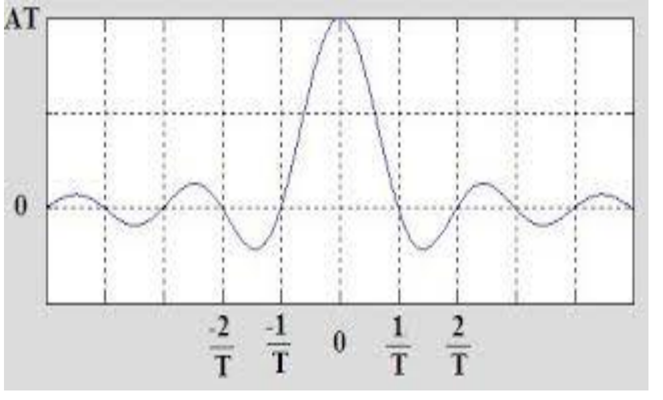
The data stream 101110, for example, will then be transformed into the following waveform:



This type of signaling using sharp waveforms (if feasible) can sole the problem of ISI since pulses corresponding to adjacent bits do not have any overlap. The problem, however, is the fact that a pulse in time-limited signal will have a spectrum expanding over all frequencies. The other problem is that the sharp transitions of these pulses cannot be physically realized.

Université Concordia University
Department of Electrical & Computer Engineering

The spectrum (Fourier Transform) of the pulse shape extending from 0 to T_b is:



It is seen that the spectrum extends from $-\infty$ to $+\infty$. Of course we can take only a few sidelobes and for example say that the require bandwidth is $\frac{k}{T_b}$ or even accept quite a lot of ICI and only take the main lobe and assign a bandwidth of $\frac{1}{T_b}$. But this is still much more than we will see is needed.

Digital Modulation
The duality property

Concordia University
Department of Electrical & Computer Engineering

- *You may probably still remember the duality theorem you have learned in your Signal and System course. If not, I strongly suggest that you review the material about LTI systems, Fourier Transforms, etc. referring to the text used for ELEC342:*

  Continuous and Discrete Time Signals and Systems by Mrinal Mandal and Amir Asif. Cambridge University Press,

Or any text on Signals and Systems. There are also a lot of material online, particularly on Wikipedia.

For example, I copied the contents of the next slide from the Internet.

The main point is that since the Fourier Transform (FT) of a square pulse in time domain is a sinc function in the frequency domain, then the FT of a sinc pulse must be a square function in the frequency domain. You may say that a sinc function in the time domain extends from -∞ to +∞ so if one sends just one symbol it occupies the line forever and interferes with all other sysmbols. But if you look more carefully to the waveform $sinc\left(\frac{t}{T}\right) = \frac{sin(\frac{\pi t}{T})}{\pi t/T}$ *we observe that its value is zero at t=kT, i.e., at any integer multiple of T. So, if we send a symbol every T seconds and are interested in the signal at multiples of T, then we have no problem.*

Digital Modulation
Pulse Shaping
(from Media Lab KyungHee University)

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(jw)e^{jwt}dw. \longleftrightarrow X(jw) = \int_{-\infty}^{\infty} x(t)e^{-jwt}dt.$$

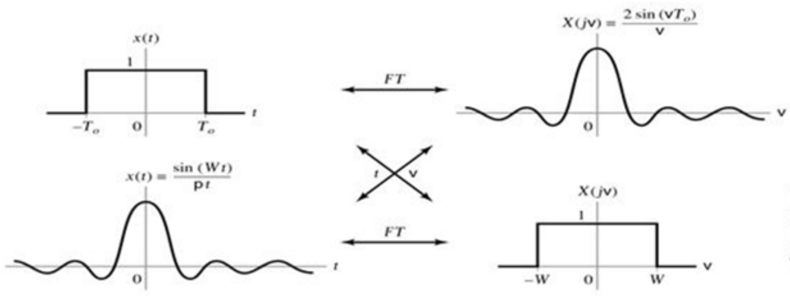


Fig. 3.73 Duality of rectangular pulses and sin

$$y(v) = \frac{1}{2\pi}\int_{-\infty}^{\infty} z(\eta)e^{jv\eta}dv.$$

if we choose $v = t$ *and* $\eta = w$, *then Eq.(3.66) implies that*

$$y(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} z(w)e^{jwt}dw.$$

$$y(t) \xleftrightarrow{FT} z(w).$$

$$y(-w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} z(t)e^{-jwt}dt,$$

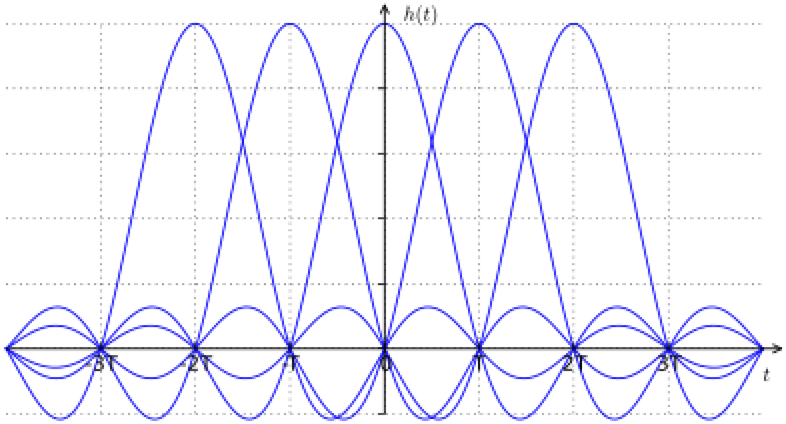
$$z(t) \xleftrightarrow{FT} 2\pi y(-w),$$

$$f(t) \xleftrightarrow{FT} F(jw),$$

$$F(jt) \xleftrightarrow{FT} 2\pi f(-w).$$

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

Note that at $t = kT, k = 0, \mp 1, \mp 2, \ldots$ only one of the pulses sent is non-zero and the rest have value zer at these points. So, there is no Intersymbol Interference (IS).

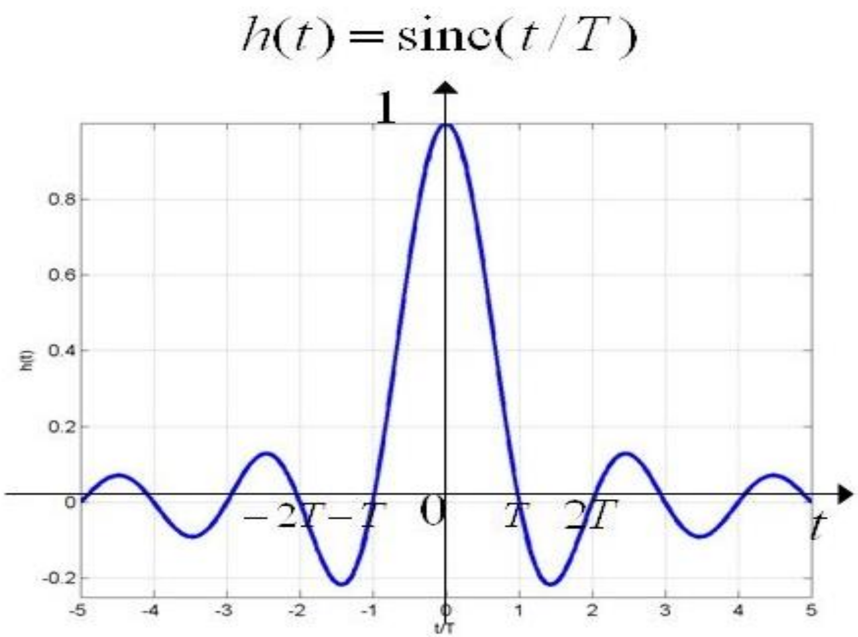


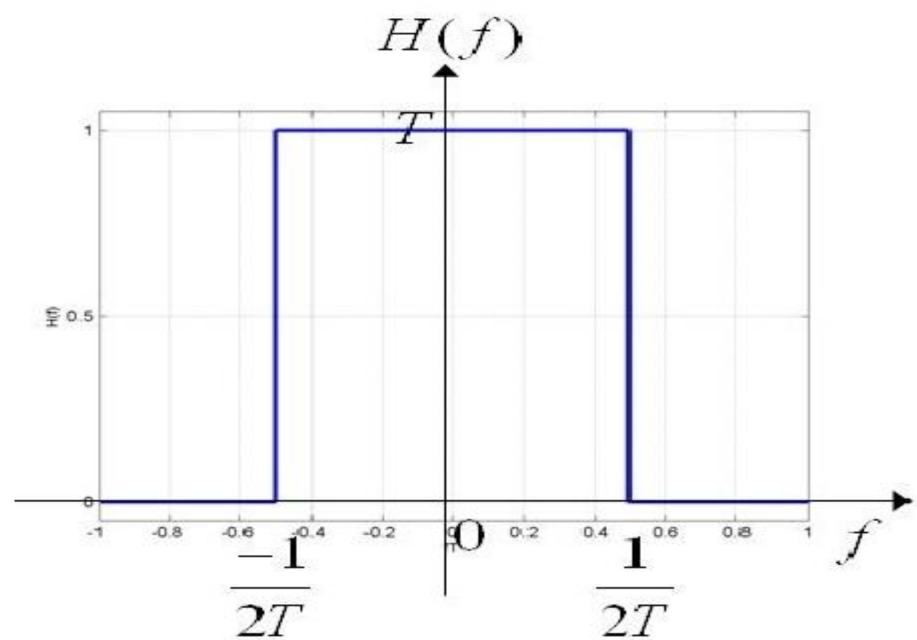
The spectrum of the sinc pulse p(t)=$sinc(t/T)$, called the Nyquist pulse, is a square function extending to 1/2T Hz. Note that the symbol rate is R_s $=1/T$. So the minimum bandwidth required for transmitting at a rate of R_s symbols per second is $W = \frac{R_S}{2}$. In the binary case every T second we send +p(t) or −p(t) depending on whether the bit to transmit is zero or one.

Digital Modulation
Pulse Shaping
Nyquist Pulse

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

The problem with the Nyquist pulse is that: The first problem is that it is not causal, i.e., not physically realizable since it extends to negative time. This means in order to send it at time t=kT, we need to know before hand whether we get a 0 or a one at time t. This of course can be remedied by delaying the pulse forming (shaping) for a few symbols and approximating the pulse with the main lobe and just a few side lobes.

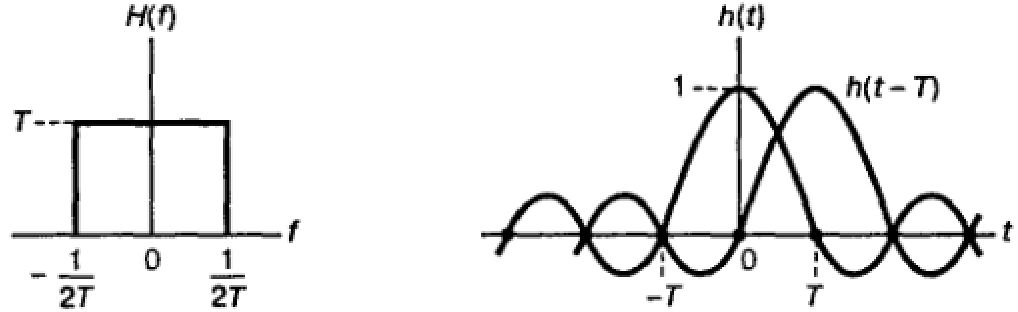




Digital Modulation
Pulse Shaping
Raised Cosine Filter

Department of Electrical & Computer Engineering

The second problem is with the slow decay of the side lobes (it decays as 1/t). This causes issues with timing. As long as the timing is perfect, i.e., the received signal is samples at T, 2T, 3T, etc. there is no problem and the pulse at (k-1)T and (k+1)T have no effect on the one sent at t=kT. But a slight error in timing not only results in lower value for the desired symbol, but also non-zero value contribution from the other bits. This causes ISI.

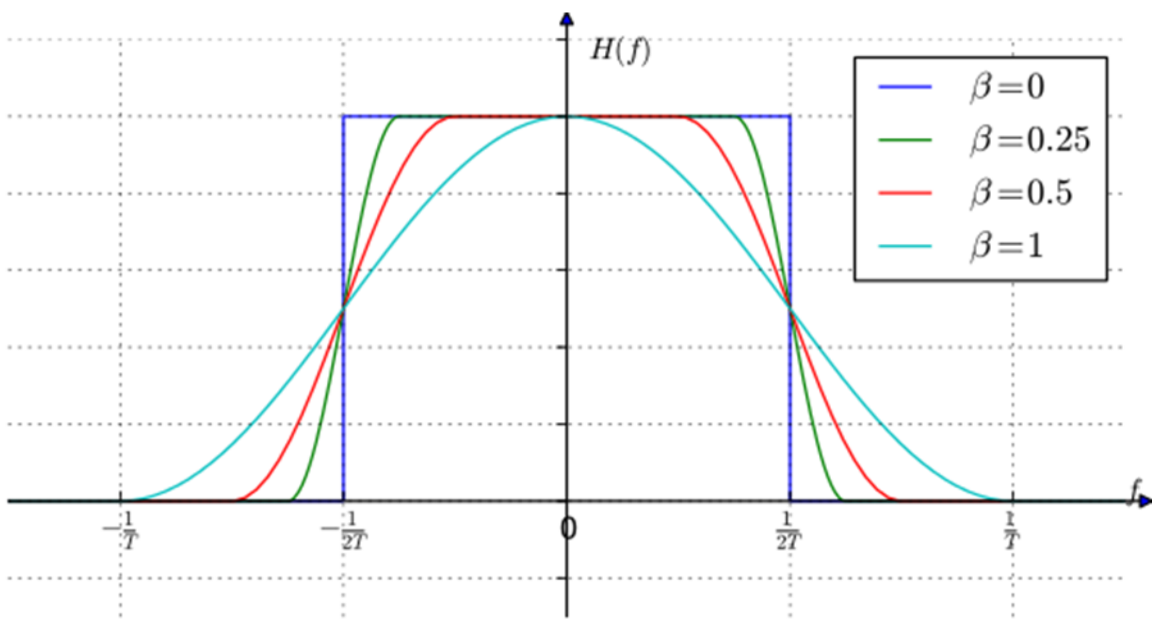


It is possible to make the decay faster by allowing the bandwidth to be more than 1/2T. It can be shown that ISI can be avoided as long as the pulse shape used has odd symmetry around the f=1/2T. One suggestion is the raised cosine (RC) filter.

Digital Modulation
Pulse Shaping
Raised Cosine Filter

The raised cosine spectrum consists of a flat part in the middle a transition band,

$$H(f) = \begin{cases} 1, & |f| \leq \frac{1-\beta}{2T} \\ \frac{1}{2}\left[1 + \cos\left(\frac{\pi T}{\beta}\left[|f| - \frac{1-\beta}{2T}\right]\right)\right], & \frac{1-\beta}{2T} < |f| \leq \frac{1+\beta}{2T} \\ 0, & \text{otherwise} \end{cases}$$



Université Concordia University
Department of Electrical & Computer Engineering

The parameter $0 \leq \beta \leq 1$ is called the roll-off factor and is a measure of how much more bandwidth we use compared to the minimum require (Nyquist). The required bandwidth is now:

$$W = (1+\beta)\frac{R_s}{2}$$

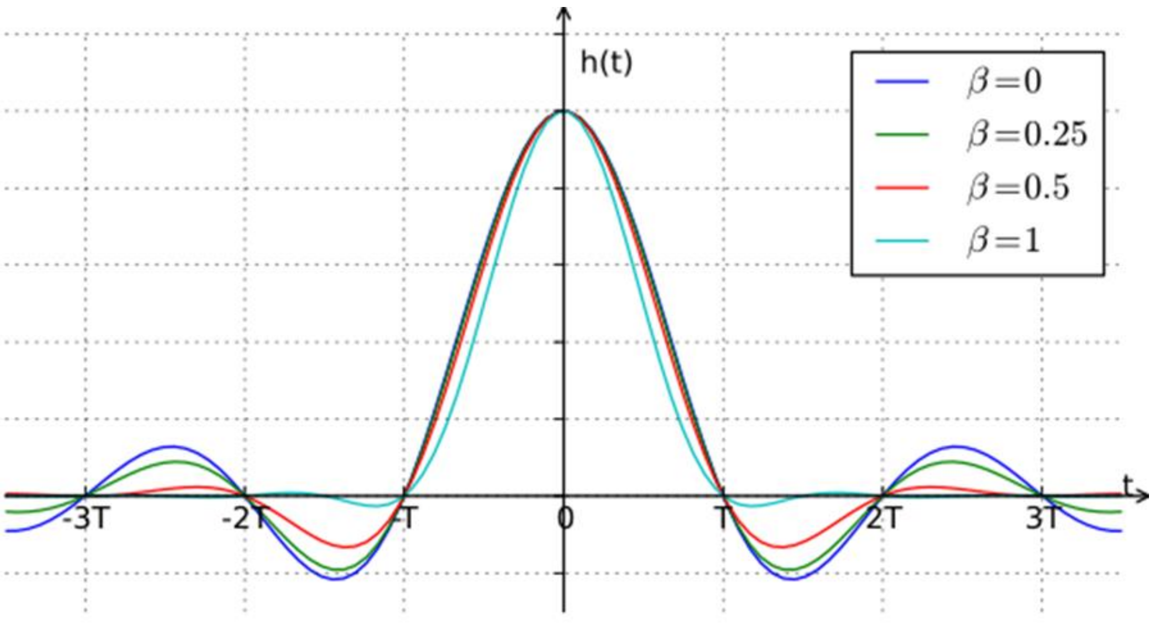
The raised cosine pulse in the time domain is,

$$h(t) = \begin{cases} \frac{\pi}{4T}\,\mathrm{sinc}\left(\frac{1}{2\beta}\right), & t = \pm\frac{T}{2\beta} \\ \frac{1}{T}\mathrm{sinc}\left(\frac{t}{T}\right)\frac{\cos\left(\frac{\pi\beta t}{T}\right)}{1-\left(\frac{2\beta t}{T}\right)^2}, & \text{otherwise} \end{cases}$$

The pulse $h(t)$ is shown in the next slide for different values of β.

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

It is seen that increasing β results in faster decay with time. The pulse decays almost as $\frac{1}{t^3}$. Note that this improvement in ISI comes at the expense of increasing the required bandwidth.

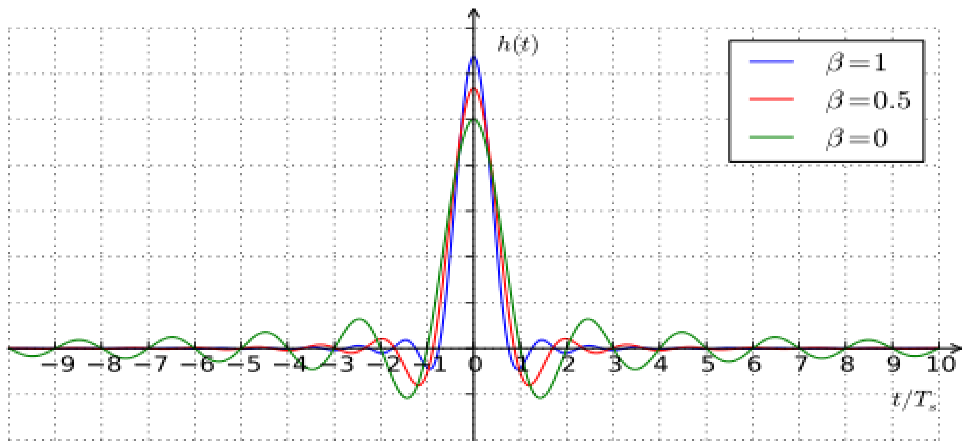


Digital Modulation
Pulse Shaping
Raised Cosine

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

Since the filter at the receiver (demodulator) has to be matched to the one at the transmitter (modulator), square root of the raised cosine filter is put at each sider. The pulse shaping filter is then called Square Root Raised Cosine (SRRC). Taking the inverse Fourier Transform of the square root of $H(f)$, we get,

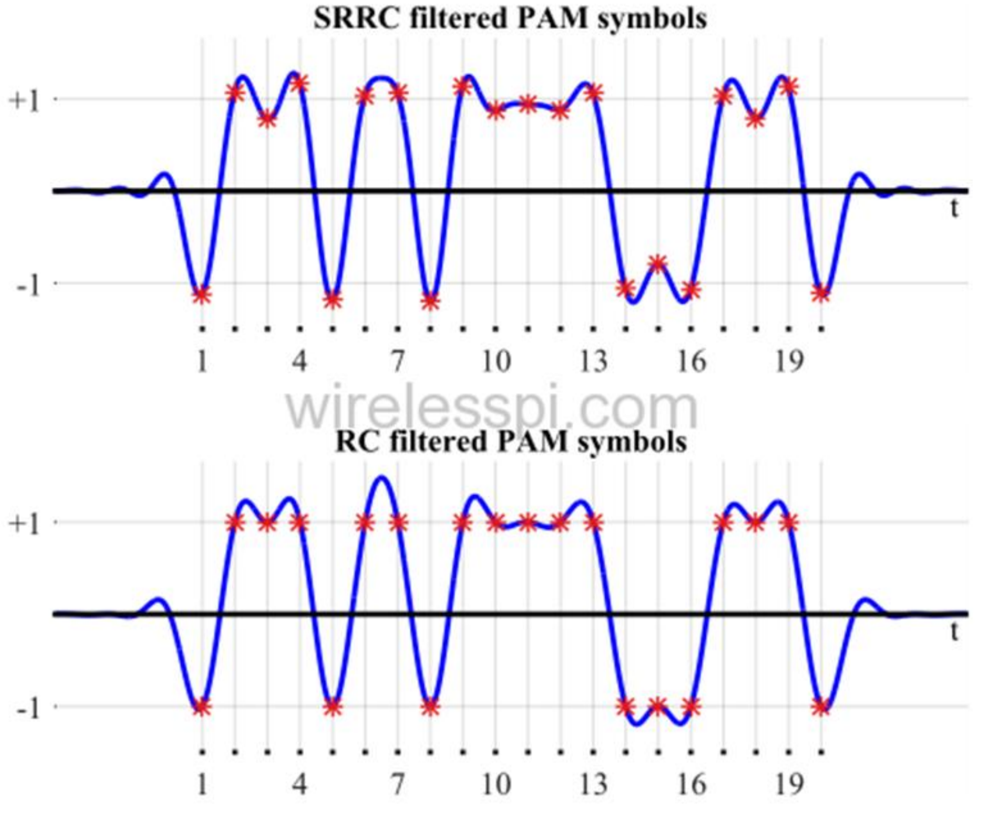
$$h_{srrc}(t) = \frac{sin\left(\pi \frac{t}{T}(1-\beta)\right) + 4\beta \frac{t}{T}\cos\left(\pi \frac{t}{T}(1+\beta)\right)}{\pi \frac{t}{T}\left(1-\left(4\beta \frac{t}{T}\right)^2\right)}$$



Digital Modulation
Pulse Shaping
Raised Cosine

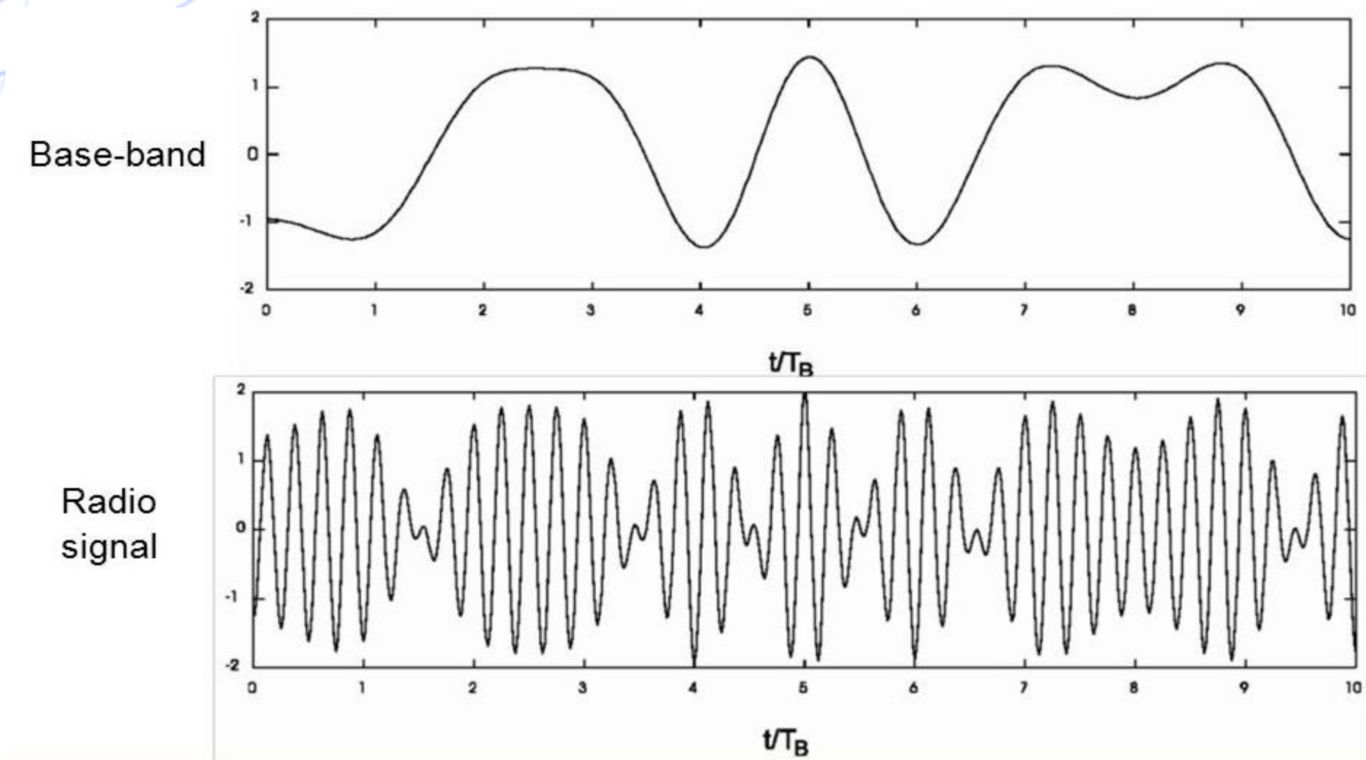
Université Concordia University
Department of Electrical & Computer Engineering

Bits will be mapped into +h(t) and –h(t) before RF modulation. This is referred to pulse shaping or Pulse Amplitude Modulation. For example bit stream 0111011011110001110 gets modulated as:



Digital Modulation
Pulse Shaping
Up/Conversion

The pulse shaped bit stream is then carrier modulated, i.e., is changed into a radio frequency signal (RF) ready for transmission. This process is also called up conversion.



Université Concordia University
Department of Electrical & Computer Engineering

Carrier modulation (up conversion) firstly, makes the signal more suitable for transmission. Note that the size of the has to be comparable with the wavelength of the signal. Usually it is at least one quarter of the wavelength.

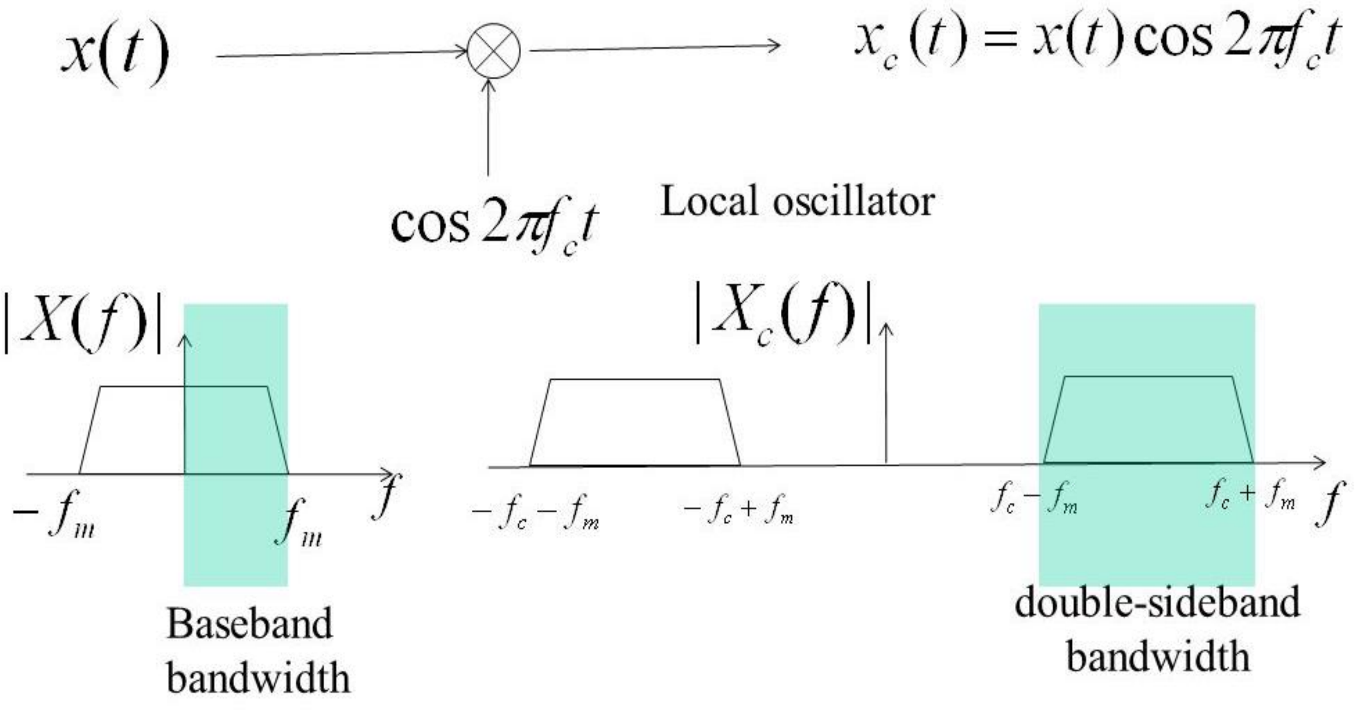
Secondly, up conversion is a tool for assigning different bands of frequency (channels) to different streams.

Note that shifting the baseband spectrum to a carrier frequency f_c results in the bandwidth to be doubled. While in the baseband the bandwidth extends from 0 to $\frac{R_s}{2}(1+\beta)$, after RF modulation, in the passband, it extends from $f_c - \frac{R_s}{2}(1+\beta)$ to $f_c + \frac{R_s}{2}(1+\beta)$. That is the required bandwidth is $W = R_s(1+\beta)$.

Using Single Side Band (SSB), i.e., filtering the upper or lower sideband of the spectrum, we can reduce the required bandwidth to very close to $W = \frac{R_s}{2}(1+\beta)$.

Concordia University
Department of Electrical & Computer Engineering

Carrier modulation (up conversion)



Université Concordia University
Department of Electrical & Computer Engineering

At the receiver, the received signal is multiplied by a sine wave $Cos2\pi f_c t$ to bring it back to the baseband. Then it is filtered by a square root raised cosine filter (called a matched filer as it is the same as the transmitter filter). Finally the matched filtered signal is sampled every T seconds and decision is made on whether it has been a one or a zero (a -1 or a +1).

If there was no noise, we would have gotten the right answer every time. However, the whole point is that noise is added to the signal and what we get at the output of the matched filter is not clear -1 or +1.

Later in the course we will talk more about the noise, but at the time being, let's accept that the noise is Gaussian (also called Normal) with power (variance) $\sigma^2 = \frac{N_0}{2}$. This is basically, the noise power density whose dimension is Watts/Hz.

This Gaussian Noise is added to the signal, therefore the name, AWGN (Additive White Gaussian Noise).

The reason for noise being almost always assumed to be Gaussian is the fact that due to a theorem in probability theory called the Central Limit Theorem, any phenomenon resulting from the addition of a lot of phenomena with the same distribution has a Gaussian distribution.

Concordia University
Department of Electrical & Computer Engineering

The reason for noise being almost always assumed to be Gaussian is the fact that due to a theorem in probability theory called the Central Limit Theorem, any phenomenon resulting from the addition of a lot of phenomena with the same distribution has a Gaussian distribution.

The noise in electrical circuits, including those used for communication, is due to the movement of many electrons behaving identically. So, it should not be a surprise if it is Gaussian.

A Gaussian random variable, x has the probability density function,

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the mean and σ^2 is the variance of x.

When the noise has zero mean, i.e., it has no DC value, we have,

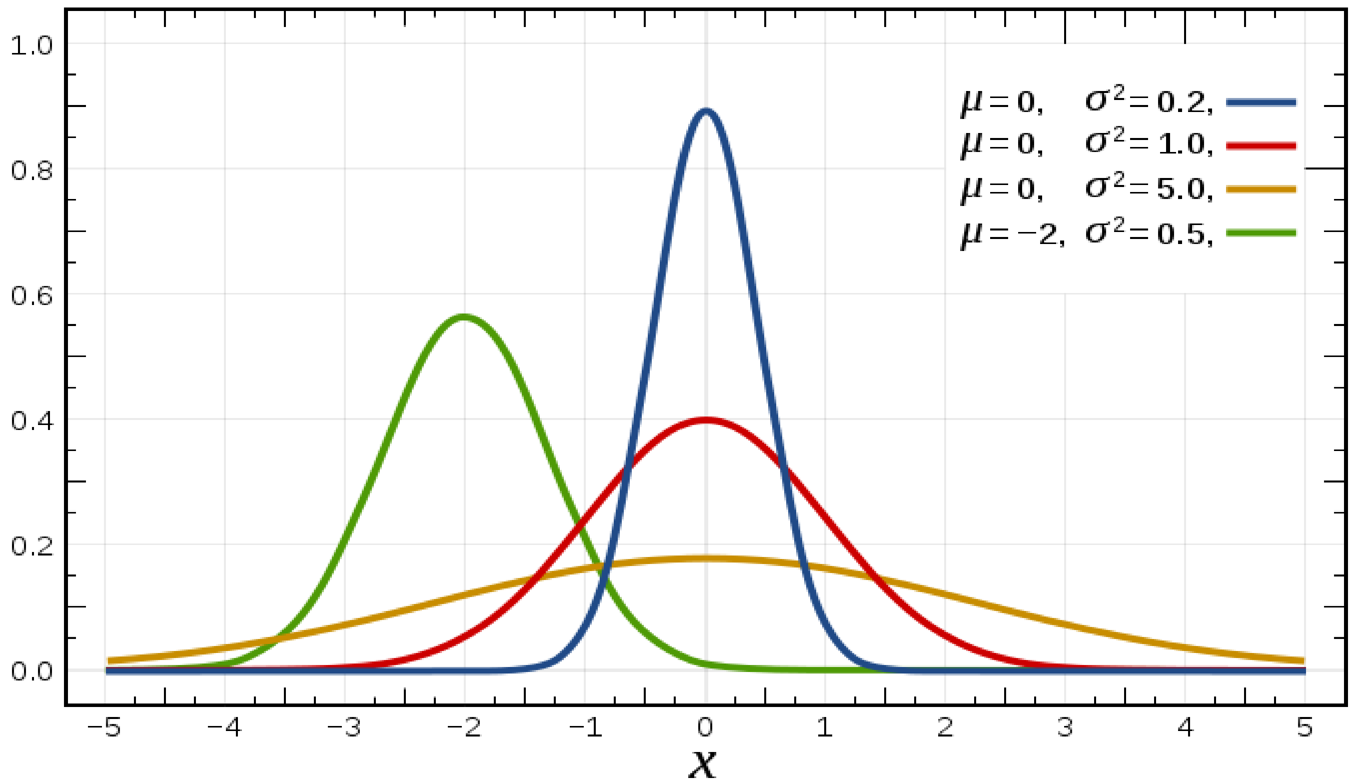
$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

Digital Modulation
Pulse Shaping
Gaussian Noise

Université Concordia University
Department of Electrical & Computer Engineering

Gaussian noise with different means and variances:

$p(x)$



Université Concordia University
Department of Electrical & Computer Engineering

The probability that a random variable with probability density function $f(x)$ is between two values a and b is,

$$P_r(a \leq x \leq b) = \int_a^b p(x)dx \,.$$

For a random variable x the probability that it is greater than some given value a is,

$$P_r(x \geq a) = \int_a^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \,.$$

Making change of variable $z = \frac{x-\mu}{\sigma}$, we get,

$$P_r(x \geq a) = R_r\left(z \geq \frac{a-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^\infty e^{-\frac{z^2}{2}} dz \,.$$

The integral,

$$Q(u) = P_r(z \geq u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{z^2}{2}} dz$$

represents the probability that a standard Gaussian random variable, i.e., one with mean zero and variance one exceeds a certain value. It is tabulated and there are also many approximations for it. The general case of a Gaussian variable with given mean and variance can be computed using it. It is called Q-function.

Concordia University
Department of Electrical & Computer Engineering

So,

$$P_r(x \geq a) = \mathrm{Q}\left(\frac{a-\mu}{\sigma}\right).$$

The following inequalities can be used to find a good approximation for Q function,

$$\frac{1}{u\sqrt{2\pi}}\left(1-\frac{1}{u^2}\right)e^{-\frac{u^2}{2}} < Q(u) < \frac{1}{u\sqrt{2\pi}}e^{-\frac{u^2}{2}}$$

A table giving values of the Q function for a wi8de range of values is shown in the next page.

UNIVERSITÉ Concordia UNIVERSITY
Department of Electrical & Computer Engineering

Table 1: Values of $Q(x)$ for $0 \leq x \leq 9$

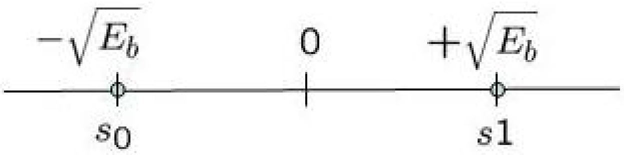
x	$Q(x)$	x	$Q(x)$	x	$Q(x)$	x	$Q(x)$
0.00	0.5	2.30	0.010724	4.55	2.6823×10^{-6}	6.80	5.231×10^{-12}
0.05	0.48006	2.35	0.0093867	4.60	2.1125×10^{-6}	6.85	3.6925×10^{-12}
0.10	0.46017	2.40	0.0081975	4.65	1.6597×10^{-6}	6.90	2.6001×10^{-12}
0.15	0.44038	2.45	0.0071428	4.70	1.3008×10^{-6}	6.95	1.8264×10^{-12}
0.20	0.42074	2.50	0.0062097	4.75	1.0171×10^{-6}	7.00	1.2798×10^{-12}
0.25	0.40129	2.55	0.0053861	4.80	7.9333×10^{-7}	7.05	8.9459×10^{-13}
0.30	0.38209	2.60	0.0046612	4.85	6.1731×10^{-7}	7.10	6.2378×10^{-13}
0.35	0.36317	2.65	0.0040246	4.90	4.7918×10^{-7}	7.15	4.3389×10^{-13}
0.40	0.34458	2.70	0.003467	4.95	3.7107×10^{-7}	7.20	3.0106×10^{-13}
0.45	0.32636	2.75	0.0029798	5.00	2.8665×10^{-7}	7.25	2.0839×10^{-13}
0.50	0.30854	2.80	0.0025551	5.05	2.2091×10^{-7}	7.30	1.4388×10^{-13}
0.55	0.29116	2.85	0.002186	5.10	1.6983×10^{-7}	7.35	9.9103×10^{-14}
0.60	0.27425	2.90	0.0018658	5.15	1.3024×10^{-7}	7.40	6.8092×10^{-14}
0.65	0.25785	2.95	0.0015889	5.20	9.9644×10^{-8}	7.45	4.667×10^{-14}
0.70	0.24196	3.00	0.0013499	5.25	7.605×10^{-8}	7.50	3.1909×10^{-14}
0.75	0.22663	3.05	0.0011442	5.30	5.7901×10^{-8}	7.55	2.1763×10^{-14}
0.80	0.21186	3.10	0.0009676	5.35	4.3977×10^{-8}	7.60	1.4807×10^{-14}
0.85	0.19766	3.15	0.00081635	5.40	3.332×10^{-8}	7.65	1.0049×10^{-14}
0.90	0.18406	3.20	0.00068714	5.45	2.5185×10^{-8}	7.70	6.8033×10^{-15}
0.95	0.17106	3.25	0.00057703	5.50	1.899×10^{-8}	7.75	4.5946×10^{-15}
1.00	0.15866	3.30	0.00048342	5.55	1.4283×10^{-8}	7.80	3.0954×10^{-15}
1.05	0.14686	3.35	0.00040406	5.60	1.0718×10^{-8}	7.85	2.0802×10^{-15}
1.10	0.13567	3.40	0.00033693	5.65	8.0224×10^{-9}	7.90	1.3945×10^{-15}
1.15	0.12507	3.45	0.00028029	5.70	5.9904×10^{-9}	7.95	9.3256×10^{-16}
1.20	0.11507	3.50	0.00023263	5.75	4.4622×10^{-9}	8.00	6.221×10^{-16}
1.25	0.10565	3.55	0.00019262	5.80	3.3157×10^{-9}	8.05	4.1397×10^{-16}
1.30	0.0968	3.60	0.00015911	5.85	2.4579×10^{-9}	8.10	2.748×10^{-16}
1.35	0.088508	3.65	0.00013112	5.90	1.8175×10^{-9}	8.15	1.8196×10^{-16}
1.40	0.080757	3.70	0.0001078	5.95	1.3407×10^{-9}	8.20	1.2019×10^{-16}
1.45	0.073529	3.75	8.8417×10^{-5}	6.00	9.8659×10^{-10}	8.25	7.9197×10^{-17}
1.50	0.066807	3.80	7.2348×10^{-5}	6.05	7.2423×10^{-10}	8.30	5.2056×10^{-17}
1.55	0.060571	3.85	5.9059×10^{-5}	6.10	5.3034×10^{-10}	8.35	3.4131×10^{-17}
1.60	0.054799	3.90	4.8096×10^{-5}	6.15	3.8741×10^{-10}	8.40	2.2324×10^{-17}
1.65	0.049471	3.95	3.9076×10^{-5}	6.20	2.8232×10^{-10}	8.45	1.4565×10^{-17}
1.70	0.044565	4.00	3.1671×10^{-5}	6.25	2.0523×10^{-10}	8.50	9.4795×10^{-18}
1.75	0.040059	4.05	2.5609×10^{-5}	6.30	1.4882×10^{-10}	8.55	6.1544×10^{-18}
1.80	0.03593	4.10	2.0658×10^{-5}	6.35	1.0766×10^{-10}	8.60	3.9858×10^{-18}
1.85	0.032157	4.15	1.6624×10^{-5}	6.40	7.7688×10^{-11}	8.65	2.575×10^{-18}
1.90	0.028717	4.20	1.3346×10^{-5}	6.45	5.5925×10^{-11}	8.70	1.6594×10^{-18}
1.95	0.025588	4.25	1.0689×10^{-5}	6.50	4.016×10^{-11}	8.75	1.0668×10^{-18}
2.00	0.02275	4.30	8.5399×10^{-6}	6.55	2.8769×10^{-11}	8.80	6.8408×10^{-19}
2.05	0.020182	4.35	6.8069×10^{-6}	6.60	2.0558×10^{-11}	8.85	4.376×10^{-19}
2.10	0.017864	4.40	5.4125×10^{-6}	6.65	1.4655×10^{-11}	8.90	2.7923×10^{-19}
2.15	0.015778	4.45	4.2935×10^{-6}	6.70	1.0421×10^{-11}	8.95	1.7774×10^{-19}
2.20	0.013903	4.50	3.3977×10^{-6}	6.75	7.3923×10^{-12}	9.00	1.1286×10^{-19}
2.25	0.012224						

From the notes given by Dr. Modhusudhon Tarafdar at Alagappa University for a Digital Communication course.

Concordia University
Department of Electrical & Computer Engineering

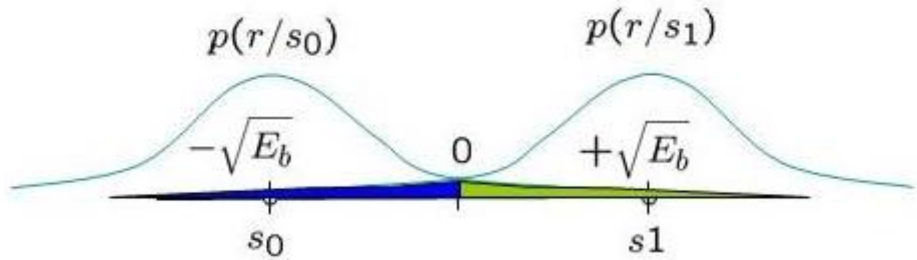
Let's get back to our demodulation problem. Assume that the transmitter amplifies the signal it transmits such that the received energy per bit is E_b Watts/bit. As we will see later, this can be calculated by dividing the transmitted power by all the losses incurred between transmitter and receiver and then divide the received power (energy per second) by the bit rate (number of bits per second).

The two signals representing 0 and 1, i.e., $\pm\frac{\sqrt{E_b}}{\sqrt{T_b/2}}\cos(2\pi f_c t)$ can be shown as two points on an axis that is scaled by the basis function $\varphi(t) = \sqrt{\frac{2}{T_b}}\cos(2\pi f_c t)\, h(t)$.



Digital Modulation
Pulse Shaping
Bit Error Rate (BER)

At the receiver, with the noise being added to the transmitted signal we get a value that can be anywhere on the axis. Of course, the distribution is according to the Gaussian distribution as shown below,



Assume that the probability of transmitting 0 and 1 is the same. In that case, it is intuitively clear that we should decide in favor of 1 when the sampled output of the matched filter is greater than zero and decide 0 otherwise.

The probability density function of the number at the input of the decision device is,

$$p(r|1) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{\left(r+\sqrt{E_b}\right)^2}{N_0}} \text{ when 0 is sent}$$

and

$$p(r|1) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{\left(r-\sqrt{E_b}\right)^2}{N_0}} \text{ when 1 is sent}$$

Digital Modulation
Pulse Shaping
Bit Error Rate (BER)

Université Concordia University
Department of Electrical & Computer Engineering

There are two types of errors that may occur. The transmitter may send a 0 and the receiver decides 1 or vice versa.

When 0 is transmitted for error to occur, the received signal has to exceed zero. This means that the noise level has to be greater than $\sqrt{E_b}$. This will take the matched filter output to the area under the green tail of the Normal curve. The probability of this type of error, say we call it type I is,

$$P_I = \frac{1}{\sqrt{\pi N_0}} \int_{E_b}^{\infty} e^{-\frac{z^2}{N_0}} dz = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

Similarly, when 1 is sent, an error occurs if the noise level is less than $-\sqrt{E_b}$. The probability of this type of error called type II is

$$P_I = \frac{1}{\sqrt{\pi N_0}} \int_{-\infty}^{-E_b} e^{-\frac{z^2}{N_0}} dz = Q\left(\sqrt{\frac{2E_b}{N_0}}\right).$$

So, the overall probability of error, also called Bit Error Rate (BER) is,

$$BER = \frac{1}{2} P_I + \frac{1}{2} P_{II} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right).$$

Concordia University
Department of Electrical & Computer Engineering

So far, we have talked only about binary communication, i.e., sending two types of symbols: 0 or 1. We send $\sqrt{\frac{2E_b}{N_0}}\cos(2\pi f_c t)$ for 1 and $-\sqrt{\frac{2E_b}{N_0}}\cos(2\pi f_c t) = \sqrt{\frac{2E_b}{N_0}}\cos(2\pi f_c t + \pi)$ for 0. This is called Binary Shift Keying (BPSK). Note that here we are sending one bit of information with each symbol. Recall that the require bandwidth is,

$$W = (1+\beta)R_s$$

As said earlier filtering the upper or lower side band we can reduce this bandwidth. But here we do not assume doing so.

In BPSK since each symbol represents a single bit, the symbol rate is equal to the bit rate, i.e., $R_s = R_b$. Therefore the required bandwidth is, $W = (1+\beta)R_b$. So, we can transmit

$$R_b = \frac{W}{(1+\beta)}$$

bits per second (bps) in W Hz. Equivalently, we can say that the bandwidth efficiency is,

$$\eta = \frac{R_b}{W} = \frac{1}{(1+\beta)} \text{ bits/sec/Hz.}$$

We can increase bandwidth efficiency, i.e., the bit rate obtained for a given band of frequency by using more levels. This means that each symbol and, consequently, each sinusoid taking more than two values. Assume that instead of $\pm\sqrt{E_b}$, we use $-3d, -d, +d$ and $+3d$ representing, 00, 01, 11 and 10, respectively.

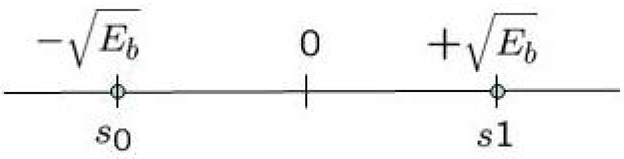
Then each symbol carries two bits. So $R_b = 2R_s$. The required bandwidth is then,

$$W = (1+\beta)\frac{R_b}{2}$$

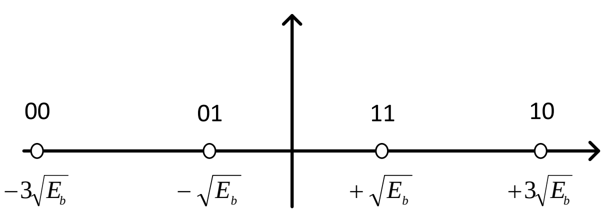
In general, using an M point constellation, each point represents $log_2 M$ bits. So,

$$W = (1+\beta)\frac{R_b}{log_2 M}.$$

This increase in bandwidth efficiency, however, comes at the expense of an increase in the required power requirement. In the case of BPSK, the energy per bit is E_b as shown below:



The distance between the two symbols is $2\sqrt{E_b}$. If we have four points, i.e., two bits/symbol, if we want to have the same error probability (BER), we need to have the same distance between the points:



Digital Modulation
Pulse Shaping
M-ary Modulation

The average energy per symbol is $E_s = \frac{(9E_b+E_b+E_b+9E_b)}{4} = 5E_b$ sine each symbol represents two bits, the energy per bit for the quaternary system is $E_b' = 2.5E_b$. So there is a tradeoff between the bandwidth and power requirement.

The points in a modulation constellation may be placed on a single line, i.e., they modulate the amplitude of a sinusoid (ASK*), or the phase (PSK**) or both (APSK***).

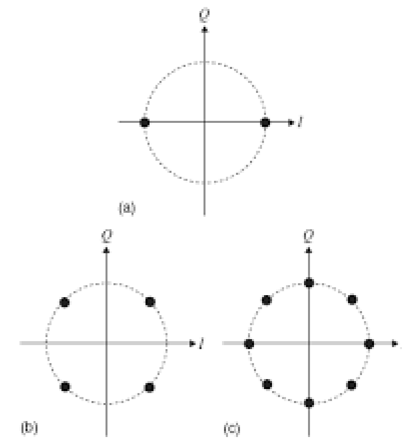
Some common modulation types are MPSK (M-ary PSK), M-QAM(M-ary Quadrature Amplitude Modulation) and APSK.

The figure on the right shows MPSK for different values of M.

The case for M=2 is BPSK that we have already see. For M=4,

We have 4PSK more commonly known as QPSK. Then there are

8PSK, 16 PSK and so on.



We have seen that the BER of the BPSK is $Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$.

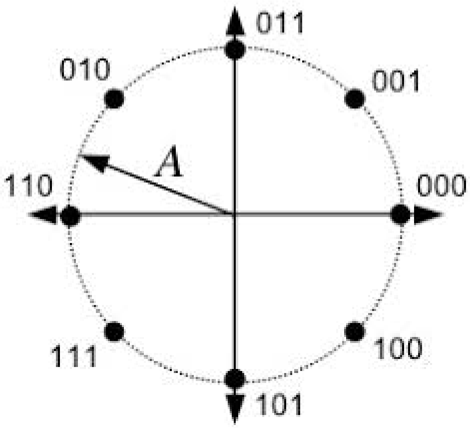
For QPSK, we have the same formula since while the energy per symbol doubles the energy per bit remains the same and also a symbol error only makes one bit error on the average.

* ASK: Amplitude Shift Keying, ** PSK: Phase Shift Keying, ***APSK: Amplitude Phase Shift Keying.

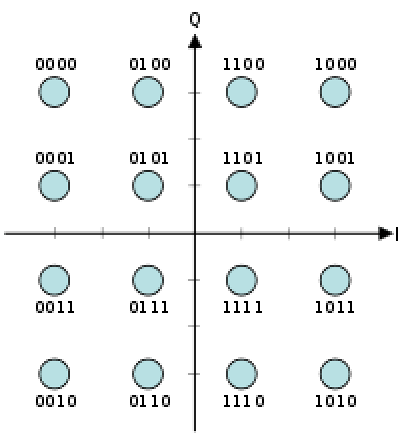
Digital Modulation
Pulse Shaping
M-ary Modulation

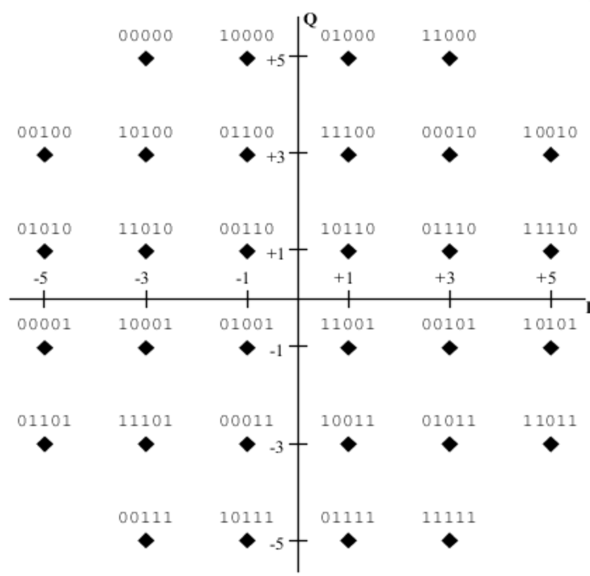
In general, the BER of MPSK is,

$$BER_{MPSK} \approx \frac{1}{log_2 M} Q\left(\sqrt{\frac{2E_b}{N_0} log_2 M} \sin\left(\frac{\pi}{M}\right)\right)$$



Another modulation technique is M-ary Quadrature Amplitude Modulation (MQAM). Following is the constellation of 16 QAM and 32QAM.





The BER of MQAM is approximated as,

$$BER_{MQAM} \approx \frac{4}{log_2 M}\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3log_2 M}{M-1}\frac{E_b}{N_0}}\right)$$

MQAM has better performance than MPSK in terms of BER when the channel is linear, i.e., amplifiers are not saturated. Constraining the operation to the linear range of amplifiers result in efficiency. For that reason MPSK is preferred where the efficiency of power amplifiers is a concern. This is due to the fact that all symbols in MPSK have the same amplitude and the information is conveyed through the changes in phase.

A compromise is to use a combination of amplitude and phase modulation, This is called Amplitude Phase Shift Keying (APSK). This scheme is one of the modulation schemes in DVB-S2 and DVB-S2X standard used

for satellite broadcasting.