ELE 635 Communication Systems

Course Notes

M. Zeytinoglu

Department of Electrical and Computer Engineering
Ryerson University
Toronto, Ontario, Canada M5B 2K3

Email: mzeytin@ee.ryerson.ca

Winter 2013

Contents

Chapter 1

Introduction

ELE 635 Communication Systems is an introductory course on communication systems. The course will focus on the *transmission of information* and will investigate techniques applicable to the analysis and synthesis of analog communication systems. In particular, we will study:

- mathematical tools (e.g. Fourier analysis);
- transmission media (e.g. communication channels);
- basic transmission and reception techniques (e.g. amplitude modulation (AM), frequency modulation (FM) ...);
- performance analysis of standard communication systems;
- implementation examples (e.g. AM/FM radio broadcasting ...).

We will study these concepts from a *systems* point of view. In the laboratory we will look at the implementation of basic communication systems elements including modulation, envelope detection, phase locked loop (PLL), PLL as an FM signal detector, etc.

1.1 Overview

Figure 1.1 presents a generic communications system. Throughout the course we will investigate

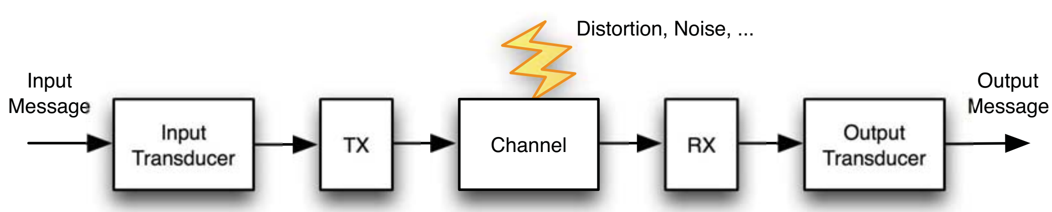


Figure 1.1: Communication system.

the transmitter (TX), channel and receiver (RX) triplet as shown in the block diagram. Basic functionality of these components are as follows.

Input Message: signal containing the raw information that we want to transmit, e.g. human voice (speech), image, ...

Input Transducer: converts the raw input signal into an electrical signal (voltage, current, ...). We will refer to the resulting signal as the baseband/input/message signal.

Transmitter (TX): converts the input signal (electronically, mechanically, physically, ...) to a format that best matches the channel characteristics (fibre optic cable, electronic wave guide, coaxial cable, ...); the transmitter uses modulation and frequency translation techniques.

Channel: Physical medium over which the signal will travel. Here, in the channel, a lot of "ugly", "nasty" things (from the message point of view) will happen, e.g. distortion, noise, interference, fading, etc.

Receiver (RX) and the Output Transducer: These components reverse the modifications implemented by the transmitter and the input transducer, respectively.

Our goal is to transmit the information contained within the input signal to the destination such that the information contained in the received signal is either identical (ideally) or closely resembles the information in the input signal. And of course, we want to achieve this objective most efficiently, economically and effectively.

In this course, we will study *analog* communication systems, whereas the 4*th* year elective courses will introduce *digital* communication systems. However, as digital communication systems also use analog signals which in turn rely on analog communication techniques, this course will indeed function as an introduction to universal communication system fundamentals.

1.2 Analog vs. Digital Information Systems

Analog: We transmit the information carrying signal in raw format (even after the modifications applied by the input transducer and the transmitter). Hence, any distortion affecting the signal in the channel will be part of the received signal and will irreversibly distort the signal. By choosing the TX and the RX appropriately we can minimize distortion effects, but they will always be part of the signal.

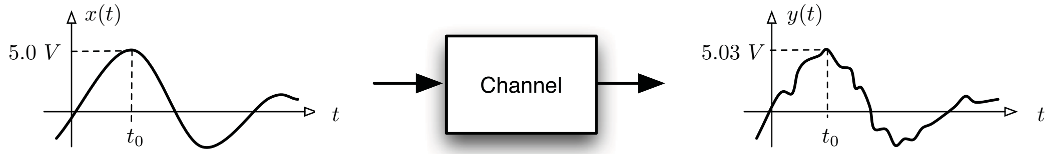


Figure 1.2: Analog signal transmission over a communication channel.

Example 1.1: If the level of the transmitted waveform at time t_0 equals 5 V, and the noise and/or distortion changes this value to 5.03 V, the distorted value is what we will see/hear/receive at the RX. There is no going back to the true voltage value of 5 V as the noise distorting the signal is not known.

Digital: In this case we transmit a "representation" of the signal $x(t)$. Once again let us consider the value of the signal at t_0:

$$x(t_0) = 5V \quad \Longrightarrow \quad \{5\}_{decimal} \equiv \{0101\}_{binary} \quad \Longrightarrow \quad \{-A,\, +A,\, -A,\, +A\},$$

where we assume that each signal value is represented by a 4-digit binary number; a $\{1\}_{binary}$ is represented by a rectangular pulse of $+A$ amplitude and a $\{0\}_{binary}$ is represented by a rectangular pulse of $-A$ amplitude. The waveforms representing $x(t_0)$ and the received waveform $y(t)$ at the output of distorting and noisy channels are shown in Figure (1.3).

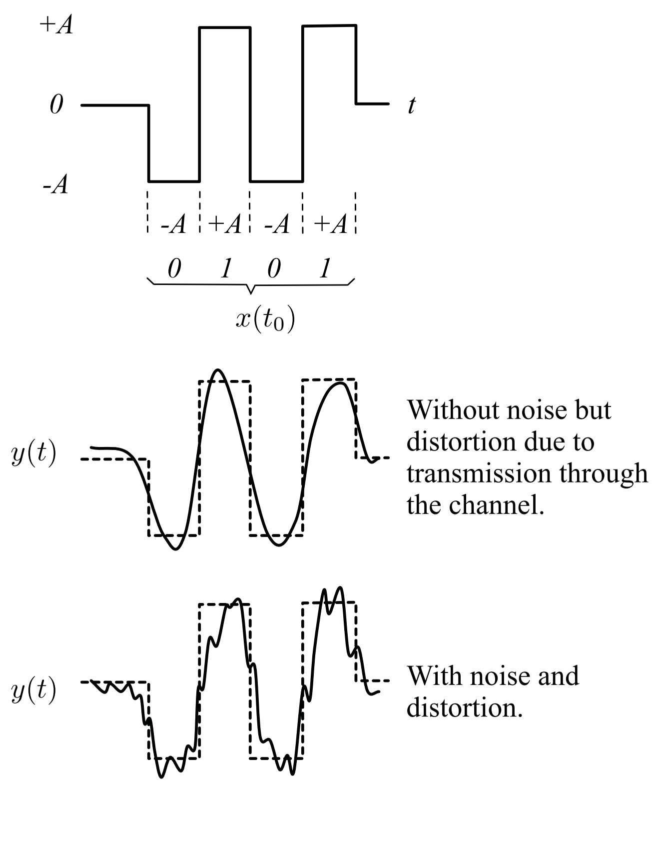


Figure 1.3: Digital signal transmission over a noisy communication channel.

As we know the format of the transmitted waveform (positive pulses representing $\{1\}_{binary}$ ... but not the order of positive and negative pulses), we can still correctly identify the "most likely" pulse train which must have been transmitted. Hence by inspecting $y(t)$ we can conclude that

$$y(t_0) = \{0101\}_{binary} = \{5\}_{decimal} = x(t_0).$$

The main advantage of the digital communication systems is the symbolic representation of the signal value. Theoretically (and practically) it is possible to receive and decode a digital signal such that $y(t) = x(t)$. This statement is a direct consequence of the underlying fact that there is only a finite number of possible signals that we can transmit (e.g. there are two possible signal levels in the case of binary coding ... M possible signal levels in the case of M-ary coding).

Example 1.2:

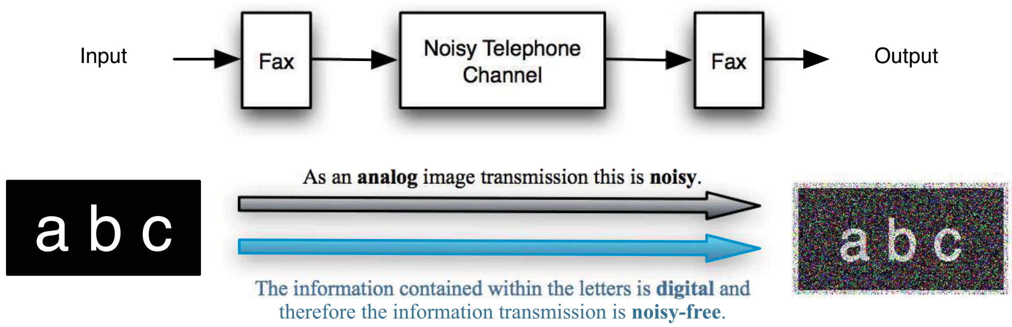


Figure 1.4: Analog vs. digital signal transmission.

In our discussion we will refer to certain fundamental concepts that characterize the behaviour and performance of communication systems. These concepts include:

- Bandwidth B;
- Signal power, noise power and signal-to-noise ratio, SNR;
- Shannon's channel capacity theorem which relates B and SNR: $C = B\log_2(1+\text{SNR})$ bits/s;
- Randomness or uncertainty;
- Redundancy;
- Modulation;
- Multiplexing.

In the following weeks we will introduce these concepts, discuss their significance and will use them in the analysis and design of communication systems.

Chapter 2

Signals and Systems: A Brief Review

In this course we will mostly work with communication systems which fit into the block diagram template shown in Figure 2.1:

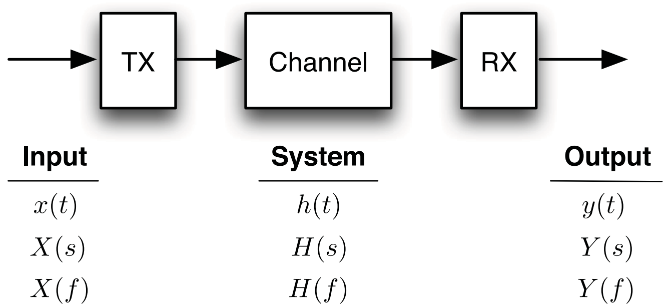


Figure 2.1: Basic communication system template.

We assume that the channel is modelled as a linear, time-invariant (LTI) system. This assumption is usually made to simplify and to introduce a degree of mathematical tractability to the analysis of the underlying communications problem. It is empirically shown that indeed a large class of communication system can be very accurately represented by appropriate LTI models. The *ELE 532 Signal and Systems* course introduced various linear system analysis tools: differential equations, difference equations, Laplace transform, Fourier transform, etc. In the subsequent semester, the *ELE 639 Control Systems* course used the Laplace transform as its principal analysis tool. Whereas in the *ELE 635 Communication Systems* course we will mostly use the Fourier transform. The Fourier transform will allow steady-state, sinusoidal analysis of the underlying signals and systems. This steady-state analysis is mostly sufficient for the study of communication systems; we will shortly justify this assertion.

2.1 Signal Measure

Given the fact that most signals of interest are rapidly varying functions of time "t", how can we measure the "size" of a given waveform? Remember that the signal $g(t)$ is not necessarily a scalar quantity. For a *dc* signal such as $g(t) = 5$ V, the concept of the size of the signal is well understood.

Even for a sinusoidal signal such as $g(t) = A\cos\omega_0 t$ we can unambiguously refer to the size of signal by referencing the amplitude of the sinusoidal oscillations. On the other hand, we have to introduce a new measure to compare the "size" of an arbitrary waveform $g(t)$. One possible measure is the signal energy defined as:

$$E_g = \int_{\infty}^{\infty} |g(t)|^2 dt. \tag{2.1}$$

Observe that for an arbitrary, possibly complex-valued signal $g(t)$

$$|g(t)|^2 = g(t)g^*(t). \tag{2.2}$$

If $g(t)$ is real-valued then $g(t) = g^*(t)$. Thus,

$$|g(t)|^2 = g(t)g^*(t) = g^2(t). \tag{2.3}$$

There is a large class of important signals for which E_g is not finite, namely *power signals*. For these signals we will use the signal power as a measure of the "size" of the signal:

$$P_g = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |g(t)|^2 dt. \tag{2.4}$$

For example, periodic signals are typically power signals. Please note that E_g and P_g do not indicate actual energy and power levels since they depend not only on the signal but also on the load. Hence, if $g(t)$ is a voltage waveform, then E_g and P_g represent the *unit* energy or power delivered across a 1-Ω resistor.

2.2 Some Useful Signal Characteristics and Operations

Periodic Signals: The waveform $g(t)$ is periodic with period T_0 if $g(t) = g(t + kT_0)$, for all t and $k \in \mathcal{Z}$ (the set of all integers: $0, \pm1, \pm2, \ldots$).

Energy Signals: The waveform $g(t)$ is an energy signal if $0 \le E_g < \infty$.

Power Signals: The waveform $g(t)$ is a power signal if $0 \le P_g < \infty$. Please note that a signal cannot simultaneously be both an energy and a power signal.

Time shifting: $g(t) \longrightarrow g(t - t_0)$:

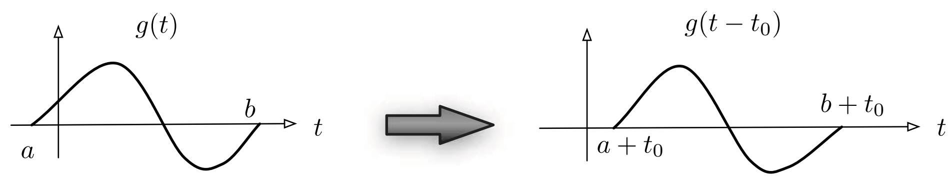


Figure 2.2: Time shifting operation.

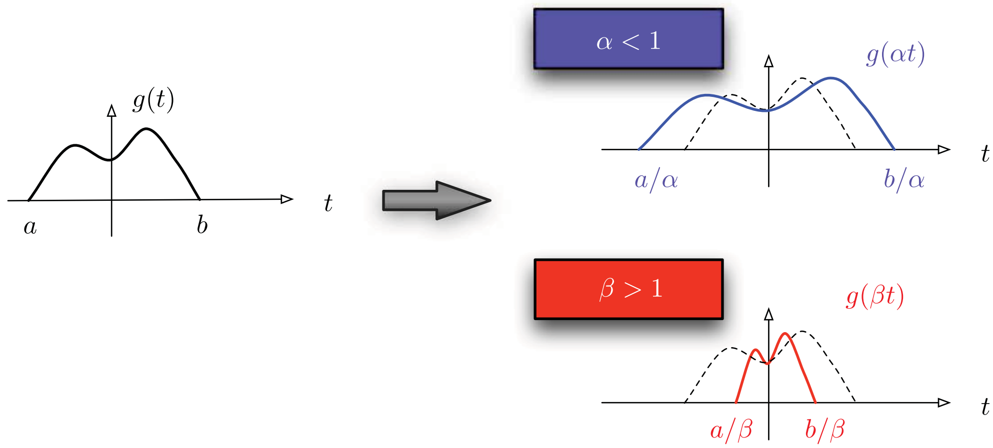


Figure 2.3: Time scaling operation.

Time Scaling: $g(t) \longrightarrow g(\alpha t)$ with $\alpha \in \mathcal{R}$. If $\alpha > 1$, then $g(\alpha t)$ represents a time-compressed (by a factor of α) version of $g(t)$, whereas if $\alpha < 1$, then $g(\alpha t)$ represents a time-stretched (by a factor of α) version of $g(t)$:

Time Reversal: $g(t) \longrightarrow g(-t)$ represents a mirror image of $g(t)$ with respect to the vertical axis $t = 0$. Similarly, $g(t_0 - t)$ with $t_0 \in \mathcal{R}$ first takes the mirror image of $g(t)$ with respect to the vertical axis and then shifts the resulting waveform by t_0 units.

2.3 Special Functions

2.3.1 Unit Impulse Signal

The unit impulse signal $\delta(t)$ is one of the most useful signals that will allow us to formulate many operations frequently encountered in the analysis of LTI systems. The unit impulse signal $\delta(t)$ (also known as the Dirac-delta function, named after the quantum mechanist P.A.M. Dirac) has zero amplitude everywhere except at $t = 0$ where its magnitude is infinitely large such that the area underneath is unity. As such $\delta(t)$ belongs to the category of *generalized functions*.

Definition: The unit impulse signal $\delta(t)$ is defined as:

$$\delta(t) = 0, \text{ if } t \neq 0; \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(t)dt = 1. \tag{2.5}$$

The Dirac-delta function is normally represented as a limiting case obtained by considering the limit of a regular function, e.g. $\delta(t) = \lim_{\tau \to 0} \text{rect}(t/\tau)$, where $\text{rect}(t/\tau)$ is a rectangular pulse of τ-duration centered at the origin.

Important Properties:

1. Symmetry: The δ-function has even symmetry, i.e., $\delta(t) = \delta(-t)$.

2. For an arbitrary function $x(t)$ and constant $K \in \Re$ we have:

$$\int \delta(t)x(t)dt = x(0);$$
$$\int \delta(t-t_0)x(t)dt = x(t_0), \quad \text{(the \textit{sifting} property)};$$
$$\int K\delta(t)dt = K.$$

3. Convolution with the δ-function:

$$\begin{aligned}\delta(t-t_0) * x(t) &= \int \delta(\lambda - t_0)x(t-\lambda)d\lambda, \\ &= x(t-\lambda)|_{\lambda=t_0}, \\ &= x(t-t_0).\end{aligned}$$

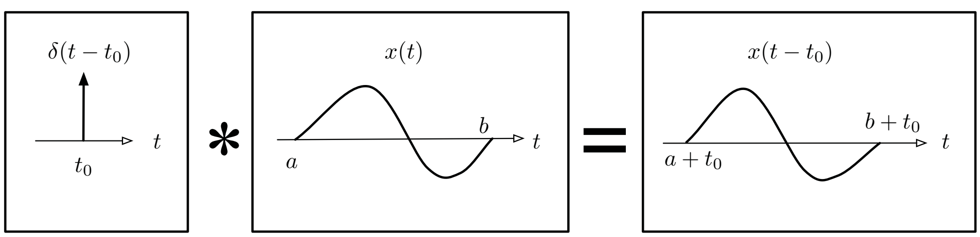


Figure 2.4: Convolution with a δ-function.

2.3.2 The Unit Rectangular Pulse Function: $\Pi(t)$

We will use the notation $\Pi(x)$[1] to refer to the *unit rectangular pulse function* of unit amplitude and unit width, centered at the origin:

$$\Pi(t) = \begin{cases} 1, & |t| < \frac{1}{2}; \\ \frac{1}{2}, & |t| = \frac{1}{2}; \\ 0, & |t| > \frac{1}{2}. \end{cases} \tag{2.6}$$

Frequently, we will use the scaled version of the unit rectangular pulse as $\Pi(t/\tau)$ where the scaling parameter τ will represent the width of the pulse centered at origin.

[1] We will also use the notation rect(t) interchangeably to represent the unit rectangular pulse function $\Pi(t)$.

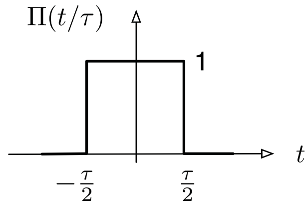


Figure 2.5: The rectangular pulse function $\Pi(t/\tau)$.

2.3.3 The Unit Triangular Pulse Function: $\Delta(t)$

The unit triangular pulse $\Delta(t)$ is a triangular shaped pulse of unit height and unit amplitude centered at the origin:

$$\Delta(t) = \begin{cases} 1 - 2|t|, & |t| \leq \frac{1}{2}; \\ 0, & |t| > \frac{1}{2}. \end{cases} \tag{2.7}$$

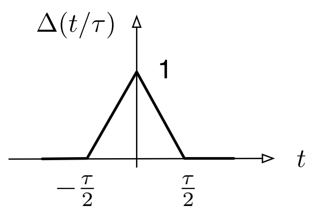


Figure 2.6: The triangular pulse function $\Delta(t/\tau)$.

2.3.4 The "Sinc" Function: sinc(t)

Another function that we will frequently use is the **sinc** function defined as:

$$\text{sinc}(t) = \frac{\sin(t)}{t}. \tag{2.8}$$

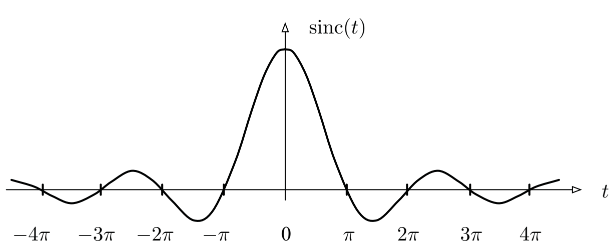


Figure 2.7: The function sinc(t).

Inspection of Equation(2.8) reveals the following key characteristics of the sinc function:

- sinc(t) has even symmetry, i.e., sinc$(t) =$ sinc$(-t)$.
- sinc$(t) = 0$ when $\sin(t) = 0$ except at $t = 0$ where it is indeterminate. This means that sinc$(t) = 0$ for $t = \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$.
- sinc$(0) = 1$. (As the denominator of $\sin(t)/t$ equals to 0 when $t = 0$, sinc(0) can be determined using the L'Hôpital's rule.
- sinc(t) is the product of an oscillating signal ($\sin t$) of period 2π and a monotonically decreasing function ($1/t$). Therefore, sinc(t) is an oscillating function with even symmetry and decreasing amplitude. It has a unit amplitude at $t = 0$ and zero crossings at integer multiples of π.

Chapter 3

Analysis and Transmission of Signals

3.1 Preliminaries

3.1.1 Fourier Series

Given a periodic waveform $x_p(t)$ with period T_0 (and the corresponding fundamental frequency $f_0 = 1/T_0$ such that $\omega_0 = 2\pi f_0$) we can expand $x_p(t)$ in a Fourier series:

$$x_p(t) = \sum_{n=-\infty}^{\infty} D_n e^{jn\omega_0 t} \tag{3.1}$$

where the coefficients of the Fourier series are evaluated as:

$$D_n = \frac{1}{T_0} \int_{T_0} x_p(t) e^{-jn\omega_0 t} dt. \tag{3.2}$$

We will refer to the expansion of the periodic waveform $x_p(t)$ in terms of harmonically-related complex exponentials as shown in Equations (3.1–3.2) as the *complex Fourier series* representation of $x_p(t)$. In general, the $\{D_n\}$ coefficients of the complex Fourier series expansion will be complex valued. We will frequently use the nottaion:

$$x_p(t) \Longleftrightarrow \{D_n\}_n \tag{3.3}$$

to represent the equivalence between the time domain (the function $x_p(t)$ itself) and the corresponding Fourier domain representation (the $\{D_n\}$ coefficients). In many cases (particularly, if the waveform $x_p(t)$ has even or odd symmetry) we may find it convenient to expand $x_p(t)$ in terms of an equivalent *trigonometric Fourier series* given as:

$$x_p(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t), \tag{3.4}$$

with

$$a_0 = \frac{1}{T_0}\int_{T_0} x_p(t)dt, \tag{3.5}$$

$$a_n = \frac{2}{T_0}\int_{T_0} x_p(t)\cos n\omega_0 t\,dt, \quad n = 1, 2, \ldots \tag{3.6}$$

$$b_n = \frac{2}{T_0}\int_{T_0} x_p(t)\sin n\omega_0 t\,dt, \quad n = 1, 2, \ldots \tag{3.7}$$

The integrals used in the evaluation of the complex or trigonometric Fourier series coefficients can be taken over any interval of T_0 duration as both the waveform $x_p(t)$ and the complex/trigonometric basis functions of the Fourier series are periodic with period T_0. Also note that the complex Fourier series expansion, Equations (3.1–3.2), and the trigonometric Fourier series expansion, Equations (3.4–3.7) are equivalent and one can be easily converted into the other.

3.1.2 Fourier Transform

Let $x(t)$ be a (possibly non-periodic) function satisfying the Dirichlet conditions[1]. We can determine the Fourier transform of $x(t)$ to be represented by $X(f)$ using the **Fourier integral**:

$$X(f) = \mathcal{F}[\,x(t)\,], \tag{3.8}$$

$$= \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt, \tag{3.9}$$

where $\omega = 2\pi f$ is the angular frequency measured in [rad/s]. We will also use the notation:

$$x(t) \xrightarrow{\mathcal{F}} X(f),$$

to represent the origin/source of the Fourier transform $X(f)$. We can "recover" $x(t)$ from its Fourier transform $X(f)$ by using the inverse-Fourier transform operation:

$$x(t) = \mathcal{F}^{-1}[\,x(t)\,], \tag{3.10}$$

$$= \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df. \tag{3.11}$$

We will also use the notation:

$$x(t) \xleftarrow{\mathcal{F}^{-1}} X(f),$$

to represent how we recover $x(t)$ from its Fourier transform $X(f)$. Equivalently, we will use the symbolic representation:

$$x(t) \iff X(f),$$

to show that $x(t)$ and $X(f)$ form a Fourier transform pair.

[1] The Dirichlet conditions required for the existence of the Fourier transform are: (1) The function $x(t)$ must be absolutely integrable, i.e., $\int_{-\infty}^{\infty}|x(t)|dt < \infty$, and (2) $x(t)$ may have only a finite number of maxima and minima and a finite number of discontinuities.

Remark: The definition of the Fourier transform (and the corresponding inverse Fourier transform) can take many different forms. You may frequently find the normalization term $1/2\pi$ preceding the Fourier integral such that the Fourier and inverse-Fourier transforms are defined as:

$$X(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt,$$
$$x(t) = \int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega.$$

The above representation of the Fourier transform arises from the use of ω as the frequency variable instead of f. Irrespective of which frequency variable is used, both representations are equivalent—using substitution of variables with $\omega = 2\pi f$ and $d\omega = d(2\pi f)$ we can easily establish the equivalence of these representations. In other definitions of the Fourier transform you may even see $1/\sqrt{2\pi}$ as the normalization factor symmetrically distributed to the definitions of both the forward and inverse transforms.

3.1.3 Laplace Transform

We will use the notation $X(s) = \mathcal{L}[\,x(t)\,]$ to refer to the Laplace transform of the function $x(t)$ such that

$$x(t) \iff X(s),$$

will form a Laplace transform pair.

3.2 Motivation

In this course our objective is the *sinusoidal, steady-state analysis* of signals and systems with particular emphasis on communication systems. In particular, we assume that the systems we will study are *linear, time-invariant* (LTI) systems. This is a simplifying assumption that nevertheless allows us to model and study a wide range of practical communication systems with a great degree of accuracy.

Let $h(t)$ be the impulse response function of an N*th* order LTI system with the transfer function $H(s) = \mathcal{L}[\,h(t)\,]$. Let $x(t)$ be the input and $y(t)$ be the corresponding output of the system. Then

$$y(t) = h(t) * x(t).$$

Let $x(t) = A\sin\omega_0 t$ such that

$$X(s) = \frac{A\omega_0}{s^2+\omega_0^2},$$

with

$$H(s) = \frac{N(s)}{(s+p_1)\,\cdots\,(s+p_N)},$$

where $N(s)$ is the Laplace transform of the numerator of $H(s)$. Assume the system poles are simple, i.e., each has multiplicity 1, real-valued and $p_n < 0$, for $n = 1, 2, \ldots, N$ to ensure the stability of the system. Then,

$$
\begin{aligned}
Y(s) &= H(s)\,X(s), && (3.12)\\
&= \frac{N(s)}{(s+p_1)\cdots(s+p_N)}\,\frac{A\omega_0}{s^2+\omega_0^2}, && (3.13)\\
&= \frac{K_0}{s+j\omega_0}+\frac{K_0^*}{s-j\omega_0}+\sum_{n=1}^{N}\frac{K_n}{s+p_n} && (3.14)
\end{aligned}
$$

By evaluating the inverse Laplace transform of $Y(s)$ given in Equation (3.14) we obtain the following expression for the output waveform:

$$
\begin{aligned}
y(t) &= K_0 e^{-j\omega_0 t} + K_0^* e^{j\omega_0 t} + \sum_{n=1}^{N} K_n e^{-p_n t}, && (3.15)\\
&= K\sin(\omega_0 t+\varphi) + \sum_{n=1}^{N} K_n e^{-p_n t}, && (3.16)
\end{aligned}
$$

where $K = A|H(\omega_0)|$ and $\varphi = \arg[\,H(\omega_0\,]$. Since $H(s)$ represents a stable system, we have:

$$\lim_{t\to\infty} K_n e^{-p_n t} = 0, \quad \text{for } n = 1, 2, \ldots, N.$$

Therefore, the steady-state output of the system represented by $y_{ss}(t)$ is described by the expression:

$$
\begin{aligned}
y_{ss}(t) &= \lim_{t\to\infty} y(t),\\
&= K\sin(\omega_0 t+\varphi),\\
&= A|H(\omega_0)|\sin\big(\omega_0 t + arg[\,H(\omega_0)\,]\big).
\end{aligned}
$$

Observations:

1. For a LTI system defined by the transfer function $H(s)$, if $x(t)$ is a sinusoid at frequency ω_0, the steady-state system output $y_{ss}(t)$ is also a sinusoid at the same frequency but with amplitude and phase values modified by $H(s)$ evaluated at $s = j\omega_0$. This approach forms the basis for most tests used to determine if a system described by $H(s)$ is linear.

2. If we restrict our analysis to sinusoidal steady-state analysis we do not need to know $H(s)$ over the entire s-plane since we only need to evaluate $H(s)$ for $s = j\omega$, i.e., along the $j\omega$-axis. In other words, all we need is the Fourier transform of the impulse response function $h(t)$ as

$$\mathcal{F}[\,h(t)\,] = H_{\mathcal{F}}(f) = H(s)\big|_{s=j2\pi f},$$

where we used the notation $H_{\mathcal{F}}(f)$ to explicitly refer to the Fourier transform of $h(t)$.

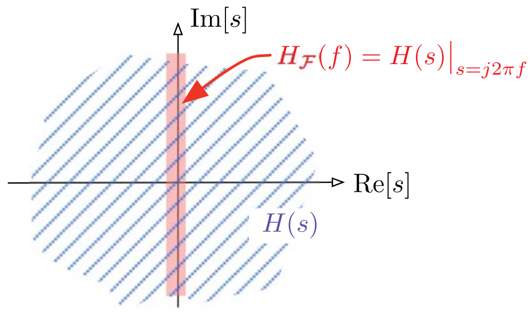


Figure 3.1: How to obtain the Fourier transform from the Laplace transform.

3. If we know $H_{\mathcal{F}}(f)$ then we can easily answer the the question: *"What is the system output when the input is $A \sin \omega_0 t$?"*. Based on the result stated above we can first evaluate $H_{\mathcal{F}}(f)$ at f_0 and then express the steady-state system output as:

$$y(t) = A|H_{\mathcal{F}}(f_0)| \sin\big(2\pi f_0 t + arg[\, H_{\mathcal{F}}(f_0)\,]\big).$$

   This result is very useful if we work all the time with sinusoidal signals only. But simple sinusoids are not very exciting signals. Do we really want to listen to a radio broadcasting consisting of test tones, i.e., single frequency sinusoids, only? As most signals of interest are highly complex, we may want to extend this important result to a much wider class of complex signals.

4. The following table provides an overview of how we can approach to the problem of determining the steady-state output of a LTI system when the input is an arbitrary signal.

Input Signal	**Output Signal**
$x_1(t)$	$y_1(t)$
$x_2(t)$	$y_2(t)$
$\alpha x_1(t) + \beta x_2(t)$	$\alpha y_1(t) + \beta y_2(t)$
$A \sin(\omega_0 t + \theta)$	$A\lvert H_{\mathcal{F}}(f_0)\rvert \sin\big(\omega_0 t + \theta + arg[\, H_{\mathcal{F}}(f_0)\,]\big)$
$e^{j\omega_0 t}$	$\lvert H_{\mathcal{F}}(f_0)\rvert e^{j(\omega_0 t + arg[\, H_{\mathcal{F}}(f_0)\,])}$
$x_p(t) = \sum_n D_n e^{jn\omega_0 t}$	$y_p(t) = \sum_n D_n \lvert H_{\mathcal{F}}(nf_0)\rvert e^{j(n\omega_0 t + arg[\, H_{\mathcal{F}}(nf_0)\,])}$
$x(t)$	$y(t) = \mathcal{F}^{-1}[\, H_{\mathcal{F}}(f) X_{\mathcal{F}}(f)\,] = \int H_{\mathcal{F}}(f) X_{\mathcal{F}}(f) e^{j\omega t} df$

   Observe that the input-output relation shown in the third line is due to the linearity of the system; $x_p(t)$ in the second last line is an arbitrary periodic function expanded in complex Fourier series, and the corresponding system output $y_p(t)$ is obtained again using the linearity of the system. And the last line generalizes the input-output relation to an arbitrary (non-periodic) input signal and expresses the system output $y(t)$ in terms of the system and input Fourier transforms.

Thus, the Fourier analysis is a convenient and very powerful tool; it will allow us to determine the output of LTI systems excited by arbitrary input signals.

5. The Fourier representation (Fourier series for periodic and Fourier transform for non-periodic signals) decomposes a waveform into a discrete or continuous sum of complex exponentials. We also recall that for real-valued waveforms positive and negative frequencies are by-products of the complex exponential representation. Negative frequencies have no physical meaning and a 2-sided spectrum of a real-valued signal can always be represented as a 1-sided spectrum.

6. The time-domain and frequency-domain representations of a signal are comparable to a representation of an image in different file formats, e.g. TIFF, GIF, PNG, Raw, etc.; they all represent the same image and carry the same image information. Furthermore, one can transfer one file format into another by using an appropriately designed file format converter, i.e., transformation.

7. The relative amplitude of frequency components at f' is proportional to $|X(f')|$ (or if we are using a Fourier series representation $|D_{n'}|$ where $n'f_0 = f'$). The units of a Fourier representation is determined by the waveform itself. Let the notation $[\,a\,]_u$ represent the units associated with the signal a. If $x(t)$ is a time waveform with $[\,x(t)\,]_u = gnats$, then the corresponding Fourier series coefficients have $[\,D_n\,]_u = gnats$ as $x(t) = \sum_n D_n e^{jn\omega_0 t}$ and $[\,X(f)\,]_u = gnats/Hz$ since $X(f) = \lim_{T_0 \to \infty}(T_0 D_n)$.

Questions: Why are we interested in steady-state , i.e. Fourier analysis? What happens to the transients? Are the transients not important?

- **Yes, we can ignore transients:** For a stable LTI system, the transient response approaches to zero as $t \to \infty$, and the system output approaches the forced response. Contrary to analysis and design of control systems, we can assume that the communication system has been in operation for a long time such that all transient signals can be ignored . If you compare this approach with that taken in a control system course, we understand why such a difference exists. The very nature of a control system means that we initiate an action to perform a task, hence we want to study the behaviour of the system as the action designed to generate a certain response is applied to the system. In a communication system, however, this is seldom the case.

  In a typical communication system the duration of transients is very short relative to the signal duration. Therefore, we can safely ignore the transients if they do not create problems during the initial start-up phase.

- **No, we cannot ignore transients:** During the initial analysis, design and testing stages we have to consider the effects of transients to ensure that they do not have any detrimental effects on the system performance, e.g. dynamic behaviour of PLL circuits.

3.3 Fourier Analysis

In this section we discuss important characteristics of the Fourier analysis and point out how the Fourier transform can be considered a limiting case of the Fourier series representation. The discussion presented in these notes does not present a rigorous derivation of important results; rather, it demonstrates key concepts by examples and observations. These notes should be used as a supplement to the course reference text.

3.3.1 Fourier Series

Let $x_p(t)$ be a periodic function with period T_0 seconds which also satisfies the Dirichlet conditions. Therefore, we can expand $x_p(t)$ in a complex Fourier series with coefficients $\{D_n\}$. As an example let us consider a periodic train of τ second wide rectangular pulses with amplitude A.

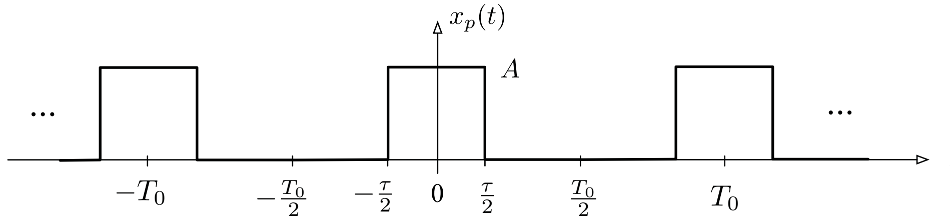


Figure 3.2: The periodic pulse train $x_p(t)$.

We can expand $x_p(t)$ in a complex Fourier series $\sum_n D_n e^{jn\omega_0 t}$ with $\omega_0 = 2\pi/T_0$ such that:

$$
\begin{aligned}
D_n &= \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x_p(t)e^{-jn\omega_0 t}dt,\\
&= \frac{A}{T_0}\int_{-\tau/2}^{\tau/2} e^{-jn\omega_0 t}dt,\\
&= \frac{A}{T_0}\left.\frac{e^{-jn\omega_0 t}}{-jn\omega_0}\right|_{-\tau/2}^{\tau/2},\\
&= \frac{A}{T_0}\left[\frac{e^{-jn\omega_0\tau/2} - e^{jn\omega_0\tau/2}}{-jn\omega_0}\right],\\
&= \frac{2A}{n\omega_0 T_0}\left[\frac{e^{j(.)} - e^{-j(.)}}{2j}\right],\\
&= \frac{A}{n\pi}\sin\Big(\frac{\pi\tau}{T_0}n\Big),\\
&= A\frac{\tau}{T_0}\,\mathrm{sinc}\Big(\frac{\pi\tau}{T_0}n\Big).
\end{aligned}
$$

Thus given the Fourier series representation

$$x_p(t) \longleftrightarrow \left\{ A\frac{\tau}{T_0}\,\text{sinc}\left(\frac{\pi\tau}{T_0}n\right) \right\}_n \tag{3.17}$$

a graphical depiction of the spectrum of $x_p(t)$ is shown in Figure (3.3).

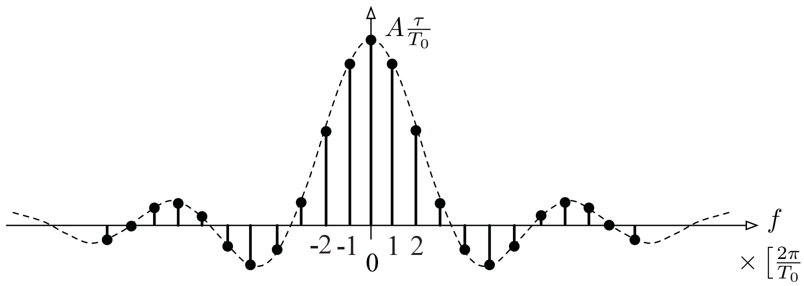


Figure 3.3: Spectrum of $x_p(t)$.

Remarks and Observations:

- **Magnitude and Phase Spectra:** We can separate the information contained in $\{D_n\}$ into a corresponding magnitude spectrum $\{|D_n|\}$ and a phase spectrum $\{\arg[D_n]\}$. In this example all D_n's are real valued; some are positive and some are negative. Observe that if α_+ is a positive real number, then $\arg[\alpha_+] = \arg[|\alpha_+|e^{j0}] = 0$. Conversely, if α_- is a negative real number, then $\arg[\alpha_-] = \arg[|\alpha_-|e^{\pm j\pi}] = \pm\pi$.

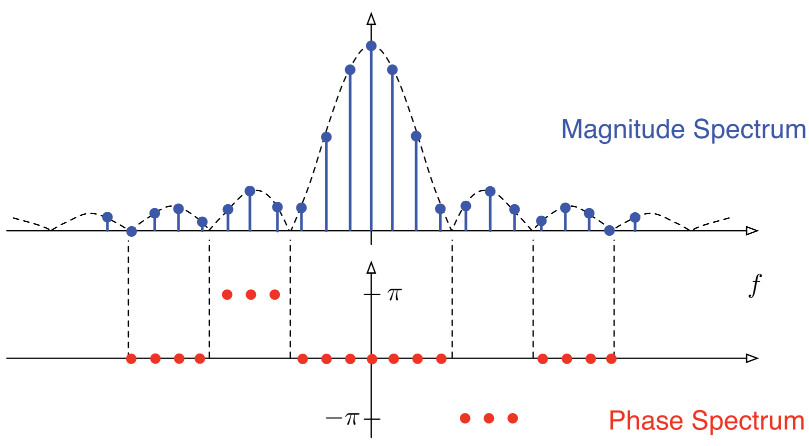


Figure 3.4: Magnitude and phase spectra of $x_p(t)$.

- **Where did the *negative* frequencies come from?** Consider a real-valued periodic signal signal $x_p(t)$ with the Fourier series expansion

$$x_p(t) = \sum_n D_n e^{jn\omega_0 t} = \cdots + D_{-n}e^{-jn\omega_0 t} + D_n e^{jn\omega_0 t} + \cdots \tag{3.18}$$

with $D_{-n} = D_n^*$. Let $D_n = \big|D_n\big|e^{j\phi_n}$ where $\varphi_n = arg[D_n]$. Observing that $D_{-n} = \big|D_n\big|e^{-j\phi_n}$, we can reformulate the two terms on the right-hand side of Equation(3.18) as:

$$D_{-n}e^{-jn\omega_0 t} + D_n e^{jn\omega_0 t} = D_n^* e^{-jn\omega_0 t} + D_n e^{jn\omega_0 t} \tag{3.19}$$

$$= \big|D_n\big|e^{-j(n\omega_0 t+\varphi_n)} + \big|D_n\big|e^{j(n\omega_0 t+\varphi_n)}, \tag{3.20}$$

$$= 2\big|D_n\big|\left[\frac{e^{jn\omega_0 t} + e^{-jn\omega_0 t}}{2}\right], \tag{3.21}$$

$$= 2\big|D_n\big|\cos(n\omega_0 t + \varphi_n). \tag{3.22}$$

Thus, any component with a negative frequency $-n\omega_0$ can be combined with its counterpart term at the positive frequency $n\omega_0$ to yield the sinusoidal component $\cos(n\omega_0 t + \phi_n)$. Hence, the negative frequency components are essential; they simply result from the complex Fourier series expansion of $x_p(t)$. If we had expanded $x_p(t)$ into a trigonometric Fourier series then the combined terms as shown in Equation (3.22) would have already been built in into the calculation of the $\{a_n\}$ and $\{b_n\}$ coefficients of the trigonometric Fourier series expansion.

- **One-sided rms-spectrum:** Let $x(t)$ be the sinusoidal signal $A\cos\omega_0 t$ with the rms value $x_{rms} = A/\sqrt{2}$. The expansion of $x(t)$ into a complex Fourier series can be easily accomplished using the Euler's formula such that

$$x(t) = \frac{A}{2}e^{j\omega_0 t} + \frac{A}{2}e^{-j\omega_0 t}.$$

Spectrum analyzers (such as the ones we will be using in the laboratory) typically display one-sided rms-spectra. For an arbitrary periodic function $x_p(t)$ with complex Fourier series coefficients $\{D_n\}$ we have the Fourier series expansion:

$$x_p(t) = \sum_{n=-\infty}^{\infty} D_n e^{jn\omega_0 t},$$

$$= |D_0| + \sum_{n=1}^{\infty}\Big(D_{-n}e^{-jn\omega_0 t} + D_n e^{jn\omega_0 t}\Big),$$

$$= |D_0| + \sum_{n=1}^{\infty} 2|D_n|\cos(n\omega_0 t + \varphi_n),$$

where as before $\varphi_n = arg[D_n]$. Therefore, the rms values corresponding to each term in the above Fourier series expansion are calculated as:

$$\big[|D_0|\big]_{rms} = |D_0|,$$

$$\big[2|D_n|\big]_{rms} = 2|D_n|/\sqrt{2},$$

$$= \sqrt{2}|D_n|.$$

The two-sided magnitude spectrum of $x_p(t)$ together with the one-sided rms-spectrum is shown in Figure (3.5).

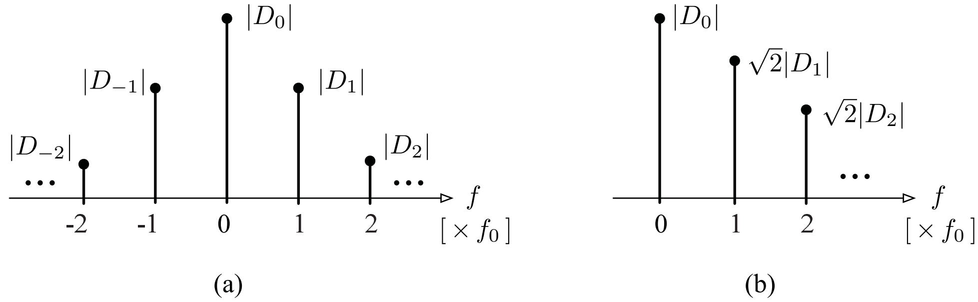


Figure 3.5: (a) Two-sided magnitude spectrum and (b) one-sided rms-spectrum of $x_p(t)$.

3.3.2 Fourier Transform

A non-periodic signal can be obtained from a periodic signal by letting the period of the signal approach to infinity. Consequently, we can show that the Fourier transform of such a non-periodic function can be derived by observing the behaviour of the Fourier series coefficients of a corresponding periodic function as the waveform period is increased. To illustrate this point let us consider the following example. Let $x_p(t)$ be a periodic train of rectangular pulses of τ second duration, amplitude A and period T_0. In Section 3.3.1 we presented $x_p(t)$ and shown that the corresponding Fourier series coefficients are given by the expression:

$$D_n = \frac{A\,\tau}{T_0}\,\text{sinc}\Big(\frac{\pi\tau}{T_0}n\Big), \quad n = 0, \pm1, \pm2, \ldots \tag{3.23}$$

Now consider the corresponding non-periodic function $x(t) = A\Pi(t/\tau)$. Its Fourier transform $X(f)$ can be easily calculated as:

$$\begin{aligned}
X(f) &= \int_{-\infty}^{\infty} A\,\Pi(\frac{t}{\tau}) e^{-j2\pi f t} dt, \\
&= A\int_{-\tau/2}^{\tau/2} e^{-j2\pi f t} dt, \\
&= A\,\frac{e^{-j2\pi f t}}{-j2\pi f}\Bigg|_{-\tau/2}^{\tau/2}, \\
&= \frac{A}{\pi f}\,\Big[\frac{e^{j\pi f\tau} - e^{-j\pi f\tau}}{2j}\Big], \\
&= A\,\frac{\sin(\pi f\tau)}{\pi f}, \\
&= A\tau\,\text{sinc}(\pi f\tau).
\end{aligned}$$

This result establishes the Fourier transform pair:

$$A\,\Pi(\frac{t}{\tau}) \iff A\tau\,\text{sinc}(\pi f\tau). \tag{3.24}$$

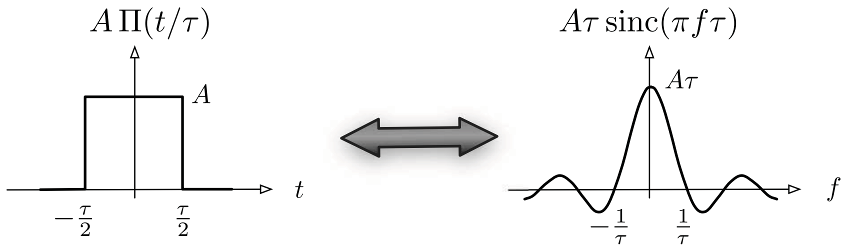


Figure 3.6: The rectangular pulse function $A\Pi(t\tau)$ and its Fourier transform.

As we compare the Fourier series coefficient given in Equation (3.23) and the Fourier transform in Equation (3.24) we observe that

$$\lim_{T_0\to\infty} T_0 \sum_{n=-\infty}^{\infty} D_n\delta(f - \frac{n}{T_0}) = X(f) \tag{3.25}$$

when we recognize that the D_n coefficients located at discrete frequency locations n/T_0 will converge to the continuous frequency variable f as $T_0 \to \infty$. The relation established in Equation (3.25) holds in general and indicates how the Fourier series and transform results are related.

Properties of the δ-function allow us to determine some useful Fourier transform pairs:

$$\mathcal{F}[\delta(t)] = \int \delta(t)e^{-j2\pi ft}dt = e^{-j2\pi ft}|_{t=0} = 1. \tag{3.26}$$

$$\mathcal{F}^{-1}[\delta(f)] = \int \delta(f)e^{j2\pi ft}df = e^{j2\pi ft}|_{f=0} = 1. \tag{3.27}$$

We can now amend our Fourier transform tables by adding the transform pairs:

$$\delta(t) \Longleftrightarrow 1 \tag{3.28}$$

$$1 \Longleftrightarrow \delta(f) \tag{3.29}$$

We can continue with the computation of other useful Fourier transform pairs by using the above results and Fourier transform properties:

$$\mathcal{F}^{-1}[\delta(f - f_0)] = \int \delta(f - f_0)e^{j2\pi ft}df = e^{j2\pi f_0 t}, \tag{3.30}$$

$$\mathcal{F}^{-1}[\delta(f + f_0)] = \int \delta(f + f_0)e^{j2\pi ft}df = e^{-j2\pi f_0 t}, \tag{3.31}$$

$$\mathcal{F}^{-1}\big[\frac{1}{2}\big(\delta(f - f_0) + \delta(f + f_0)\big)\big] = \frac{1}{2}\big(e^{j2\pi f_0} + e^{-j2\pi f_0}\big) \tag{3.32}$$

$$= \cos 2\pi f_0 t, \tag{3.33}$$

which results in the following entry in the Fourier transforms table:

$$\cos 2\pi f_0 t \iff \frac{1}{2}\big[\delta(f + f_0) + \delta(f - f_0)\big]. \tag{3.34}$$

Similarly, by first expressing $\sin 2\pi f_0 t$ as $(e^{j2\pi f_0 t} - e^{-j2\pi f_0 t})/2j$ we can show that:

$$\sin 2\pi f_0 t \iff \frac{j}{2}\big[\delta(f + f_0) - \delta(f - f_0)\big]. \tag{3.35}$$

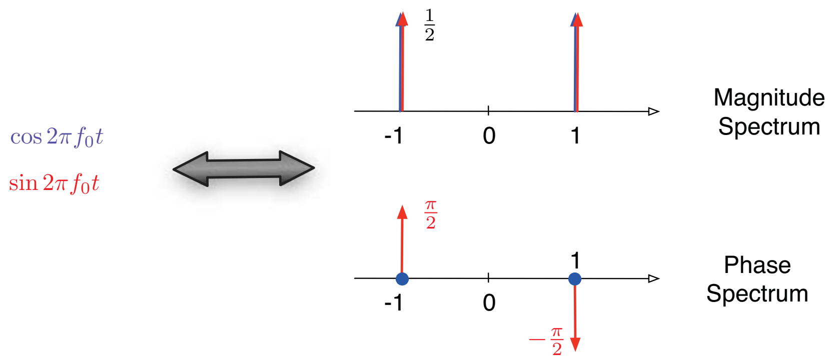


Figure 3.7: Fourier series expansion of a periodic train of δ-functions.

As expected the Fourier transforms of the trigonometric functions $\cos$ and $\sin$ have the same magnitude but different phase spectra.

Observe that τ in the definition of the single pulse function $x(t)$ defines the pulsewidth. And the first zero crossing of the corresponding Fourier transform is at $f = 1/\tau$. Thus, if we consider the main lobe of the sinc function, i.e., $[0, 1/\tau]$ as the range of frequencies in $x(t)$, we observe that this bandwidth is inversely proportional to τ and thus it increases with decreasing τ. This result is expected since decreasing τ values describe shorter pulse duration (in time-domain), and of course shorter duration pulses imply larger signal bandwidth (in frequency-domain). This reciprocal relation between signal duration and signal bandwidth will be one of the important relations in the analysis and design of communication systems and protocols.

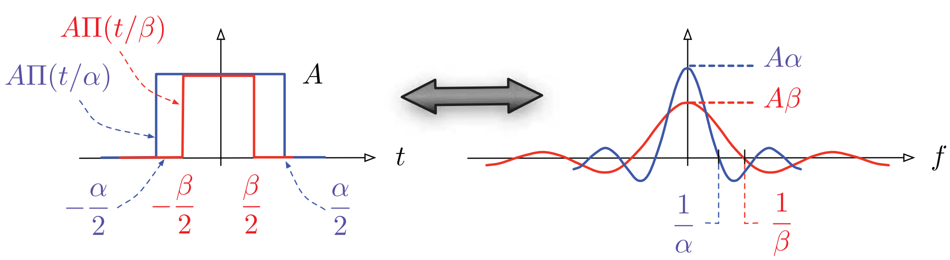


Figure 3.8: Changes in the Fourier transform of $\Pi(t\tau)$ as a function of the pulse width.

3.3.3 Fourier Transform of a Periodic Function

Let $x_p(t)$ be a periodic function with period T_0. $x_p(t)$ has the Fourier series expansion $\{D_n\}_n$ where

$$D_n = \frac{1}{T_0} \int_{T_0} x_p(t) e^{-j2\pi nt/T_0} dt.$$

Let us also consider $x(t)$ as a non-periodic signal which is equivalent to one period of $x_p(t)$ as shown in Figure (3.9). We observe that

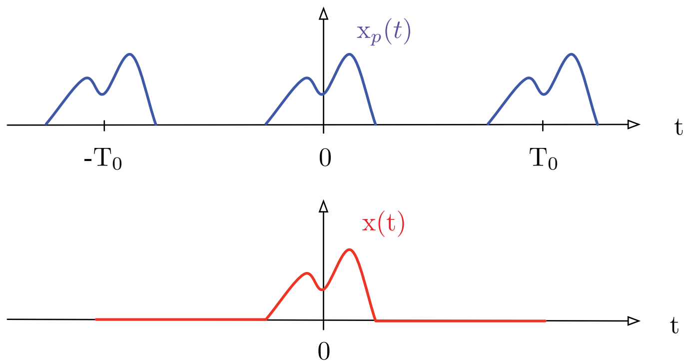


Figure 3.9: Extraction of $x(t)$ from $x_p(t)$.

$$x(t) = \begin{cases} x_p(t), & |t| \leq T_0/2; \\ 0, & \text{otherwise}; \end{cases} \tag{3.36}$$

and $x_p(t) = \sum_n x(t - nT_0)$. Since $x(t)$ is a non-periodic function we can compute its Fourier transform as

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt.$$

But we also have

$$\begin{aligned} D_n &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x_p(t) e^{-j2\pi n t/T_0} dt, \\ &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-j2\pi n t/T_0} dt, \\ &= \frac{1}{T_0} \int_{-\infty}^{\infty} x(t) e^{-j2\pi n t/T_0} dt, \\ &= \frac{1}{T_0} X(\frac{n}{T_0}). \end{aligned}$$

Using the linearity of the Fourier transform we can also write

$$\mathcal{F}\big[x_p(t)\big] = \mathcal{F}\Big[\sum_n D_n e^{j2\pi n t/T_0}\Big], \tag{3.37}$$

$$= \sum_n D_n \, \delta(f - n f_0), \tag{3.38}$$

$$= \frac{1}{T_0} \sum_n X(\frac{n}{T_0}) \, \delta(f - n f_0). \tag{3.39}$$

Example 3.1: Let $x_p(t)$ be a periodic train of δ-functions with unit amplitude defined as: $x_p(t) = \sum_n \delta(t - nT_0)$, where T_0 is the period, $f_0 = 1/T_0$ is the fundamental frequency and $\omega_0 = 2\pi f_0$. As $x_p(t)$ is a periodic function, it can be expanded in an exponential Fourier series with coefficients:

$$
\begin{aligned}
D_n &= \frac{1}{T_0}\int_{T_0} x_p(t)e^{-jn2\pi f_0 t}dt, \\
&= \frac{1}{T_0}\int_{T_0} \delta(t)e^{-jn2\pi f_0 t}dt, \\
&= \frac{1}{T_0}.
\end{aligned}
$$

Thus, the Fourier series expansion of $x_p(t)$ is also a train of δ-functions with constant amplitude $1/T_0$ uniformly spaced at integer multiples of the fundamental frequency f_0.

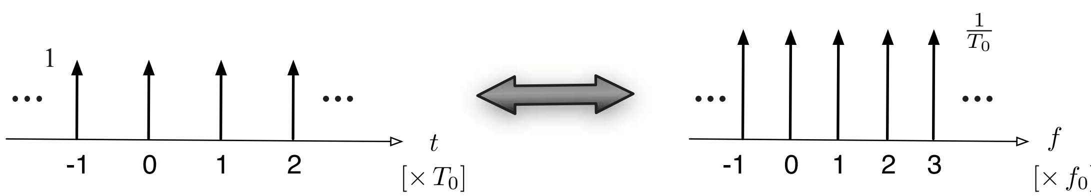


Figure 3.10: Fourier series expansion of a periodic train of δ-functions.

Example 3.2: Let $x_p(t)$ be a periodic train of rectangular pulses of τ second duration, amplitude A, period T_0 and Fourier series coefficients:

$$
D_n = \frac{A\,\tau}{T_0}\,\mathrm{sinc}\Big(\frac{\pi\tau}{T_0}n\Big), \quad n = 0, \pm 1, \pm 2, \ldots
$$

Let $x(t) = A\Pi(t/\tau)$ be the non-periodic waveform extracted from $x_p(t)$ with the Fourier transform

$$
X(f) = A\tau\,\mathrm{sinc}(\pi f \tau).
$$

Using these previously established results and the relation presented in Equation (3.39) we can express the Fourier transform of $x_p(t)$ as:

$$
\begin{aligned}
\mathcal{F}\big[x_p(t)\big] &= \frac{1}{T_0}\sum_n X\big(\frac{n}{T_0}\big)\,\delta(f - nf_0), \\
&= \frac{A\,\tau}{T_0}\sum_n \mathrm{sinc}\big(\frac{\pi n \tau}{T_0}\big)\,\delta(f - nf_0).
\end{aligned}
$$

Example 3.3: Let $x_p(t) = \sum_n \delta(t - nT_0)$ be a periodic train of unit impulse functions spaced T_0 seconds apart. To determine its Fourier transform $X_p(f)$, we first isolate one period of $x_p(t)$

such that $x(t) = \delta(t)$ which has the Fourier transform $X(f) = \mathcal{F}[\delta(t)] = 1$. Therefore,

$$\begin{aligned}\mathcal{F}\big[x_p(t)\big] &= \frac{1}{T_0}\sum_n X(\frac{n}{T_0})\,\delta(f - nf_0),\\ &= \frac{1}{T_0}\sum_n \delta(f - nf_0).\end{aligned}$$

Hence, the Fourier transform of a periodic train of unit impulse functions spaced T_0 seconds apart is also a train of impulse functions scaled in amplitude by $1/T_0$ and spaced $f_0 = 1/T_0$ Hz apart. This is an important result that we will refer to when discussing sampling of analog waveforms. Example 2.1 illustrates this result and depicts the Fourier transform of a periodic train of impulse functions in Figure (3.10).

3.4 Properties of the Fourier Transform

There are many useful properties of the Fourier transform. You have seen these properties (and their variants in the context of different frequency transformations such as the Laplace transform). We will not discuss the derivation of these properties in any detail, it will be your responsibility to read and understand them. Please refer to *Section 3.3* of the course reference text for a detailed discussion, and to *Table 3.2 Properties of Fourier Transform Operations* presented on page 123 of course reference text. We will only discuss the **frequency-shifting/modulation** property as modulation is the basic operation that underlies almost all communication systems, as in *amplitude modulation* (AM) and *frequency modulation* (FM).

Frequency-Shifting property

Let $x(t)$ be a waveform with the Fourier transform $X(f)$. Then

$$x(t)\,e^{j2\pi f_0 t} \iff X(f - f_0). \tag{3.40}$$

Proof:

$$\mathcal{F}\big[\,x(t)\,e^{j2\pi f_0 t}\,\big] = \int_{-\infty}^{\infty} x(t)\,e^{j2\pi f_0 t}\,e^{-j2\pi f t}dt, \tag{3.41}$$

$$= \int_{-\infty}^{\infty} x(t)\,e^{-j2\pi(f-f_0)t}dt, \tag{3.42}$$

$$= X(f - f_0). \tag{3.43}$$

Thus, the time-domain multiplication of $x(t)$ with the complex exponential $e^{j2\pi f_0 t}$ will result in frequency-shifting of $X(f)$ such that it will be centered on f_0. Similarly, if were to multiply $x(t)$ with $e^{-j2\pi f_0 t}$ then $X(f)$ will be shifted in frequency and centered on $-f_0$. We can now use this property together with the linearity of the Fourier transform to determine the effect of multiplying a signal with a sinusoid. Using Euler's identity we first expand $\cos 2\pi f_0 t = (e^{j2\pi f_0 t} + e^{-j2\pi f_0 t})/2$ such that:

$$\mathcal{F}\big[\,x(t)\cos 2\pi f_0 t\,\big] = \int_{-\infty}^{\infty} x(t)\cos 2\pi f_0 t\,dt, \tag{3.44}$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x(t)\left(e^{j2\pi f_0 t} + e^{-j2\pi f_0 t}\right) e^{-j2\pi f t} dt, \tag{3.45}$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \left(x(t)\, e^{-j2\pi(f-f_0)t} + x(t)\, e^{-j2\pi(f+f_0)t}\right) dt, \tag{3.46}$$

$$= \frac{1}{2}\big[\,X(f-f_0) + X(f+f_0)\,\big]. \tag{3.47}$$

Similarly, the effect of multiplying $x(t)$ with $\sin 2\pi f_0 t = (e^{j2\pi f_0 t} - e^{-j2\pi f_0 t})/2j$ becomes:

$$\mathcal{F}\big[\,x(t)\,\sin 2\pi f_0 t\,\big] = \frac{1}{2j}\big[\,X(f-f_0) - X(f+f_0)\,\big], \tag{3.48}$$

$$= \frac{1}{2}\big[\,X(f-f_0)e^{-j\pi/2} + X(f+f_0)e^{j\pi/2}\,\big]. \tag{3.49}$$

Of course, we could have derived these results using the Fourier transforms of sinusoids shown in Equations (3.34–3.35) together with the *frequency convolution* property, $\mathcal{F}[x_1(t)x_2(t)] = X_1(f) * X_f(f)$. Figure (3.11) demonstrates the effects of multiplying a signal with a sinusoid. If the spectrum of the signal $x(t)$ covers the frequency band $[-B, B]$, then $\mathcal{F}[x(t)\cos\omega_0 t]$ and $\mathcal{F}[x(t)\sin\omega_0 t]$ are both centered at $\pm f_0$ and occupy the frequency bands $[-B-f_0,\ -f_0+B]$ and $[f_0-B,\ f_0+B]$. Frequency-shifting the signal up and down the frequency band is known as **modulation**.

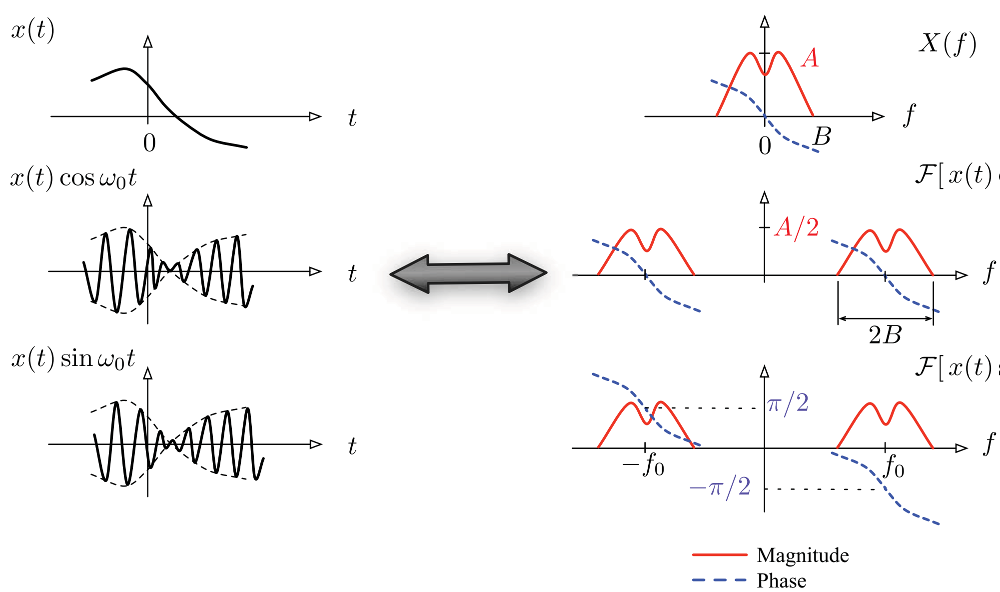


Figure 3.11: Demonstration of the frequency-shifting property.

Example 3.3: We now want to extend the preceding discussion to demonstrate how we can combine two signals for transmission over a communication channel. Let $x_1(t)$ and $x_2(t)$ be two waveforms with spectra $X_1(f)$ and $X_2(f)$, respectively, and let $y(t) = x_1(t)\cos 2\pi f_1 t + x_2(t)\cos 2\pi f_2 t$. We want to determine any constraints on the modulating frequencies f_1 and f_2 such that the spectra of $X_1(f)$ and $X_2(f)$ can be recovered from $Y(f) = \mathcal{F}[y(t)]$, i.e., we should be able to identify the spectra of the constituent signals within $Y(f)$.

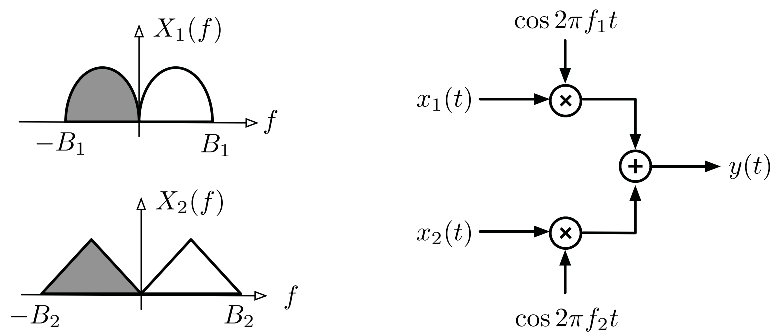


Figure 3.12: Combining signals per Example 3.3.

The first step of our analysis will be to sketch the spectrum of the composite signal $Y(f)$, identify all critical frequencies and attempt to formulate constraints on the modulating frequencies f_1 and f_2. In particular, a key consideration for identifying $X_1(f)$ and $X_2(f)$ in $Y(f)$, is that the frequency-shifted versions of $X_1(f)$ and $X_2(f)$ should not overlap:

Case 1: $f_1 > f_2$

$$f_2 + B_2 \le f_1 - B_1 \implies f_1 - f_2 \ge B_1 + B_2;$$

Case 2: $f_1 < f_2$

$$f_1 + B_1 \le f_2 - B_2 \implies f_2 - f_1 \ge B_1 + B_2.$$

Combining these observation into one inequality yields the result:

$$\left| f_1 - f_2 \right| \ge B_1 + B_2.$$

This example illustrates *frequency division multiplexing*, a technique that allows multiple signals to share the same channel: each signal is allocated part of the frequency spectrum.

3.5 Bandwidth

In our discussion of the frequency-shifting/modulation property we have observed that to transmit a signal $x_1(t)$ which covers the frequency band $[0, B]$ (we consider only the positive frequency values to determine the frequency content of a signal or system) over a communications channel, the channel should accommodate the $[0, B]$ frequency band, i.e., the channel should have a minimum bandwidth of B Hz. On the other hand, to transmit the modulated waveform $x(t)\cos\omega_0 t$ we

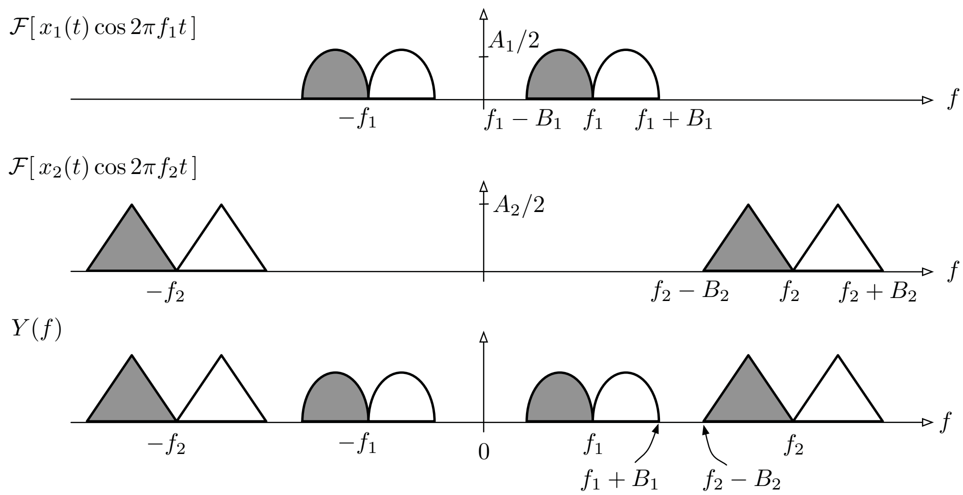
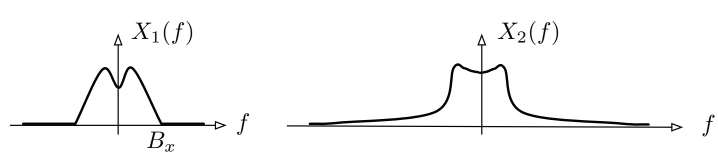


Figure 3.13: Spectra of the modulated and combined waveforms in Example 3.3 (Case 2).

require a communications channel with twice the bandwidth as the modulated waveform occupies the frequency band $[f_0 - B, f_0 + B]$.

The concept of *bandwidth* is an important element in the study of communication systems. We will now proceed to define and discuss the concept of bandwidth and the classification of signals based on their frequency contents. In particular, **bandwidth** is a measure of the extent of significant spectral components of the signal for *positive* frequencies. The definition of bandwidth for systems is a measure of the range of frequencies thac the system can handle. What we mean by "*handled*" will be explained as we discuss the concept of *distortionless transmission*. However, as a brief introduction to the concept of the bandwidth of a system, you may think system bandwidth as being the frequency-domain equivalence of dynamic range.

Baseband/Lowpass Signals: A baseband/lowpass waveform has significant spectral magnitude components for frequencies in the vicinity of the frequency origin ($f = 0$) with the spectral magnitude components being negligible elsewhere. Observe that if a signal $x(t)$ is **strictly band-limited**, then $|X(f)| = 0$ for $f > B_x$ as demonstrated in Figure (3.14a). However, Figure (3.14b) also demonstrates a baseband/lowpass signal; in this case $|X_2(f)|$ rapidly approaches zero yet never becomingzero beyond a certain frequency. Below, we will discuss in some detail how to measure the bandwidth of such not strictly band-limited signals.



(a) Strictly band-limited (b) Not strictly band-limited

Figure 3.14: Spectral characteristics of baseband/lowpass signals.

Bandpass Signals: A bandpass waveform has spectral magnitude components that are non-zero for frequencies in some band concentrated at $\pm f_c$ with $f_c \gg 0$. The spectral magnitude components are negligible elsewhere. We recall that if $x(t)$ is a baseband signal with band-

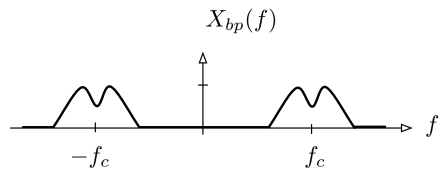


Figure 3.15: Spectrum of a bandpass signal.

width B-Hz, then $x(t)\cos\omega_c t$ and $x(t)\sin\omega_c t$ are bandpass signals both with bandwidth $2B$. Further, the modulated waveforms $x(t)\cos\omega_c t$ and $x(t)\sin\omega_c t$ occupy the frequency band $|f \pm f_c| \leq 2B$. Therefore, a generic bandpass signal can be represented as

$$x_{bp}(t) = x_c(t)\cos\omega_c t + x_s(t)\sin\omega_c t, \tag{3.50}$$

where $x_c(t)$ and $x_s(t)$ are baseband signals each with bandwidth B-Hz. Observe that the spectrum of $x_{bp}(t)$ is not necessarily symmetric with respect to $\pm f_c$. We can also express $x_{bp}(t)$ as:

$$x_{bp}(t) = E(t)\cos\big(\omega_c t + \varphi(t)\big), \tag{3.51}$$

with

$$E(t) = \sqrt{x_c^2(t) + x_s^2(t)}, \tag{3.52}$$

$$\varphi(t) = \arctan\left[\frac{x_s(t)}{x_c(t)}\right]. \tag{3.53}$$

Since $x_c(t)$ and $x_s(t)$ are baseband signals, so are $E(t)$ and $\varphi(t)$. In particular $E(t)$ is the slowly-varying (baseband/lowpass) *envelope* signal and $\varphi(t)$ is the slowly-varying *phase* signal. Consequently, the bandpass signal $x_{bp}(t)$ will appear as a high-frequency signal with slowly varying amplitude and slowly varying frequency.

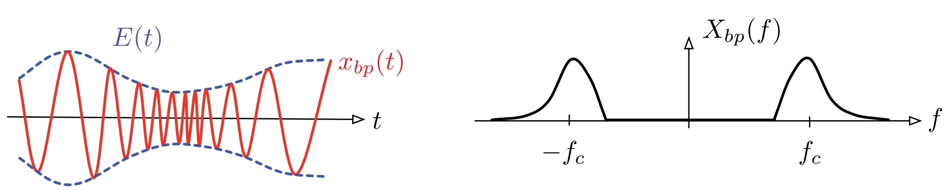


Figure 3.16: Time- and frequency-domain display of x_{bp} described in Equations (3.50–3.53).

Strictly Band-limited Signals: If a signal is **strictly band-limited**, then the definition of bandwidth is obvious. A strictly band-limited signal will have non-zero spectral magnitude components only over a finite frequency band. Therefore, for a strictly band-limited baseband signal $x_{lp}(t)$ with $|X_{lp}(f)| = 0$ for $|f| \geq B_x$, the signal bandwidth equals B_x-Hz. Similarly for strictly band-limited bandpass signal $x_{bp}(t)$ with $|X_{bp}(f)| = 0$ for $|f - f_c| \geq B_x$, the signal bandwidth equals $2B_x$-Hz.

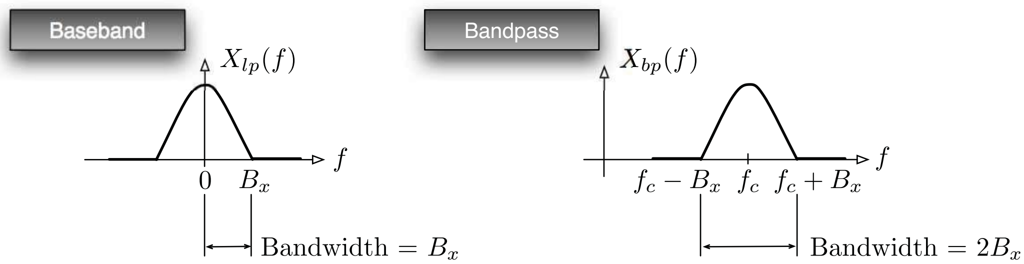


Figure 3.17: Bandwidth measurement for strictly band-limited signals.

Bandwidth Measurement for Not Strictly Band-limited Signals: If the signal is **not** strictly band-limited, then we can use different criteria to measure the signal bandwidth. We will now introduce two of the most commonly used criteria.

- **Main-Lobe Bandwidth Measurement:** In this case, the bandwidth is defined as the bandwidth of the main-lobe of the magnitude spectrum of the signal with the understanding that the main-lobe of the magnitude spectrum will include the significant frequency components. For example, if $x_{lp}(t) = A\Pi(t/\tau)$ then $X_{lp}(f) = A\tau\text{sinc}(\pi\tau f)$ such that the main-lobe of magnitude spectrum $|X_{lp}(f)|$ covers the baseband frequencies $[0, 1/\tau]$ and therefore the main-lobe bandwidth of $x_{lp}(t)$ is $(1/\tau)$-Hz. In the case of a bandpass signal such as $x_{bp}(t) = x_{lp}(t)\cos\omega_c t = A\Pi(t/\tau)\cos\omega_c t$, we measure the frequency band occupied by the main-lobe of the modulated waveform (centered about f_c). In this particular example the main-lobe of the magnitude spectrum $|X_{bp}(f)|$ covers the frequency band $[f_c-(1/\tau), f_c+(1/\tau)]$. Therefore, the main-lobe bandwidth of x_{bp} is $2/\tau$.

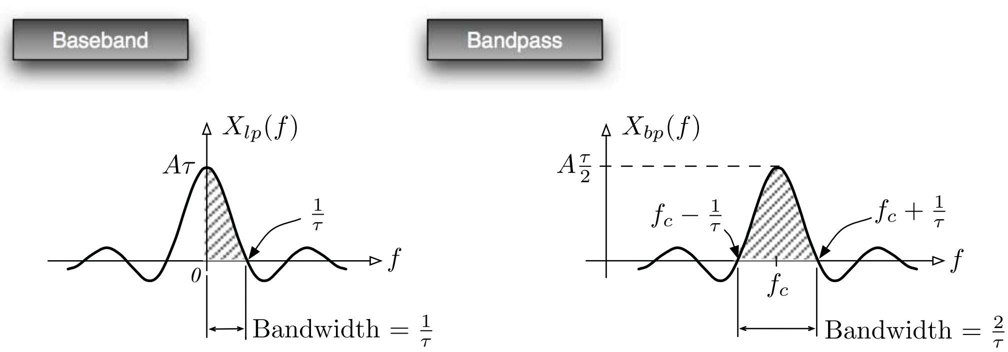


Figure 3.18: Bandwidth measurement for not strictly band-limited signals.

- **3-dB/Half Power Bandwidth Measurement:** An alternative measure of signal bandwidth is the so called 3-dB bandwidth. Typical (but not all) lowpass signals have magnitude spectra which are maximum at the origin and decrease monotonically with increasing frequency. Similarly, typical bandpass signal have magnitude spectra which are maximum at the centre frequency f_c, which decrease monotonically as the frequency variable moves away from f_c. For such signals we measure the 3-dB bandwidth as the frequency where the power spectrum drops 3-dB relative to its maximum (if the signal is $x(t)$ then $|X(f)|^2$ is its power spectrum). Let $x_{lp}(t)$ be lowpass/baseband signal with power spectrum $|X_{lp}(f)|^2$ such that $\max_f |X_{lp}(f)|^2 = |X_{lp}(0)|^2$. Observe that the 3-dB drop in the power spectrum relative to its maximum is equivalent to halving the maximum power measured in decibels[2] expressed using the unit notation [dB].

$$\begin{aligned} 10\log_{10}\left[\frac{|X_{lp}(0)|^2}{2}\right] &= 10\log_{10}\left[|X_{lp}(0)|^2\right] - 10\log_{10}2, \\ &= 10\log_{10}\left[|X_{lp}(0)|^2\right] - 3\text{ dB}. \end{aligned}$$

Figure (3.19) depicts the measurement of the 3-dB bandwidth for baseband signals. We

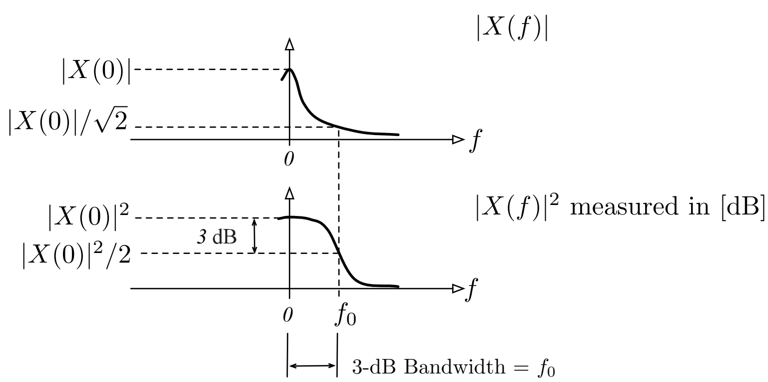


Figure 3.19: 3-dB Bandwidth measurement for baseband signals.

measure the 3-dB bandwidth of bandpass signals using a process similar to the one we used to measure the main-lobe bandwidth, namely the measurements are taken with respect to the centre frequency of the bandpass signal. Figure (3.20) demonstrates this process.

- There are other definitions of bandwidth measurement criteria, e.g. the *rms*-bandwidth. We will define and explain such criteria as required.

Remarks:

- The bandwidth and time-duration of a signal are inversely related and cannot be independently altered, i.e., $[B_x][\text{time-duration}] \geq K$ where B_x is the signal bandwidth and K is constant. This inequality is another manifestation of the **Heisenberg uncertainty principle**.

[2] For a given quantity x let x_{lin} be its measurement using a linear scale and let x_{dB} be its measurement using the dB-scale, then $x_{\text{dB}} = 10\log_{10}[\,x_{\text{lin}}\,]$.

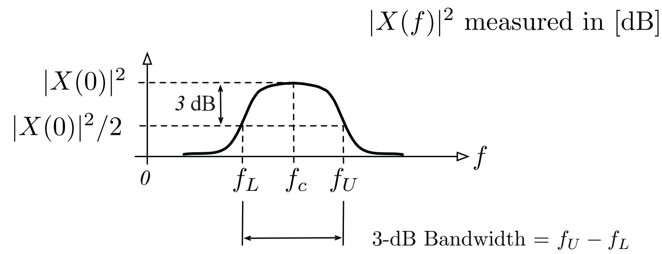


Figure 3.20: 3-dB Bandwidth measurement for bandpass signals.

- If x_1 has a bandwidth of B_1-Hz and x_2 has a bandwidth of B_2-Hz, then x_1x_2 has a a bandwidth of $(B_1 + B_2)$-Hz. We can easily validate this result by using the frequency convolution property of the Fourier transform, namely $\mathcal{F}[x_1(t)x_2(t)] = X_1(f) * X_f(f)$.

3.6 Signal Transmission Through a Linear System

Let $h(t)$ be the impulse response of a LTI system. By definition, $h(t)$ is the response of the system at rest, i.e., $h(t)$ is the system output observed when all system initial conditions are zero with a unit impulse function $\delta(t)$ applied to the input of the system at time $t = 0$. The impulse response function $h(t)$ fully describes the system and allow us to determine the output of the system for an arbitrary input function using the convolution operation: if $x(t)$ is the input, then the system output equals $y(t) = h(t) * x(t)$. Where does the *convolution* operation come from?

1. If the system input equals $K\delta(t - t_0)$, the sytem output is $Kh(t - t_0)$. This result is simply due to the LTI characteristic of the system (a scaled and time-shifted unit impulse function will generate a scaled and time-shifted impulse response function as the system output).

2. We can approximate any arbitrary wave form $x(t)$ as

$$x(t) \approx \sum_{\lambda} \left[x(\lambda)\Delta\lambda \right] \delta(t - \lambda).$$

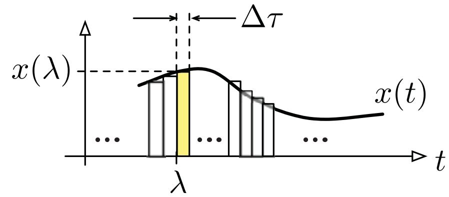


Figure 3.21: Approximation of an arbitrary waveform $x(t)$.

3. If the input to the system is $[x(\lambda)\Delta\lambda]\,\delta(t-\lambda)$, then the system output equals $[x(\lambda)\Delta\lambda]\,h(t-\lambda)$ where we used the definition of the impulse response and the LTI characteristics of the system. Using the linearity of the system once again, we observe that if $\sum_\lambda [x(\lambda)\Delta\lambda]\,\delta(t-\lambda)$ is the system input then $\sum_\lambda [x(\lambda)\Delta\lambda]\,h(t-\lambda)$ will be the corresponding output.

4. The accuracy of the approximation of $x(t)$ will improve as $\Delta\lambda \to 0$. In this case we will have:

$$\begin{aligned} y(t) &= \lim_{\Delta\lambda\to 0}\Big[\sum_\lambda [x(\lambda)\Delta\lambda]\,h(t-\lambda)\Big], \\ &= \int_{-\infty}^{\infty} x(\lambda)h(t-\lambda)d\lambda, \\ &= h(t) * x(t). \end{aligned}$$

Important system properties can be deduced from the impulse response function $h(t)$.

Causality: If $h(t) = 0$ for $t < 0$, then the system will be causal. The causality condition is necessary for system realizability.

BIBO Stability: A system is *bounded-input, bounded-output* (BIBO) stable if $\max_t |y(t)| < \infty$ when $\max_t |x(t)| < \infty$. We can show that the system will be BIBO stable, if $\int |h(t)|dt < \infty$.

3.7 Frequency Response

Earlier we provided a physical interpretation of what $X(f)$ represents: for a given waveform $x(t)$ (voltage, current,...), $X(f)$ represents a density function describing the contribution of the continuous frequency components to the composition of $x(t)$. We will now use this interpretation and extend it to the description of a LTI system. In particular, we will use the term **frequency response** to describe a LTI system. In Section 3.2 we have shown that a sinusoidal input (complex or real) to a LTI system generates a amplitude-scaled and phase-shifted sinusoid at the same frequency as the input. Let $x(t) = e^{j\omega_0 t}$ be the input to a LTI system decribed by the impulse response function $h(t)$, then the system output is given by the exporession:

$$y(t) = |H(f_0)|e^{j(2\pi f_0 t + arg[\,H(f_0)\,])}, \tag{3.54}$$

where

$$H(f) = \mathcal{F}\big[\,h(t)\,\big], \tag{3.55}$$

is the **frequency response** function of the system. $H(f)$ exerts its characteristics on the system output by modifying the amplitude and phase of the frequency components in $x(t)$.

Example 3.4:

Hence, $H(f)$ summarizes how the system will affect different frequency components as they pass through the system. Using Fourier analysis, we can write

$$Y(f) = H(f)X(f).$$

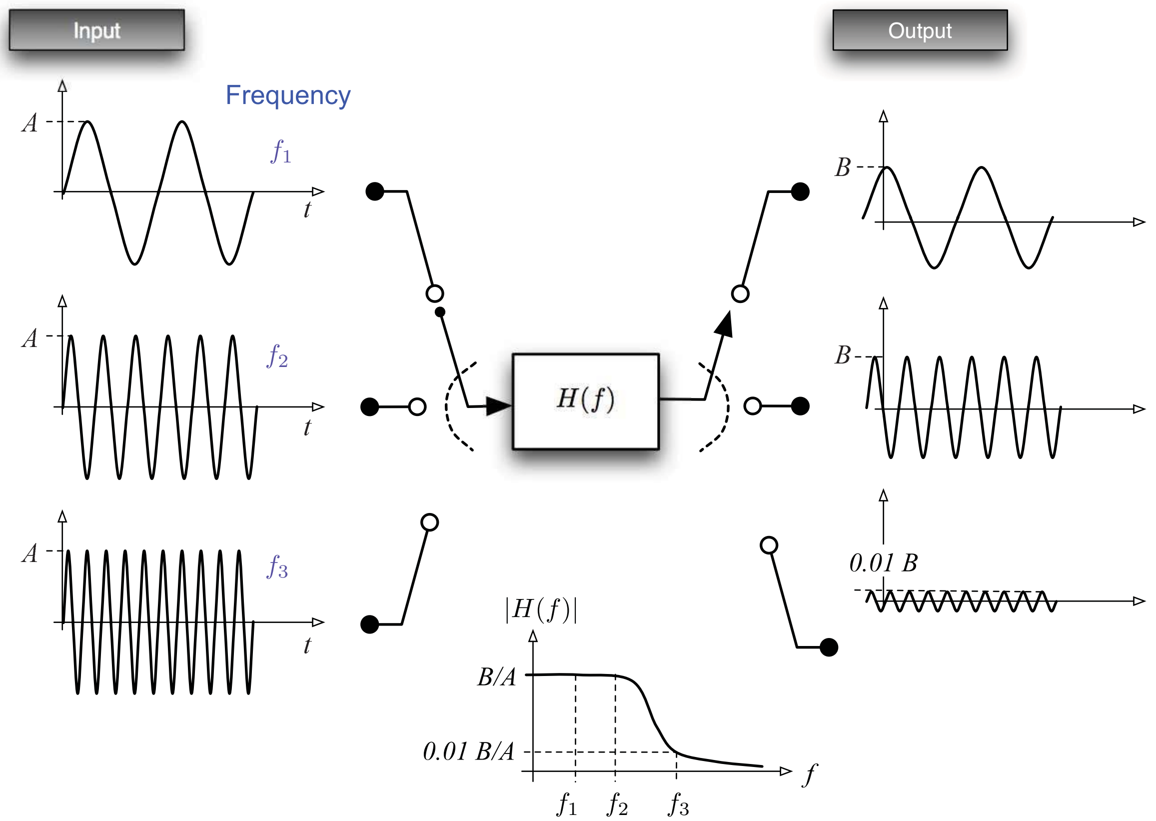


Figure 3.22: Determining the frequency/magnitude response of a LTI system.

In our analysis the most important elements are $H(f)$, the system bandwidth B_h, the bandwidth of the input signal B_x, and how B_x is related to B_h. As a further example, consider a case as demonstrated in the above figure where a unit pulse function of τ-second duration is applied to a lowpass system with adjustable bandwidth.

3.8 Distortionless Transmission

Distortionless transmission is an ideal case that is physically impossible to implement (as we will show). However, studying the properties of such an ideal case allows us to establish a benchmark against which we can measure the performance of non-ideal but realizable systems.

Definition: A **distortionless system** with input $x(t)$ and $y(t)$ satisfies the input-output relation:

$$y(t) = K\,x(t - t_0), \tag{3.56}$$

where $K \in \Re$ is the gain factor, and $t_0 \in \Re$ is the system delay.

The frequency domain characteristics of a distortionless system can be obtained by taking the Fourier transform of Equation (3.56) and then determining the corresponding systems function

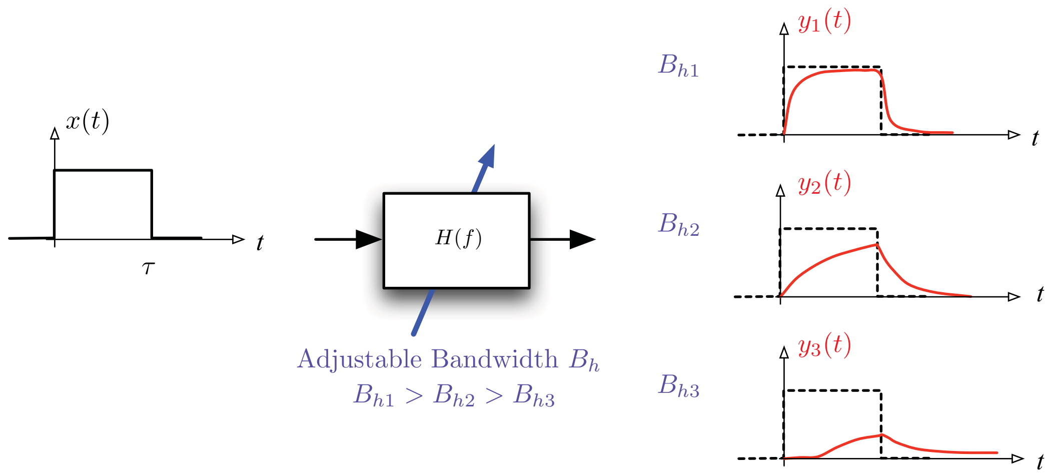


Figure 3.23: Effect of system bandwidth on the system response to a unit pulse.

$H(f)$. As we have

$$Y(f) = H(f)X(f),$$

with

$$Y(f) = \mathcal{F}[\,y(t)\,], \tag{3.57}$$

$$= \mathcal{F}[\,Kx(t-t_0)\,], \tag{3.58}$$

$$= K\,\mathcal{F}[\,x(t-t_0)\,], \tag{3.59}$$

$$= KX(f)e^{-j2\pi f t_0}, \tag{3.60}$$

it then follows:

$$H(f) = \frac{Y(f)}{X(f)} = K\,e^{-j2\pi f t_0}. \tag{3.61}$$

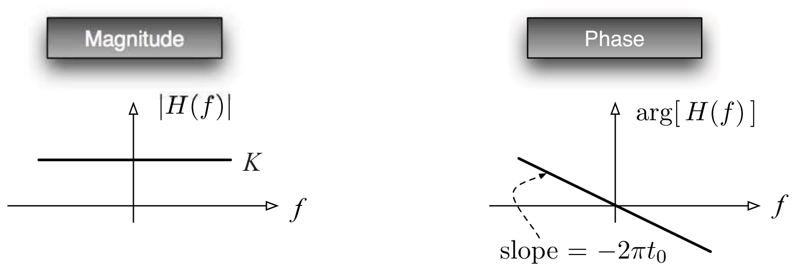


Figure 3.24: Magnitude and phase response of an ideal distortionless transmission system.

Since having a physically realizable with $H(f) = K$ for $-\infty < f < \infty$ is not possible, in practice we require that $|H(f)| = K$ only over a specific frequency band, let's say $[f_L,\ f_U]$. In this case, we may consider the $H(f)$ of the corresponding "ideal band-limited" distortionless system will be the same as depicted in Figure (3.24) but seen through a narrow window over the frequency band $[f_L,\ f_U]$. Figure (3.25) demonstrates this behavior.

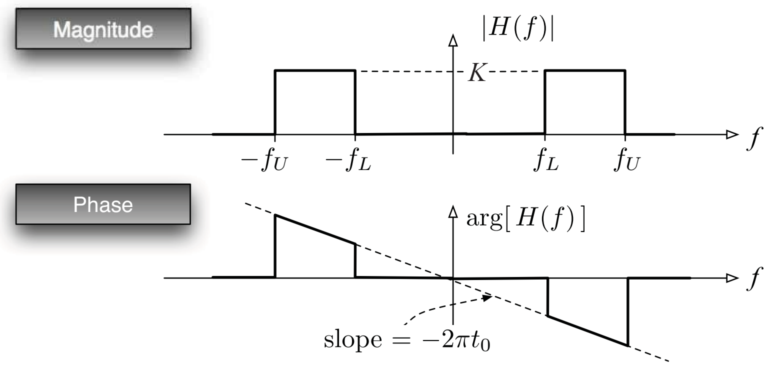


Figure 3.25: Frequency and phase response functions of an ideal distortionless transmission system band-limited to the frequency band $[f_L,\ f_U]$.

3.8.1 Approximation for Distortionless Transmission

In many cases achieving exactly a distortionless transmission system even over a limited frequency range may not be a realizable option. As an example consider the following RC-lowpass system:

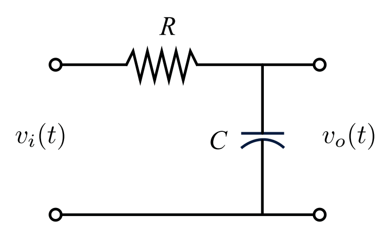


Figure 3.26: The RC-lowpass system.

with

$$H_{rc}(f) = \frac{1}{1 + j2\pi f RC}$$

such that

$$|H_{rc}(f)| = \frac{1}{\sqrt{1 + (2\pi f RC)^2}}, \quad \text{and} \quad \arg[H_{rc}(f)] = -\tan^{-1}(2\pi f RC). \tag{3.62}$$

We can also express the magnitude and phase response functions in terms of the 3-dB bandwidth parameter f_0. Let f_0 be the the 3-dB bandwidth of H_{rc} such that:

$$|H_{rc}(f_0)|^2 = \frac{|H_{rc}(0)|^2}{2} \quad \Longrightarrow \quad \frac{1}{2} = \frac{1}{1 + (2\pi f_0 RC)^2} \quad \Longrightarrow \quad f_0 = \frac{1}{2\pi RC}. \tag{3.63}$$

Equation (3.62) can then be re-written in terms of f_0:

$$|H_{rc}(f)| = \frac{1}{\sqrt{1 + (f/f_0)^2}}, \qquad \text{and} \qquad \arg[H_{rc}(f)] = -\tan^{-1}(f/f_0). \tag{3.64}$$

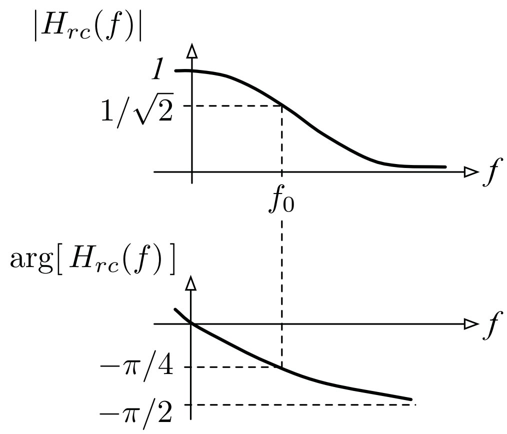


Figure 3.27: Magnitude and phase spectra of the RC-lowpass system.

Clearly, $H_{rc}(f)$ does not represent a distortionless system. Therefore, let us consider the case where we want to have a system which satisfies the requirements for distortionless transmission within specified constraints. In particular:

- Consider only the magnitude response, $|H_{rc}(f)|$;
- Allow maximum deviation from from distortionless transmission to be no more than 10% for $f \in [0,\ 10]$ kHz.

As $|H_{rc}(f)|$ is a monotonically decreasing function of f with $|H_{rc}(f)|_{max} = |H_{rc}(0)| = 1$, we can solve for the system parameter RC (or f_0) such that:

$$\frac{1}{\sqrt{1+(f_1/f_0)^2}} = 0.9 \quad \text{with} \quad f_1 = 10\,\text{kHz} \quad \text{and} \quad f_0 = \frac{1}{2\pi RC}. \tag{3.65}$$

Solving for for the system parameter RC we obtain $RC = 7.7 \times 10^{-6}$. We can now approximate the output of H_{rc} for $f \in [0,\ 10]$ kHz by taking the phase response into consideration:
We determine the phase response at $f_1 = 10$ kHz, by evaluating $\arg[H_{rc}(f_1)] = -\tan^{-1}(f_1/f_0) = -0.451$, where $f_0 = \left(2\pi \times 7.7 \times 10^{-6}\right)^{-1}$. We can now approximate $\arg[H_{rc}]$ over the frequency band $[0, 10]$ kHz by a straight-line (linear phase response) with slope:

$$\text{slope} = \frac{\arg[H_{rc}(f_1)] - \arg[H_{rc}(0)]}{f_1 - 0} = \frac{-0.451}{10^4} = -0.451 \times 10^{-4}. \tag{3.66}$$

From Equation (3.61) we recognize that the system delay t_0 and the slope of the linear phase response (or its approximation as we are discussing right now) are related through the relation:

$$\text{slope} = -2\pi t_0. \tag{3.67}$$

Solving for t_0 we obtain $t_0 = 7.2\ \mu$s. To conclude, the RC-lowpass system described by the frequency response function $H_rc(f)$ approximately provides distortionless transmission for $f \in [0, 10]$ kHz such that the the system output is given by the expression:

$$y(t) \approx Kx(t - t_0),$$

with $K \in [0.9, 1]$ and $t_0 = 7.2\ \mu$s.

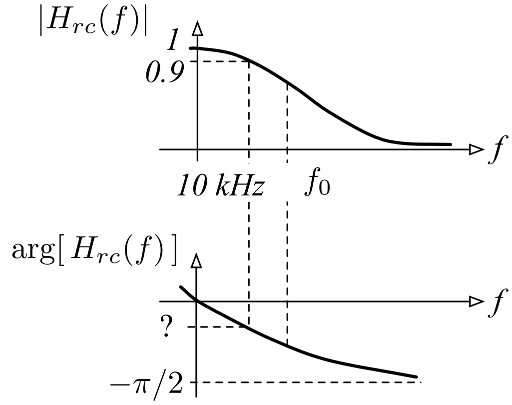


Figure 3.28: Approximating the RC-lowpass system for a distortionless transmission system.

3.8.2 Response of an Ideal Lowpass Filter

Let us consider an ideal lowpass filter (LPF) with the magnitude and phase response function: The

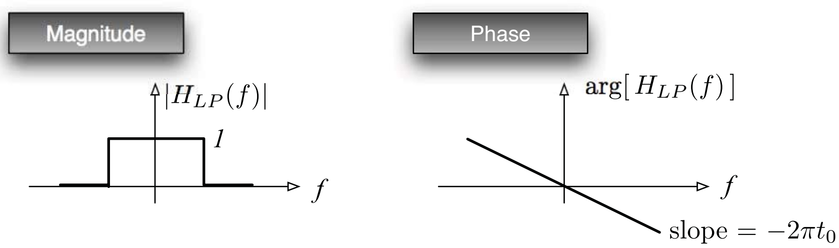


Figure 3.29: Magnitude and phase response of an ideal lowpass filter.

corresponding frequency response is then given by the expression:

$$H_{LP}(f) = \Pi\left(\frac{f}{2B}\right) e^{-j2\pi f t_0}. \tag{3.68}$$

Using the table of Fourier transforms (Lathi and Ding, Table 3.1, p. 107) and the table of properties of Fourier transform operations (Lathi and Ding, Table 3.2, p. 123) and we determine the impulse response function of the ideal lowpass filter from Equation (3.68):

$$h_{LP}(t) = 2B\,\text{sinc}\big(2\pi B(t-t_0)\big). \tag{3.69}$$

Observations:

- $h_{LP}(t)$ represents a **non-casual** system, therefore an ideal lowpass filter is not realizable.
- As $B \to \infty$, $h_{LP}(t) \to \delta(t-t_0)$, such that $y(t) \approx x(t-t_0)$. If the input is a single pulse, then the output will be identical to the input delayed by t_0. (While we are discussing the

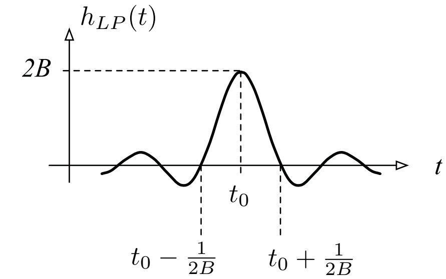


Figure 3.30: Impulse response function of an ideal lowpass filter.

case where the input is a single pulse, this observation will be correct with any arbitrary input.) However, for a signal with $B < \infty$, the input pulse will be spread-out as a result of transmission through a lowpass filter with finite B.

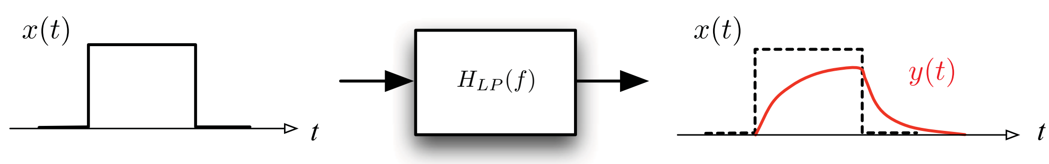


Figure 3.31: Response of a lowpass filter to a rectangular pulse.

- We can truncate $h_{LP}(t)$ such that $h_{LP}(t) = 0$, for $t < 0$. This operation will convert the non-causal ideal lowpass filter into a casual and therefore realizable approximation.
- In particular, for system realizability we require
  - In time domain: $h(t) = 0$, for $t < 0$—causality;
  - In frequency domain: $\int |\log H(\omega)|/(1+\omega^2)d\omega < \infty$—Paley-Wiener condition.

  Observe that any $H(f) = 0$ over a finite frequency band violates the Paley-Wiener condition.

In realizable filters it is not possible to have the passband and stopband ranges adjacent to each other (this would imply a $\Pi(f)$ type of a frequency response which would result in a corresponding impulse response function with infinite support, and hence will result in a non-causal system). Also, rather than having $|H(f)| = K$ for $f \in$ *passband* and $|H(f)| = 0$ for $f \in$ *stopband*, we specify maximum and minimum attenuation levels specified over these frequency bands. In the context of frequency-selective filters, we will refer to:

- passband(s), transition band(s), stopband(s);
- maximum deviation/tolerance over passband(s);
- minimum attenuation/tolerance over stopband(s).

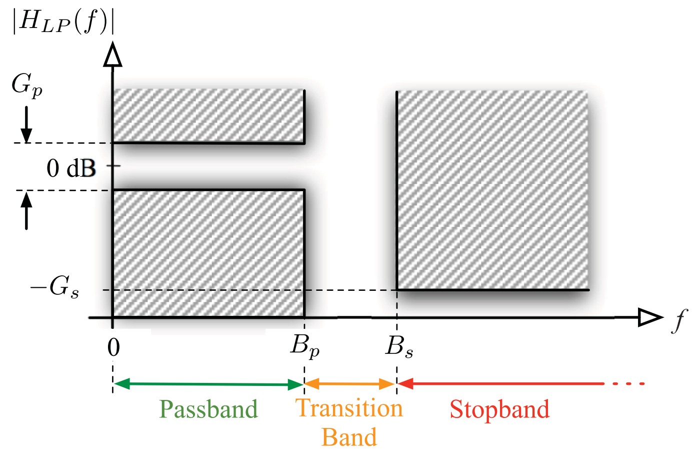


Figure 3.32: Specifications of a realizable lowpass filter.

Frequency-selective filters are mostly specified in terms of their magnitude response functions. Once the filter is designed to satisfy the magnitude response constraints, we study the corresponding phase response to ensure that the phase distortion (caused by non-linear phase response) remains within acceptable limits. As an example consider a typical, realizable lowpass filter with the magnitude response constraints specified as shown. The magnitude response function should remain outside of any cross-hatched regions shown in Figure (3.32). Figure (3.32) also shows that the following parameters describe a realizable lowpass filter:

Passband: The frequency range $[0, B_p]$ Hz. (We provide information only for $f > 0$ as magnitude response function of the filter will have even symmetry about $f = 0$).

Passband Tolerance, G_p: This parameter is typically measured in dB and represents the maximum allowable deviation from ideal *distortionless transmission* conditions, i.e. $|H(f) = 1$ or 0 dB. Ideally, we want G_p to be as close to 1 (or 0-dB) as possible. This parameter is also know as the **ripple factor**.

Transition Band: The frequency range $[B_p, B_s]$ Hz.

Stopband: The frequency range $[B_s, \infty)$ Hz.

Stopband Attenuation, G_s: This parameter represents the minimum attenuation over the stopband; it is typpically specified in dB. Iddeally, we want G_s to be as close to 0 (or if specified in dB, as large as possible). Please note that G_s measures attenuation, therefore it is a positive quantity. For example, a $G_s = 40$ dB implies that the lowpass filter will exert a minimum of 40 dB attenuation (measured relative to the passband gain) over the stopband.

In general, the filter order increases when [G_p is small] or [G_s is large] or the transition band, $[B_p, B_s]$, is narrow or the stopband edge frequency B_p is small. The most commonly utilized frequency-selective filter structures are lowpass, highpass, bandpass, bandstop and allpass filters. For example, a typical bandpass filter may have the following magnitude response specifications as shown in Figure (3.33). There are many analog filter design methods/families. Some of the

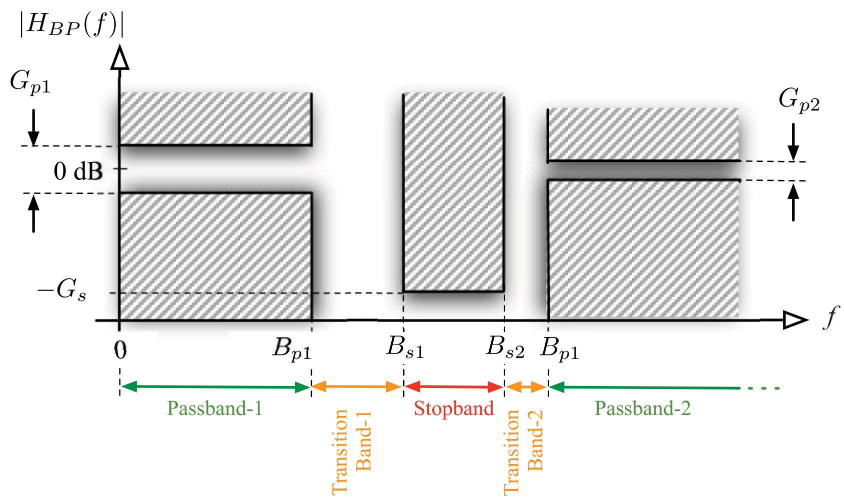


Figure 3.33: Specifications of a realizable bandstop filter.

most commonly used ones methods include: Butterworth, Chebchev-I, Chebchev-II, Elliptic and Bessel-Thompson among others. Each of these filter design methodss result in polynomials that define the pole-zero locations of the filter in the s-plane. When you study a text on analog filter design techniques, you will notice that all design techniques refer to lowpass filters only. So, how can we design the other types of frequency-selective filters?

Discussion:

- A comparison of the frequency responses for lowpass and highpass filters (to simplify the discussion assume that both systems have phase responses that are identical to zero over the entire frequency band) yield the relationship $H_{HP}(f) = 1 - H_{LP}(f)$.

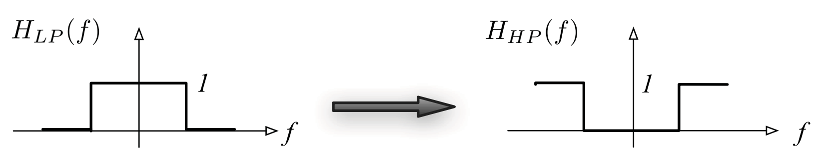


Figure 3.34: Lowpass to highpass transformation.

- Bandpass and bandstop filters are combinations of lowpass and highpass filters.
- While most common filter structures are frequency-selective, there are other “filters” as well: allpass filters (used as phase conditioners or to create minimum-phase systems required for system inevitability), phase-shift system/Hilbert transformer (used to create analytic signals), differentiator, etc..
- Digital filter design techniques open up many other possibilities. The design of frequency-selective *infinite impulse response* (IIR) digital filters rely mostly on existing analog filter design methodologies (e.g., Butterworth, Chebchev-I/II, ...). On the other hand, *finite impulse*

response (FIR) digital filters allow to achieve perfect linear phase (in addition to meeting the magnitude response specifications) and are particularly suitable for efficient implementation on high-speech digital signal processors.

3.9 Some Practical Considerations for Bandpass Systems

In a previous section we have analyzed an RC-circuit as a simple first-order lowpass filter. We now extend the discussion to an RLC-circuit and define the simplest bandpass system.

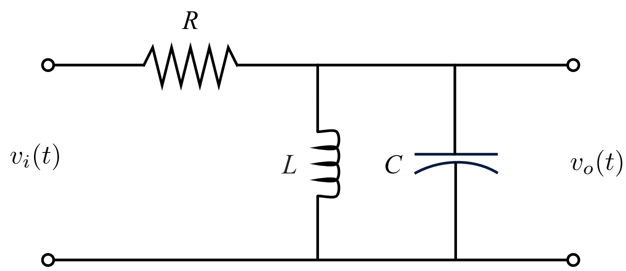


Figure 3.35: The RLC-Parallel resonant system.

Parallel Resonant/Tank/Tuned Circuit: The transfer function of the RLC-circuit equals:

$$H(s) = \frac{sL}{s^2(RLC) + sL + R}. \tag{3.70}$$

We can also express the frequency response function of the RLC circuit in an equivalent form as:

$$H(f) = \frac{1}{1 + j\,Q\,\big(\frac{f}{f_0} - \frac{f_0}{f}\big)}, \tag{3.71}$$

where

$$Q = R\sqrt{\frac{C}{L}}, \quad \text{(Quality Factor)}; \tag{3.72}$$

$$f_0 = \frac{1}{2\pi}\,\frac{1}{\sqrt{LC}}, \quad \text{(Resonant Frequency)}. \tag{3.73}$$

From the magnitude response curve shown in Figure (3.36) we observe that

$$B = f_U - f_L = \frac{f_0}{Q} \tag{3.74}$$

with

$$f_U = f_0\big(1 + \frac{1}{2Q}\big); \tag{3.75}$$

$$f_L = f_0\big(1 - \frac{1}{2Q}\big). \tag{3.76}$$

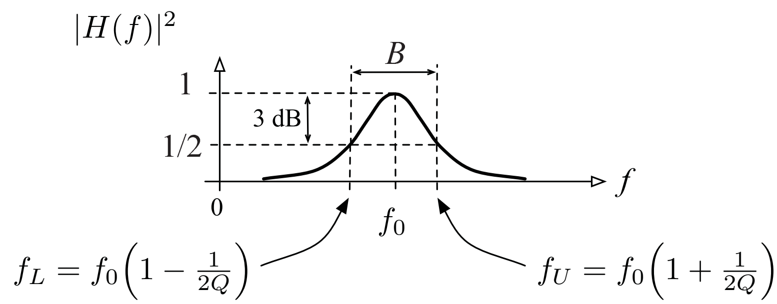


Figure 3.36: The magnitude response of the parallel resonant system.

- Observe that the quality factor parameter Q defines the bandwidth of the RLC-circuit. As Q increases, the system becomes more selective.
- Practical circuits usually have $Q \in [10, 100]$. In other words, the 3-dB bandwidth of a bandpass system falls within 1%–10% of its centre frequency f_0.
- The quantity B/f_0 is known as the the fractional bandwidth of a bandpass system. Designing a distortionless bandpass system can be very challenging if $B/f_0 \gg 1$ or $B/f_0 \ll 1$. For ease of implementation, a bandpass system should have $0.01 \ll B/f_0 \ll 0.1$. Thus, large

Frequency Band	**Carrier Frequency f_0**	**Bandwidth B**
Longwave radio	100 kHz	2 kHz
Shortwave radio	5 MHz	100 kHz
VHF	100 MHz	2 MHz
μWave	5 GHz	100 MHz
Optical	5×10^{14} Hz	10^{13} Hz

Table 3.1: Selected carrier frequencies and nominal bandwidth (based on $B/f_0 \approx 0.02$). [A.B. Carlson, *Communication Systems*, Third Edition, McGraw-Hill, Inc., page 195]

bandwidths to be used for signal transmission (more data payload with faster data rates) require higher frequencies.

- Earlier we defined lowpass/baseband/bandpass signals and systems. One additional terminology we will frequently use is that of a **narrowband** system.

  Definition 3.1. A bandpass system is a **narrowband system** if its bandwidth is small compared to its centre frequency.

- In communication problems, the information source outputs usually a baseband signal (e.g. a TTL waveform from a digital circuit, voice signal from a microphone, a video signal from a

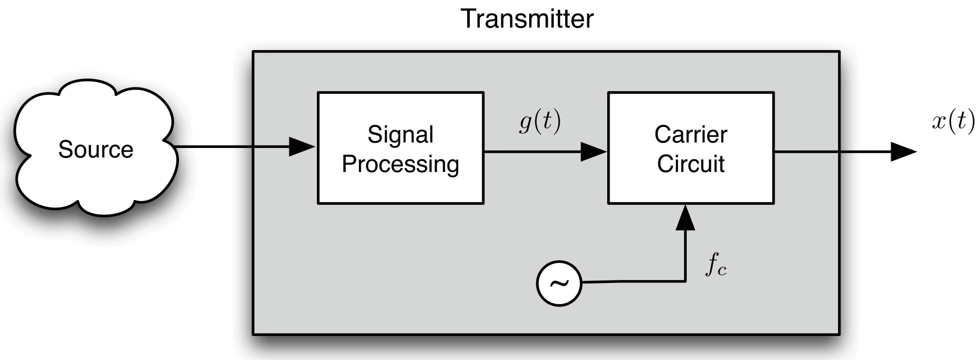


Figure 3.37: Preparing the baseband signal for transmission.

camera ...) In a communication system, it is the transmitter that modulates a high-frequency carrier signal with the baseband information signal, so that the fractional bandwidth of resulting modulated waveform will be compatible with the above discussed ranges.

3.10 Signal Distortion Over a Communication Channel

3.10.1 Linear Distortion

In its simplest form we define **linear distortion** as any deviation from the conditions of distortionless transmission we defined earlier.

Example 3.5: Let $x(t)$ be arbitrary, periodic signal with a Fourier series expansion $\{D_n\}_n$. If all frequency components in $x(t)$ as defined by the Fourier series coefficients are scaled and delayed by the same amount, then the system output corresponding to $x(t)$ will also be scaled and delayed. However, if the individual components are scaled and/or delayed differently (e.g. in the case of a non-uniform magnitude response or a non-linear phase response), then the output of the system will be distorted relative to the system input. This type of

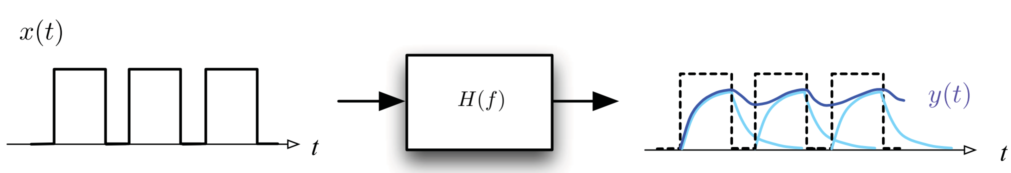


Figure 3.38: Effets of linear distortion on TDM signals.

distortion can create significant problems for time-division multiplexed (TDM) signals as signals occupying adjacent TDM channels will mutually interfere.

3.10.2 Non-Linear Distortion

Let $H(f)$ be the frequency response of a system with $x(t)$ and $y(t)$ be the system input and output signals, respectively. Let $y(t) = g[x(t)]$ with $g[.]$ being a non-linear function. If we expand $g[.]$ in a McLaurin series:

$$y(t) = g\big[\, x(t) \,\big] = a_0 + a_1 x(t) + a_2 x^2(t) + \cdots \; .$$

Observe that if $x(t)$ is band-limited to B Hz, then $x^n(t)$ is band-limited to nB Hz. Consequently, the spectrum of the system output is spread out well beyond the spectrum of the input signal. If spectrum-spreading distortion occurs as a result of non-linear distortion, it may likely create problems with frequency-division multiplexed systems.

Example 3.6: Let $x(t) = A\cos\omega_1 t + B\cos\omega_2 t$ with

$$y(t) = g[x(t)] = x(t) + 2x^2(t).$$

Substituting the expression for $x(t)$ into the system input-output equation (after some simple arithmetic operations and use of trigonometric identities) results in the following expression for the system output:

$$\begin{aligned}
y(t) &= \big[\, A\cos\omega_1 t + B\cos\omega_2 t \,\big] + 2\left[A\cos\omega_1 t + B\cos\omega_2 t\right]^2, \\
&= A\cos\omega_1 t + B\cos\omega_2 t + 2A^2\cos^2\omega_1 t + 2B^2\cos\omega_2 t + 4AB\cos\omega_1 t\cos\omega_2 t, \\
&= A\cos\omega_1 t + B\cos\omega_2 t + A^2 + A^2\cos 2\omega_1 t + B^2 + B^2\cos 2\omega_2 t \\
&\qquad + 2AB\cos(\omega_1+\omega_2)t + 2AB\cos(\omega_1-\omega_2)t.
\end{aligned}$$

If $f_1 = 4$ kHz and $f_2 = 3$ kHz, the input and output spectra will be as shown: Observe that

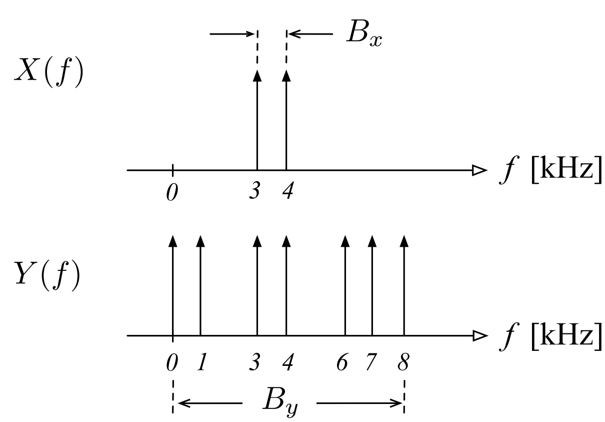


Figure 3.39: Spectra of $x(t)$ and $y(t) = g[x(t)]$. (Magnitudes of spectral components are not drawn to scale.)

$x(t)$ requires a transmission bandwidth of $B_x = 1$ kHz whereas $y(t)$ requires a transmission bandwidth of $B_y = 8$ kHz due to the spectrum-spreading effects of the non-linear system. Most importantly, we observe that the system output contains frequency components that are not part of the system input. This is an example of so-called **intermodulation distortion**

(IM distortion). This approach (feeding an input signal consisting of two closely spaced sinusoids and searching the system output for frequency components that are not part of the system input) is the foundation of many test systems designed to measure the IM distortion. As such, the IM distortion is an inherent measure of the system non-linearity.

3.10.3 Multipath Effects

Multipath transmission takes place when the transmitted signal arrives to the receiver by two or more paths with different delays. Multipath transmission may happen in cases when we transmit signal over a cable that has impedance mismatch with the transmitter and/or receiver structures. More frequently, we encounter multipath transmission in mobile communication systems where the receiving antenna receives not only a direct signal but also its reflections from nearby objects. The direct and reflected signals may interfere with other constructively or destructively.

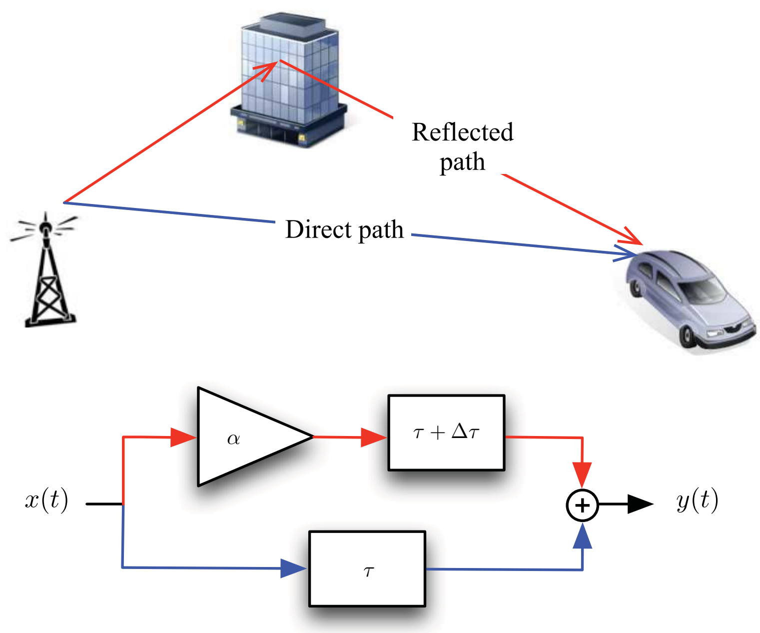


Figure 3.40: Modelling multipath distortion.

3.10.4 Fading Channels

Fading channels are non-stationary channels where the channel transmission characteristics vary randomly due to changing propagation conditions. In particular the channel may exert random attenuation of the signal, a phenomenon known as **fading**. Many communication systems employ **automatic gain control** subsystems to counter the effects of such fading channels.