Part 1: Introduction to Signals

Outline

I. Elementary Signal Concepts
- A. Signal definition and signal descriptions
- B. Elementary signal characteristics
  - 1. Signal support characteristics
  - 2. Signal value characteristics (Signal statistics)
  - 3. Signal shape characteristics
- C. Two-dimensional signals

II. Elementary Signal Operations
- A. Elementary operations on one signal
- B. Elementary operations on two or more signals
- C. Effects of operations on characteristics

III. Signal Similarity Measures
- A. Mean-squared error
- B. Signal correlation

IV. Introduction to Basic Signal Processing Tasks
- A. Signal recovery/noise reduction
- B. Signal detection/classification/recognition
- C. Signal digitization

V. Appendix A: Complex-Valued Signals

VI. Appendix B: Trigonometric Identities and Complex Exponentials

VII. Problems

I. Elementary Signal Concepts

Reading Assignment: Chapter 1 and these notes. It is recommended that you review these notes every now and then throughout the term. Some of these elementary concepts will be needed much later in the course, and some will be well understood only after you have had more experience with signals.

A. Signal Definition and Signal Descriptions

What are layperson examples of "signals" in common use?

Definition: A "signal" or "waveform" is a quantity that varies with time (or space), and typically conveys **information**.

In precise mathematical terms, a signal is a function of time. That is, for each value of time t there is number called the[1] **signal value at time** t.

Since signals are functions, and functions are constructs of mathematics, we will use the language of mathematics throughout the course, starting now!

Notation: We typically use lower case letters like x, y, and s, or subscripted letters like $x_1(t)$ to represent signals, *i.e.*, functions of time.

Usually we show time t as the argument of such function, as in $x(t)$.

Beware of the Ever-Present Notational Ambiguity: When you see "$x(t)$" written, sometimes the writer intends you to think of the value of the signal at the specific time t, as in $x(3.1)$, and sometimes $x(t)$ means the whole signal—that is, the writer intends you to think about the whole signal, *i.e.*, the signal values at all times. When it is essential that readers think about the whole signal, authors will sometimes write x or $\{x(t)\}$ or $x(\cdot)$ instead of $x(t)$.

Continuous-Time and Discrete-Time Signals: If the time variable ranges over a continuum of values, we say that the signal is **continuous-time**. If the time variable ranges over a discrete set of values we say the signal is **discrete-time.**

More specifically, unless stated otherwise, we assume that the time values of every continuous-time signal range over the set of all real numbers from $-\infty$ to $+\infty$. In mathematical terms we say that the **domain** of a function $x(\cdot)$ is the interval $(-\infty, \infty)$, which we denote by the symbol $\mathbb{R}$ as a shorthand.

Similarly, unless stated otherwise, we assume that the time values of every discrete-time signal range over the set of all *integers*: $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$, which we denote by the symbol $\mathbb{Z}$ as a shorthand. That is, the domain of the signal (function) is the set of all integers. When dealing with discrete-time signals it is most common to use one of the symbols $i, j, k, l, m,$ or n to denote time rather than t. It is also common to put the time variable inside square brackets "[ ]", rather than in ordinary parentheses. For instance, the following are examples of the notation used for discrete-time signals: $x[n]$, $y[k]$, $z_1[m]$.

Signal Descriptions: Sometimes signals are described with formulas and sometimes they cannot be so described.

Example. Continuous-time signals described with formulas:

$$x(t) = \mathrm{e}^{-t}, \qquad y(t) = 3\sin(47t), \qquad z(t) = \begin{cases} 2, & t < 0 \\ t^2, & 0 \le t \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

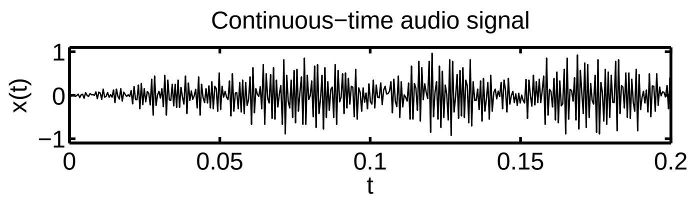
In EE, we often define functions **piecewise** using braces.

Example. The "touch tone 2" signal:

$$x(t) = \cos(2\pi 1336 t) + \cos(2\pi 697 t).$$

[1] **This font** is used when a technical term is used or introduced for the first time.

Example. A continuous-time signal that is not describable with a formula:

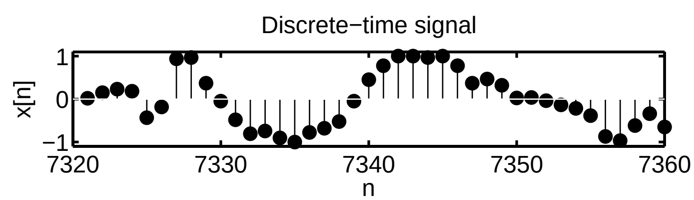


The signal shown above is part of a recording of a train whistle. Almost everything that one would hear is embodied in the function plotted above. The only missing information is the "volume" (notice the lack of units for $x(t)$), which in practice would depend on the settings of the amplifiers and the efficiencies of the speakers, etc. But the *information* in the signal is independent of the volume; the sound would be recognizable (by a human with unimpaired hearing) as a train whistle for any reasonable amplifier settings. (But can your eyes tell it is a train whistle from the plot?)

Example. Discrete-time signals described with formulas:

$$x[n] = \left(\frac{1}{2}\right)^n, \qquad y[n] = 3\sin(47n), \qquad z[n] = \begin{cases} 2, & n < 0 \\ n^2, & 0 \le n \le 10 \\ 0, & \text{otherwise.} \end{cases}$$

Example. A discrete-time signal that is not (easily) describable with a formula:



Are signals described by formulas more "real" or "authentic" than signals that are not so describable?
What does it mean to "describe a signal with a formula?"

Over the centuries, it has been found useful to give names to certain basic mathematical operations, such as '+', '-', '×', '/', x, $\ln(x)$, e, $|x|$, etc., and certain basic functions, such as $\sin(x)$, $\cos(x)$, $\Gamma(x)$, etc. To "describe a signal with a formula" is simply to say that it can be expressed in terms of previously defined operations and formulas. A signal that is not describable by a formula may simply be a function waiting to be blessed with its own name. Or it may be a function that has not previously occurred and may never occur again. Generally, we do not consider signals described by formulas to be any more real or authentic than those that are not so describable.

Note that a formula describing a signal can be quite complex, as in

$$s(t) = \sum_{k=1}^{N} a_k \cos(2\pi f_k t + \phi_k)$$

where $N, a_1, \ldots, a_N, f_1, \ldots, f_N, \phi_1, \ldots, \phi_N$ are "signal parameters", *i.e.*, constants or variables that one needs to know to fully determine the signal. It will be important to develop the skill of being able to work with complex signal formulas. For example, when you see the summation sign $\sum$, you should recognize that it is just an abbreviation for a sum of N terms. Indeed, to better understand the signal described by a summation, it is often useful to write it in its unabbreviated form, *e.g.*,

$$s(t) = a_1 \cos(2\pi f_1 t + \phi_1) + a_2 \cos(2\pi f_2 t + \phi_2) + \cdots + a_N \cos(2\pi f_N t + \phi_N).$$

Discrete-Time Signals from Continuous-Time Signals via Sampling

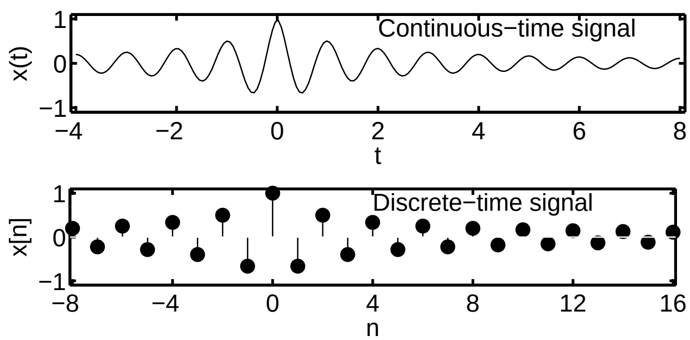
Frequently discrete-time signals are produced by **sampling** a continuous-time signal. If $x(t)$ is a continuous-time signal and T_{s} is a positive number then the discrete-time signal produced by sampling $x(t)$ with **sampling interval** T_{s} is the signal $x[n]$ defined by

$$x[n] = x(nT_{\mathrm{s}})\,.$$

Example. If $x(t) = \frac{\cos(2\pi t)}{1+|t|}$ and $T_{\mathrm{s}} = 1/2$ then

$$x[n] = x(nT_{\mathrm{s}}) = x(n(1/2)) = \frac{\cos(2\pi n(1/2))}{1+|n/2|} = \frac{\cos(\pi n)}{1+|n/2|} = \begin{cases} \frac{1}{1+|n/2|}, & n \text{ even} \\ \frac{-1}{1+|n/2|}, & n \text{ odd} \end{cases} = \frac{(-1)^n}{1+|n/2|}.$$

The continuous-time signal $x(t)$ and the discrete-time signal $x[n]$ produced by sampling $x(t)$ are shown below.



Example. If $T_{\mathrm{s}} = 1.5$ and $x(t) = \begin{cases} \sqrt{2t-4}, & t \geq 2 \\ 0, & \text{otherwise,} \end{cases}$ then

$$x[0] = x(0) = 0, \qquad x[1] = x(1.5) = 0, \qquad x[2] = x(3) = \sqrt{2}, \qquad x[3] = x(4.5) = \sqrt{5}, \qquad x[4] = x(6) = \sqrt{8}, \text{ etc.}$$

To express this compactly, consider the following manipulations:

$$x[n] = x(nT_{\mathrm{s}}) = \begin{cases} \sqrt{2(nT_{\mathrm{s}})-4}, & (nT_{\mathrm{s}}) \geq 2 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} \sqrt{3n-4}, & 1.5n \geq 2 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} \sqrt{3n-4}, & n \geq 2 \\ 0, & \text{otherwise.} \end{cases}$$

Notice that we replace *all* of the t's by nT_{s}, since that is what $x(nT_{\mathrm{s}})$ means. Also we wrote $n \geq 2$ rather than $n \geq 4/3$ in the final expression since n must be an integer.

The quantity $f_{\mathrm{s}} = 1/T_{\mathrm{s}}$ is called the **sampling frequency** or **sampling rate**, because it represents the frequency or rate (in samples per second) at which samples are taken.

As will be discussed a great deal later in the course, we often work with continuous-time signals by working with their samples, *i.e.*, with a discrete-time signal produced by sampling. For example, we often display continuous-time signals (approximately) by displaying their samples.

However, there are also discrete-time signals that are *not* obtained by sampling any continuous-time signal.
Example. Consider the signal $x[n]$, where $x[n]$ denotes the height of the nth person standing in a certain line.

B. Elementary Signal Characteristics

A starting point of any field of study is to *categorize* the objects under study. We examine **signal characteristics**[2] next.

We will emphasize the characteristics of continuous-time signals. There are discrete-time versions of each of these that will be presented later.

Perhaps the most important signal characteristic in this course is a signal's **spectrum**, which has to do with the "frequency content" of the signal. (Something like how a prism shows the components of white light.) This topic is so important that we will not discuss it here. Rather, beginning with Chapter 3, it will be a focus of much of the remainder of the class.

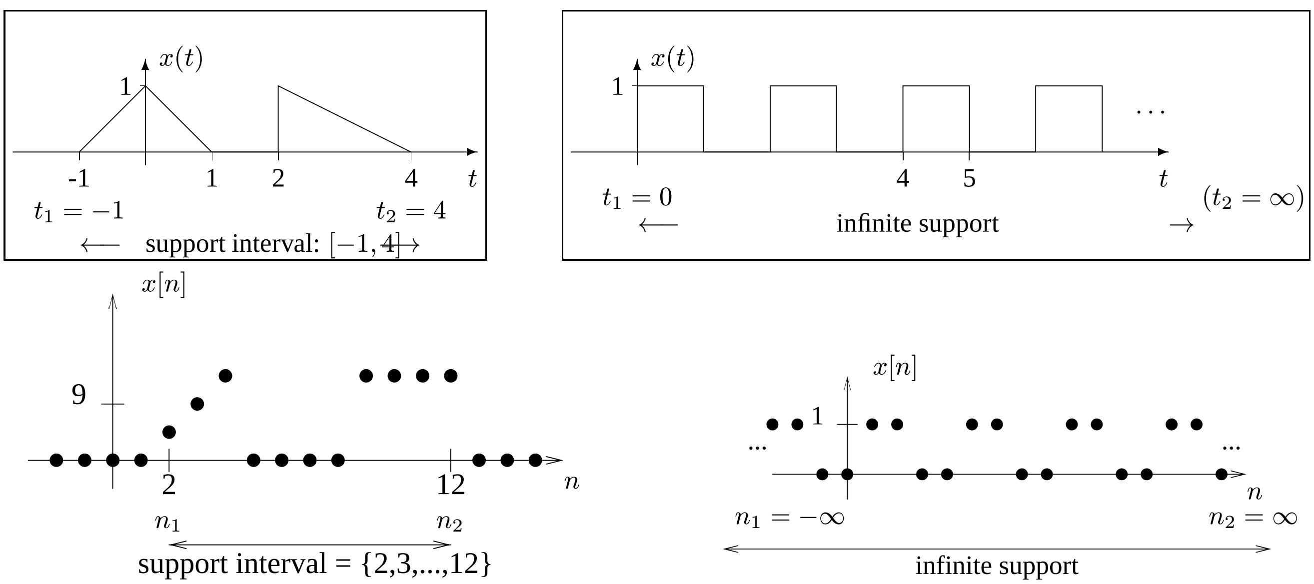
1. Signal Support Characteristics

These are signal characteristics related to the *time* axis.

C1. **Support Interval**:
Roughly speaking the **support interval** of a signal is the set of times such that the signal is not zero. We often abbreviate and say simply **support** or **interval** instead of support interval.

- More precisely the support interval of a continuous-time signal $x(t)$ is the smallest time interval[3] $[t_1, t_2]$ such that the signal is zero outside this interval.
- For a discrete-time signal $x[n]$, the support interval is a set of consecutive integers: $\{n_1, n_1+1, n_1+2, \ldots, n_2\}$. Specifically, n_1 is the largest integer such that $x[n] = 0$ for all $n < n_1$, and n_2 is the smallest integer such that $x[n] = 0$ for all $n > n_2$.

Example. Here are some signals and their supports.



C2. **Duration**:
The **duration** or **length** of a signal is the length of its support interval.

- For continuous-time signals, duration = $t_2 - t_1$.
- What is the duration of a discrete-time signal? duration = $n_2 - n_1 + 1$.

Some signals have finite duration and others have infinite duration.
Example. The two signals on the left above have finite duration, whereas the two signals on the right have infinite duration.

Outside of EECS 206, one may occasionally encounter situations where signals are considered to be **undefined** at times outside their support interval. However, within EECS 206, unless explicitly stated otherwise, we assume that signal values are 0 outside the support interval. Indeed, we will often define a signal simply by describing its values in some interval, with the presumption that the signal is zero for all times outside this interval.

[2]You do not need to memorize all of these. Rather you need to be aware of the existence of these characteristics, so you can look up and apply the appropriate ones at the appropriate times.

[3]Intervals can be **open** as in (a, b), **closed** as in $[a, b]$, or half-open, half-closed as in $(a, b]$ and $[a, b)$. For continuous-time signals, in almost all cases of practical interest, it is not necessary to distinguish the support interval as being of one type or the other.

Example. If we introduce a signal as

$$x(t) = t^2, \qquad 1 \le t \le 2,$$

then it should be understood that $x(t) = 0$ for $t < 0$ and for $t > 2$.

Pulses: Signals with short duration are often called **pulses.** Note that "short" is a subjective or relative designation.

Example. The square wave signal above can be considered to be a train of rectangular pulses.

Negative times and time zero: In some of the examples above the signal interval included negative times. What is the significance of negative time? To answer this, one must first answer the question: What is **time zero**? Basically, time zero is just some convenient reference time. Accordingly, a negative time simply represents a time prior to the reference time. For example, a radar system sends a pulse and waits to record the return times of reflections of this pulse from distant objects. It is usually convenient to let "time zero" be the time at which the original pulse was transmitted. Then $t = -1.8$ means 1.8 units of time before the reference time.

---

2. Signal Value Characteristics, a.k.a. Signal Statistics

We now consider characteristics that are related to the *values* that a signal $x(t)$ takes.

C3. **Maximum and minimum values**:
If $x(t)$ denotes some generic signal, then it has a **maximum value**

$$x_{\max} = \max_t x(t) \qquad \text{or } \max_n x[n]$$

and a **minimum value**

$$x_{\min} = \min_t x(t) \qquad \text{or } \min_n x[n]\,.$$

If these are both finite, *i.e.*, if $x_{\max} < \infty$ and $x_{\min} > -\infty$, then the signal is called **bounded**.

Example. The signal $x(t) = 3\mathrm{e}^{-|t|}$ has $x_{\min} = 0$ and $x_{\max} = 3$, so it is bounded.
Example. The signal $x[n] = \sqrt{|n + 4|}$ has $x_{\min} = 0$ and $x_{\max} = \infty$, so it is **unbounded**.

What do negative vs. positive signal values represent? The answer depends on the application. As an example, when a microphone responds to a sound, there is usually a diaphragm that moves back and forth, tracking the fluctuations in air pressure that constitute the sound. When the diaphragm is pushed one way, the microphone produces a positive voltage; when pulled the other way, it produces a negative voltage.

C4. **Absolute value**:
Quite often, when a signal has values that are both positive and negative, we are interested in a measure of the signal strength apart from its positive or negative sign. With signal strength in mind, one can compute its **magnitude** or **absolute value**, denoted $|x(t)|$.

C5. **Squared value, a.k.a. instantaneous power**:
In most physical situations, the square of $x(t)$, *i.e.*, $x^2(t)$, is a more useful measure of signal strength at time t than is its magnitude $|x(t)|$, because $x^2(t)$ is proportional to the instantaneous power in the signal $x(t)$ at time t, and because power is a physical quantity of fundamental importance. For such reasons, we often refer to $x^2(t)$ as the **instantaneous power** of $x(t)$ at time t. However, one must remember that the actual power is some constant times $x^2(t)$, where the constant depends on the specific physical situation. For example, if $x(t)$ represents the current in amperes flowing at time t through a resistor with resistance R ohms, then the instantaneous power absorbed by the resistor is $Rx^2(t)$ watts.

C6. **Energy**:
Although $x^2(t)$ is a useful measure of signal strength at a particular time instant t, often we need a measure of the total **energy** of a signal, or of a portion of the signal. The energy of a signal $x(t)$ in the interval $[t_1, t_2]$ is defined as follows:

$$E(x) = \int_{t_1}^{t_2} x^2(t)\, \mathrm{d}t.$$

Signals of infinite duration often have infinite energy (over their entire support). For such signals, average power (defined below) is usually a more relevant quantity than energy.

Example. Suppose the signal (expressed in volts) $v(t) = a\mathrm{e}^{-t/\tau},\ t \geq 0$ is applied across a 1 Ohm resistor. The energy dissipated in the resistor is

$$E(v) = \int_0^\infty v^2(t)\,\mathrm{d}t = \int_0^\infty \left[a\mathrm{e}^{-t/\tau}\right]^2 \mathrm{d}t = a^2 \int_0^\infty \mathrm{e}^{-2t/\tau}\,\mathrm{d}t = a^2 \frac{\tau}{2}.$$

(Compared to the expressions for mean value and MS value below, no "limit" is needed here since the expression $\int_0^\infty$ is already a shorthand for a limit.)

Notice that our final expression is a *formula*, not a number. This will be the case frequently in this course. From this formula, we see that increasing the amplitude a of the applied voltage will increase the dissipated energy by the square of a. And if we increase the decay constant τ, then the energy will also increase due to slower signal decay. These are among the types of relationships we are interested in exploring.

Example. The energy used in each pulse transmitted by a cellular phone will determine the battery life and can affect the voice quality.

C7. **Average or mean value**:
Given N values x_n, $n = 1, \ldots, N$, such as the ages of N people, the average of those values (*e.g.*, the average age) is simply

$$\frac{1}{N} \sum_{n=1}^{N} x_n.$$

Similarly, any signal also has an average value, also called a mean value. For discrete-time signals we often do not index from 1 to N but rather from some n_1 to some n_2, so our definition of the **average** or **mean value** of $x[n]$ is

$$\mathrm{M}(x) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x[n]\,.$$

For a continuous-time signal with support interval $[t_1, t_2]$, we can think of finely sampling the signal over that interval, computing the average of the samples using the above formula, and then taking the limit as the number of samples increases. This leads to the following integral form for the **average** or **mean value** of a continuous-time signal $x(t)$ over the interval $[t_1, t_2]$:

$$\mathrm{M}(x) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t)\ \mathrm{d}t.$$

In words: the mean value is the "area under the curve" divided by the signal duration.

If the interval over which the average is sought is infinite, then the average needs to be defined as a limit. For example, the average of the signal over the interval $[0, \infty)$ is:

$$\mathrm{M}(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\ \mathrm{d}t,$$

and the average over the interval $(-\infty, \infty)$ is

$$\mathrm{M}(x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^T x(t)\ \mathrm{d}t.$$

Again the "area divided by duration" concept applies over each interval of the form $[-T, T]$, and we take the limit as the interval length increases.

When evaluating these expressions, one must *first* compute the integral, then divide by T or $2T$, and *then* take the limit.

When a signal average is mentioned but an interval is not specified, we mean the average over the entire support of the signal.

Example. Consider the finite-length signal $x(t) = \begin{cases} 1 - |t|, & |t| \leq 1 \\ 2 - t/2, & 2 \leq t \leq 4 \end{cases}$ shown above.
The support is $[-1, 4]$ and the mean value is

$$\mathrm{M}(x) = \frac{1}{5} \int_{-1}^4 x(t)\ \mathrm{d}t = \frac{1}{5}\left[\int_{-1}^0 (1+t)\,\mathrm{d}t + \int_0^1 (1-t)\,\mathrm{d}t + \int_2^4 (2 - t/2)\,\mathrm{d}t\right] = \frac{1}{5}\,[1/2 + 1/2 + 1] = \frac{2}{5}.$$

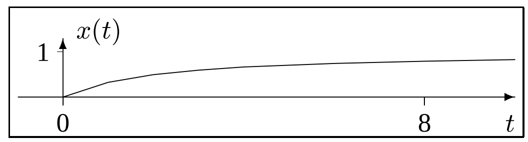
Example. Consider the finite-length signal $x(t) = \begin{cases} 3t^2, & 1 \le t \le 2 \\ 5-t, & 4 \le t \le 5. \end{cases}$ The support is $[1, 5]$ and the mean value is

$$\mathrm{M}(x) = \frac{1}{4}\int_1^5 x(t)\,\mathrm{d}t = \frac{1}{4}\left[\int_1^2 3t^2\,\mathrm{d}t + \int_4^5 5-t\,\mathrm{d}t\right] = \frac{1}{4}\,[7+1/2] = \frac{15}{8}.$$

Example. Consider the (eternal) sinusoid: $x(t) = \cos(2\pi f t)$, with $f \neq 0$. The support is $\mathbb{R}$ and the mean value is

$$\begin{aligned}\mathrm{M}(x) &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)\,\mathrm{d}t = \lim_{T\to\infty} \frac{1}{2T}\,\frac{1}{2\pi f}\sin(2\pi f t)\Big|_{-T}^{T} \\ &= \lim_{T\to\infty}\frac{1}{2T}\,\frac{\sin(2\pi f T) - \sin(2\pi f(-T))}{2\pi f} = \lim_{T\to\infty}\frac{1}{2T}\,\frac{\sin(2\pi f T)}{\pi f} = 0.\end{aligned}$$

The average value of an eternal sinusoid is zero, which is apparent from its plot.



Example. Consider the signal $x(t) = \begin{cases} t/(t+2), & t \ge 0 \\ 0, & \text{otherwise.} \end{cases}$

The support is $[0, \infty)$ and the mean value is

$$\begin{aligned}\mathrm{M}(x) &= \lim_{T\to\infty}\frac{1}{T}\left[\int_0^T \frac{t}{t+2}\,\mathrm{d}t\right] = \lim_{T\to\infty}\frac{1}{T}\left[\int_0^T 1 - \frac{2}{t+2}\,\mathrm{d}t\right] = \lim_{T\to\infty}\frac{1}{T}\left[t - 2\log(t+2)\Big|_0^T\right] \\ &= \lim_{T\to\infty}\frac{1}{T}\,[T - 2\log(T+2) + \log(2)] = \lim_{T\to\infty}\left[1 - \frac{2}{T}\log(T+2) + \frac{\log(2)}{T}\right] = 1.\end{aligned}$$

In electrical systems, $\mathrm{M}(x)$ is often called the **DC value**, where DC stands for **direct current**.

Example. Consider a **constant signal**, also called a **DC signal** (even if there is no "current"), such as $x(t) = c$. In this case $\mathrm{M}(x) = c$.

Example. Typically a microphone signal has average value equal to zero, or very nearly so, since acoustical signals usually oscillate nearly symmetrically about zero.

C8. **Mean-squared value, a.k.a. average power**:
Whereas $x^2(t)$ is an excellent measure of signal strength at an individual time instant t, quite frequently we need an aggregate measure of signal strength that applies to the whole signal, or to the signal over some specified time interval. In such cases, we will typically use the **mean-squared value** (MSV). Specifically, the MSV of a signal $x(t)$ over the interval $[t_1, t_2]$ is

$$\mathrm{MS}(x) = \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} x^2(t)\,\mathrm{d}t.$$

- This is also called the **average power** of $x(t)$ over the interval $[t_1, t_2]$.
- As with the definition of $\mathrm{M}(x)$, the definition of MS value needs to incorporate a limit when the support interval is infinite.
- When no interval is specified, the entire support interval is intended.

Example. Mean-squared value is useful when measuring the strength of the signal received by a radar antenna. If MS is large in an interval equal to the length of a radar pulse, then we assume that a reflected pulse has been received during this interval, and determine that this pulse is due to an object whose distance is the elapsed time since the original pulse was transmitted times the speed of light. If MS is very small, then we can assume that no reflected pulse has been received during this interval, *i.e.*, there is no object at the corresponding distance.

Example. Mean-squared value is used by electric utility companies to determine how much to charge you for the electricity they have supplied. This is because the amount of fuel required by them to supply your electricity is proportional to the mean-squared value of the current supplied to your home.

Example. Mean-squared value is often used as a signal quality measure. For example, suppose $x(t)$ is the signal coming from the leftmost of two microphones that are recording an orchestral concert, and suppose $y(t)$ is the signal fed to the left speaker of your stereo receiver after transmission by an FM radio station. Let $e(t) = x(t) - y(t)$ denote the difference between the two signals, which we consider to be an error signal, since ideally these two signals would be identical if the recording system, radio system, and speakers were perfect. Then the MSV of $e(t)$ is a good measure of the quality of the system that records and transmits $x(t)$ to you. In this context, it is usually called **mean-squared error.**

Example. If $x(t) = t/2$ for $0 < t < 2$, then $\mathrm{MS}(x) = \frac{1}{2}\int_0^2 (t/2)^2 \, \mathrm{d}t = 1/3$.

Energy vs average power

By comparing the previous definitions of energy and average power, we see that energy is the integral of instantaneous power. It is also the average power multiplied by the length of the interval. Alternatively, average power is energy divided by the length of the interval over which it is computed. A little thought will convince you that it is *energy* for which an electric utility company actually charges.

Since signal energy and average power (MSV) are related by a constant, the choice of which to focus on is often a matter of taste. If you focus on one, you can easily compute the other. The exception is for infinite duration signals with infinite energy; for such signals average power is generally more useful.

C9. **RMS Value**:
A closely related quantity is the **root mean-squared value** (RMSV), which is simply

$$\mathrm{RMS}(x) = \sqrt{\mathrm{MS}(x)} = \sqrt{\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x^2(t)\, \mathrm{d}t}.$$

On the one hand, RMSV is nicer than MSV in that its value is easier to interpret because its units are like a typical signal value, whereas the value of the MSV is harder to interpret because its units are like the square of a typical signal value. On the other hand, the MS value is slightly easier to work with, because it avoids the square root.

Example. Find the RMS value of

$$x[n] = \begin{cases} 3 + 1/2^n, & n \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

We first find the MS value:

$$\begin{aligned} \mathrm{MS}(x) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (3 + 1/2^n)^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} 3^2 + 6\frac{1}{2}^n + \frac{1}{4}^n \\ &= \lim_{N\to\infty} \frac{1}{N} \left[3^2 N + 6\frac{1 - 1/2^N}{1 - 1/2} + \frac{1 - 1/4^N}{1 - 1/4} \right] = \lim_{N\to\infty} \left[3^2 + 12\frac{1 - 1/2^N}{N} + \frac{4}{3}\frac{1 - 1/4^N}{N} \right] = 3^2. \end{aligned}$$

So $\mathrm{RMS}(x) = 3$, which is intuitive from the ***(Picture)*** of $x[n]$.

C10. **Variance and Standard Deviation**:
(These characteristics will be needed later in the course. You can skim them now, and return to them when needed.)

The mean-squared value of the difference between a signal and its average value is called the **variance** of x. The variance of a signal $x(t)$ over the interval $[t_1, t_2]$ is[4]:

$$\sigma^2(x) = \mathrm{MS}(x - \mathrm{M}(x)) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} (x(t) - \mathrm{M}(x))^2\, \mathrm{d}t.$$

The square root of the variance is called the **standard deviation**:

$$\sigma(x) = \mathrm{RMS}(x - \mathrm{M}(x)) = \sqrt{\sigma^2(x)} = \sqrt{\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} (x(t) - \mathrm{M}(x))^2\, \mathrm{d}t}.$$

The variance and standard deviation of a signal are useful measures of how "variable" is the signal. A signal with small variance or standard deviation stays close to its average value most of the time, whereas a signal with large variance or standard deviation does not. As with MSV vs. RMSV, standard deviation values are usually easier to interpret because their units are commensurate with signal units. On the other hand, variances are often easier to compute and work with.

Example. Output noise in an audio amplifier with no input signal.

Relationship Between Mean-Squared Value, Variance and Average Value:

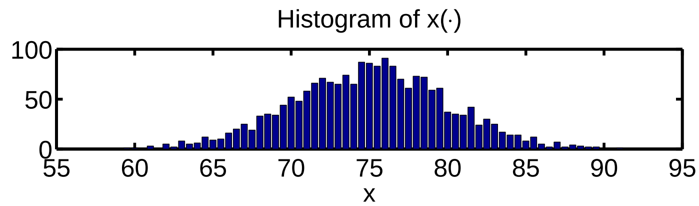
The following is a useful relationship.

$$\mathrm{MS}(x) = \sigma^2(x) + \mathrm{M}^2(x)\,.$$

Its derivation is left as a exercise.

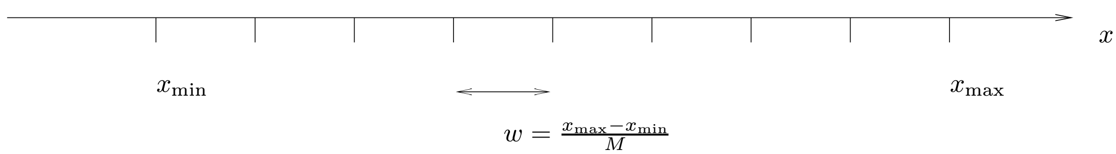
[4]Using the symbol σ^2 for variance and σ for standard deviation is traditional.

C11. **Signal Value Distribution and Histograms**:
The minimum, maximum, average, and mean-squared value are each numbers that each tell us something about the values that appear in the signal. The **signal value distribution** gives a more complete picture. Before introducing it, let us review the general meaning of the word **distribution**. As one example, consider the collection of grades resulting from an exam. If we speak of the "distribution of these grades," we mean a plot like that shown below. The horizontal axis shows the possible grades, and the height of the plot above a given grade is proportional to the number of exam papers with that grade. As another example, consider the distribution of incomes of residents of Michigan. Again this is a plot like the one shown below. In this case, the horizontal axis shows the possible incomes, and the height of the plot above a given income is proportional to the number of people with that income.



One may similarly consider the distribution of many, many quantities. Not surprisingly, in signals and systems, we are often interested in the distribution of values of a signal $x(t)$, which we call its **signal value distribution**. That is, for a given signal $x(t)$ we want a plot whose horizontal axis shows the signal values and whose height above a given signal value is proportional to the frequency with which that value[5] occurs in the signal.

How do we plot the signal value distribution of a signal $x(t)$? The most common way is make and plot a **histogram**. Specifically, we divide the range of signal values from $x_{\min}$ to $x_{\max}$ into M equal width **bins**, as illustrated below, where M is some integer, usually in the range 10 to 1000.

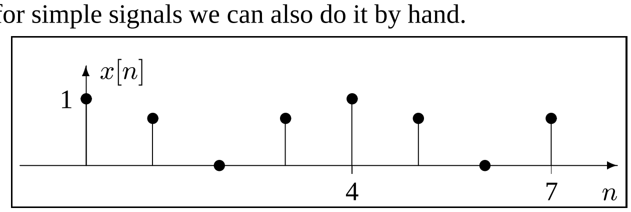


- If the signal is discrete-time, we count the number of signal values that lies within each bin:

$$N_m = \text{\# of signal values in } [x_{\min} + (m-1)w, x_{\min} + mw), \qquad m = 1, \ldots, M.$$

  We then make a "bar plot" showing each count N_m above the bin, as illustrated above.
- If the signal is continuous-time, then we apply this procedure using many *samples* of the signal. That is, we apply the procedure to the set of values $x(0)\,, x(T_s)\,, x(2T_s)\,, x(3T_s)\,, \ldots,$ where T_s is the sample spacing.

Example. We usually use a computer to compute histograms, but for simple signals we can also do it by hand.

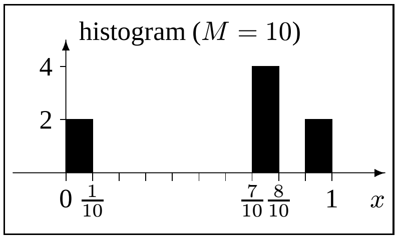


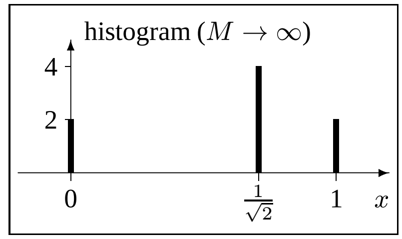
Consider the signal $x[n] = \begin{cases} \left|\cos\left(\frac{2\pi}{8}n\right)\right|, & n = 0, \ldots, 7 \\ 0, & \text{otherwise.} \end{cases}$

What values does $x[n]$ take? $\{0, 1/\sqrt{2} \approx 0.707, 1\}$

[5]Strictly speaking it is not the frequency of individual values that matter. Rather, for any value x, we want the frequency with which signal values lie in a small neighborhood of x, say from $x - \Delta$ to $x + \Delta$, Δ is a small constant.

By counting, the histogram of $x[n]$ for $M = 10$ is as follows.
If we let $M \to \infty$, then we get a histogram like that shown on the right. For signals that only take a small number of values, it is natural to consider the limit as $M \to \infty$.

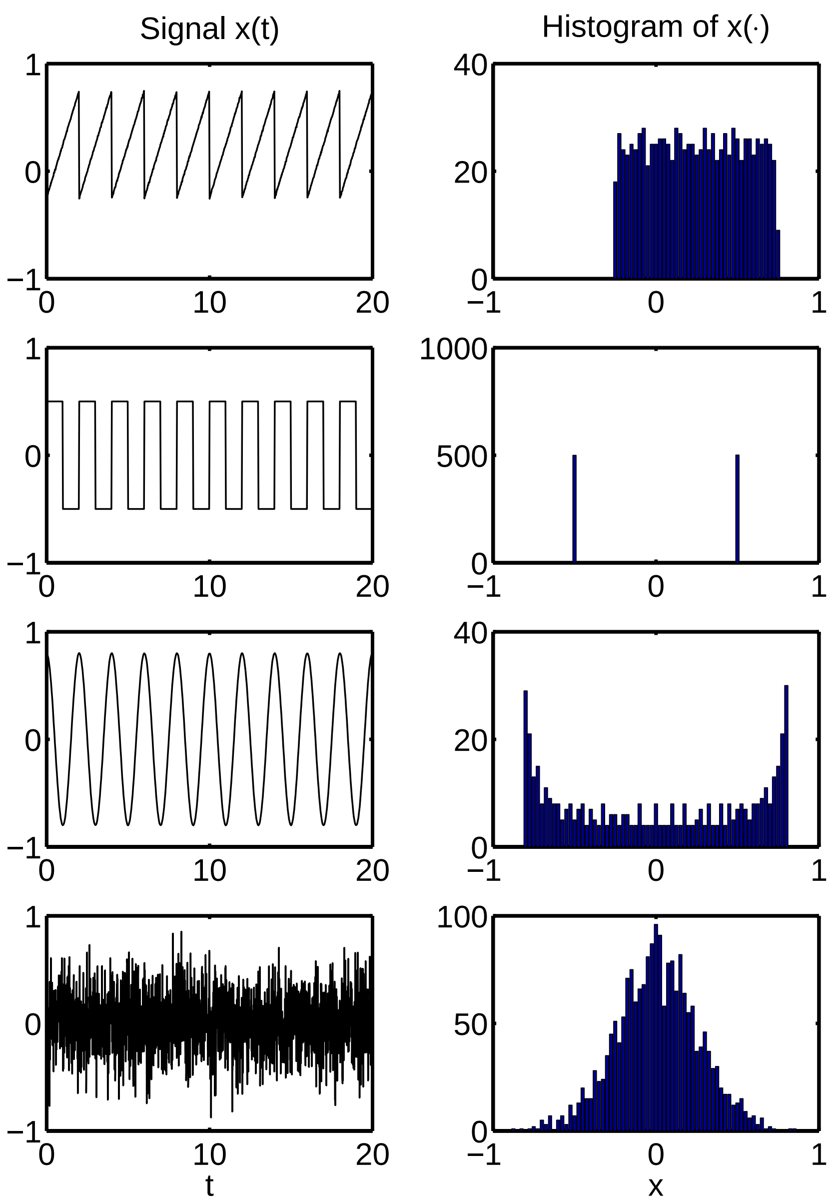




What if we considered 10 cycles of the signal ($n_2 = 79$)? (Scales vertical axis by 10.)
What if we considered 100 samples/cycle instead of just 8? Use MATLAB!

<u>Example</u>. Several signals and their signal value distributions are shown below.

These histograms were computed with MATLAB using the command `hist(x,M)`, where `x` is a vector containing signal samples, and `M` is the desired number of bins. M=100 usually works well.



Approximate signal characteristics from histograms

We now justify the statement made earlier that the signal value distribution gives a more complete picture of the signal values than its minimum, maximum, average and mean-squared values. We do this by showing that these latter quantities can be determined, at least approximately, from a histogram.

First, the minimum and maximum values will be readily apparent from the histogram. For example, the maximum value is approximately equal to the largest bin center for which the histogram is not zero.
Why approximate?

Next, let us show how to compute the average value $\mathrm{M}(x)$ from the histogram. Let $x[1]$, $x[2]$,..., $x[N]$ denote the signal samples. If the histogram has M bins, then the width of each bin will be $w = (x_{\max} - x_{\min})/M$. The first bin is the interval $[x_{\min}, x_{\min} + w)$, the second bin is the interval $[x_{\min} + w, x_{\min} + 2w)$, and so on, and the last bin is the interval $[x_{\min} + w(M-1), x_{\max}]$. Let c_m denote the center of the mth bin. That is,

$$c_m = x_{\min} + w(m - 1/2), \qquad m = 1, \ldots, M.$$

Let N_m denote the number of signal values that lie in the mth bin. Of course $\sum_{m=1}^{M} N_m = N$. Then the histogram is simply a bar plot of the points $(c_m, N_m), m = 1, \ldots, M$. The (exact) average value of the N signal samples is

$$\mathrm{M}(x) = \frac{1}{N} \sum_{n=1}^{N} x[n]\,.$$

Now we observe that we can approximately compute the sum in the above in a different manner. Since there are N_m signal values in the mth bin, we know that there are N_m signal values that approximately equal c_m. The sum of these values is approximately $N_m c_m$. Making this approximation for each of the bins leads to

$$\sum_{n=1}^{N} x[n] \approx N_1 c_1 + N_2 c_2 + \cdots + N_M c_M.$$

Thus, the following formula gives an approximation to the signal's average value:

$$\mathrm{M}(x) \approx \frac{1}{N} \sum_{m=1}^{M} N_m c_m = \sum_{m=1}^{M} \left(\frac{N_m}{N} \right) c_m.$$

That is, the average signal value $\mathrm{M}(x)$ is approximately the weighted average of the c_m's (the bin centers), where the weight multiplying c_m is the fraction of samples that lie in the mth bin.

Similarly, one may show that

$$\mathrm{MS}(x) \approx \sum_{m=1}^{M} \left(\frac{N_m}{N} \right) c_m^2$$

Then from the mean and the mean-squared value, one may directly compute the RMS value, the variance and the standard deviation.

The mean value, mean-squared value, RMS value, variance and standard deviation for a continuous-time signal are each approximately equal to the corresponding quantity for the discrete-time signal produced by sampling the continuous-time signal. Thus, they too may be estimated from a histogram.

As the number of bins M increases, the approximation improves, *cf.* Riemann approximations to integrals.

Example. From the $M = 10$ histogram of the signal $x[n]$ considered earlier, the approximate mean is

$$\mathrm{M}(x) \approx \frac{1}{8}[2 \cdot 0.05 + 4 \cdot 0.75 + 2 \cdot 0.95] = 0.6125.$$

For comparison, the exact mean is $\mathrm{M}(x) = \frac{1}{8}[2 \cdot 0 + 4 \cdot \frac{1}{\sqrt{2}} + 2 \cdot 1] = 0.6035$. The approximation error is smaller than the bin width.

In summary, for both discrete-time and continuous-time signals, all of the basic signal value characteristics can be determined, at least approximately, from the signal value distribution.

Summary of Signal Value Characteristics

The following table shows the definitions of the signal characteristics mentioned previously, with the exception of signal value distribution, which is not easily summarized in table form. It also lists the analogous characteristics for discrete-time signals.

These are all mathematically defined quantities, but each one is important due to some *physical considerations* in EE systems.

Characteristic	**Continuous-time signal** $x(t)$	**Discrete-time signal** $x[n]$
support interval	$[t_1, t_2]$	$\{n_1, n_1 + 1, \ldots, n_2\}$
duration	$t_2 - t_1$	$n_2 - n_1 + 1$
maximum value:	$x_{\max} = \max_t x(t)$	$x_{\max} = \max_n x[n]$
minimum value:	$x_{\min} = \min_t x(t)$	$x_{\min} = \min_n x[n]$
magnitude:	$\lvert x(t) \rvert$	$\lvert x[n] \rvert$
squared value (instantaneous power)	$x^2(t)$	$x^2[n]$
mean value:	$\mathrm{M}(x) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x(t) \, \mathrm{d}t$	$\mathrm{M}(x) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x[n]$
mean-squared value (average power)	$\mathrm{MS}(x) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x^2(t) \, \mathrm{d}t$	$\mathrm{MS}(x) = \frac{1}{n_2 - n_1 + 1} \sum_{n=n_1}^{n_2} x^2[n]$
RMS value:	$\mathrm{RMS}(x) = \sqrt{\mathrm{MS}(x)}$	
variance:	$\sigma^2 = \mathrm{MS}(x - \mathrm{M}(x))$	
standard deviation:	$\sigma = \sqrt{\mathrm{MS}(x - \mathrm{M}(x))}$	
relationship:	$\mathrm{MS}(x) = \sigma^2(x) + \mathrm{M}^2(x)$	
energy:	$\mathrm{E}(x) = \int_{t_1}^{t_2} x^2(t) \, \mathrm{d}t$	$\mathrm{E}(x) = \sum_{n=n_1}^{n_2} x^2[n]$

Computing continuous-time signal characteristics approximately using sums instead of integrals

This note supplements some formulas given in the Lab 1 background material without derivation.

Recall from the development of the **Riemann integral** in calculus that if N is large, then

$$\int_a^b f(t)\,\mathrm{d}t \approx \frac{b-a}{N} \sum_{k=1}^N f(t_k)\,,$$

where the interval $[a,b]$ is partitioned into N segments with left endpoints given by $t_k = a + \frac{k-1}{N}(b-a),\ k = 1,\ldots,N$.

We can use this approximation to compute signal characteristics like mean and energy of continuous-time signals by first **sampling** those signals and then using the discrete-time formulas for signal characteristics, *with a slight modification* in some cases.

Energy

Consider first the problem of calculating the **energy** of a continuous-time signal $x(t)$ from its samples $\{x[n]\}$ defined by

$$x[n] = x(nT_{\mathrm{s}})\,,$$

where T_{s} denotes the sampling rate.

Further assume that $t_1 = n_1 T_{\mathrm{s}}$ and $t_2 = (n_2+1)T_{\mathrm{s}}$ for some integers n_1 and n_2, so that $N = n_2 - n_1 + 1$ is the total number of samples, and thus $t_2 - t_1 = NT_{\mathrm{s}}$.

Applying the Riemann approximation:

$$\begin{aligned} \mathrm{E}(x) &= \int_{t_1}^{t_2} x^2(t)\ \mathrm{d}t \approx \frac{t_2-t_1}{N} \sum_{k=1}^N x^2\left(t_1 + \frac{k-1}{N}(t_2-t_1)\right) \\ &= T_{\mathrm{s}} \sum_{k=1}^N x^2(n_1 T_{\mathrm{s}} + (k-1)T_{\mathrm{s}}) = T_{\mathrm{s}} \sum_{n=n_1}^{n_2} x^2[n]\,. \end{aligned}$$

In summary, we can approximate the energy of a sampled signal as follows:

$$\boxed{\mathrm{E}(x) = \int_{t_1}^{t_2} x^2(t)\ \mathrm{d}t \approx T_{\mathrm{s}} \sum_{n=n_1}^{n_2} x^2[n] = \texttt{Ts * sum(x.\^{}2)}.}$$

Notice how the "extra" factor T_{s} comes out front due to the "width of the rectangle" in the Riemann approximation.

Mean value

Now consider instead the **mean** signal value. Applying the Riemann approximation:

$$\begin{aligned} \mathrm{M}(x) &= \frac{1}{t_2-t_1} \cdot \int_{t_1}^{t_2} x(t)\ \mathrm{d}t \approx \frac{1}{t_2-t_1} \cdot \frac{t_2-t_1}{N} \sum_{k=1}^N x\left(t_1 + \frac{k-1}{N}(t_2-t_1)\right) \\ &= \frac{1}{N} \sum_{k=1}^N x(n_1 T_{\mathrm{s}} + (k-1)T_{\mathrm{s}}) = \frac{1}{n_2-n_1+1} \sum_{n=n_1}^{n_2} x[n]\,. \end{aligned}$$

In summary

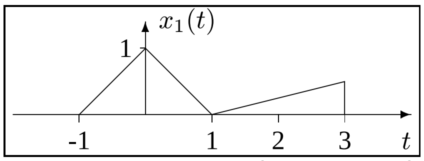
$$\boxed{\mathrm{M}(x) = \frac{1}{t_2-t_1} \int_{t_1}^{t_2} x(t)\ \mathrm{d}t \approx \frac{1}{n_2-n_1+1} \sum_{n=n_1}^{n_2} x[n] = \texttt{mean(x)}.}$$

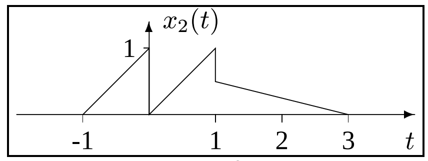
Notice how this time there is no "extra" factor `Ts` in the final MATLAB expression because it cancels out due to the **normalization** by the signal duration. Similarly, one can show that $\mathrm{MS}(x) \approx$ `mean(x.^2)`.

I.B.3. Signal Shape Characteristics

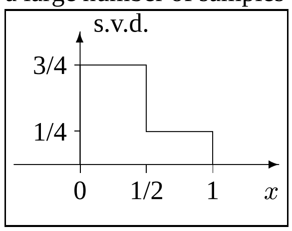
In this section we consider signal characteristics related to what we loosely call signal "shape." The signal value characteristics considered previously have nothing to do with signal shape, as one can see by noticing that very different signals can have the same signal value distribution, and consequently, the same min, max, average and mean-squared values. One may also observe that interchanging or time-reversing segments of a signal would have no effect on signal value characteristics, but definitely would affect signal shape.

Example. The following two signals have the same signal value distribution.





Here is the signal value distribution for the limit of a large number of samples and a large number of bins:

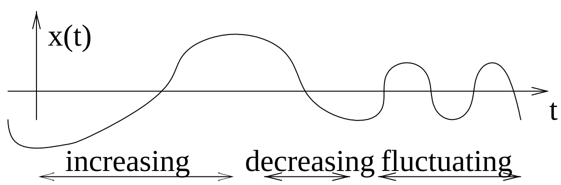


In discussing signal shape characteristics, we will first focus on continuous-time signals and later comment briefly on the analogous characteristics for discrete-time signals.

Local shape characteristics

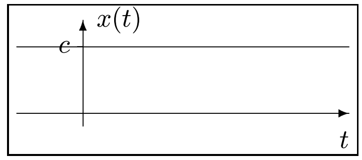
When studying a signal $x(t)$, we often examine segments of it to see if it is *increasing*, *decreasing* or *fluctuating*.

Example.

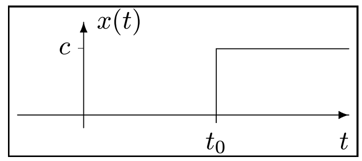


Common signal shapes

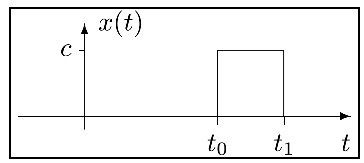
The following is a list of some common signal shapes. These shapes can occur by themselves, or as segments of signals. That is, they may be thought of as local characteristics. The symbols b, c, d, t_0 and t_1 represent parameters that one must specify for the signals to be completely determined.



- **constant**: $x(t) = c$

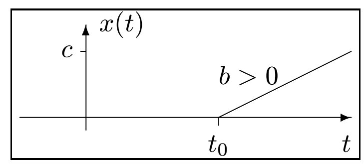


- **step**[6]: $x(t) = \begin{cases} c, & t \geq t_0 \\ 0, & t < t_0 \end{cases}$

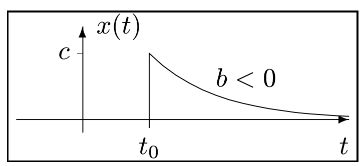


- **rectangular pulse**[7]: $x(t) = \begin{cases} 0, & t < t_0 \\ c, & t_0 \leq t \leq t_1 \\ 0, & t > t_1 \end{cases}$

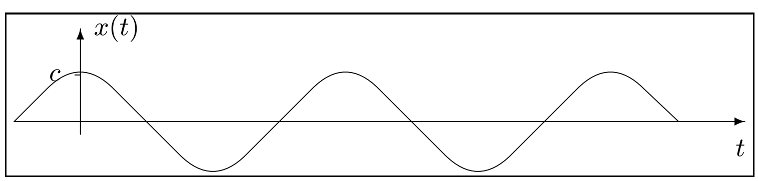
- **ramp**: $x(t) = \begin{cases} 0, & t < t_0 \\ c(t - t_0), & t \geq t_0, \end{cases}$ which is increasing if $c > 0$, decreasing if $c < 0$.



- **exponential**: $x(t) = \begin{cases} 0, & t < t_0 \\ c\mathrm{e}^{b(t-t_0)}, & t \geq t_0, \end{cases}$ which is increasing if $b > 0$, decreasing if $b < 0$, constant if $b = 0$.



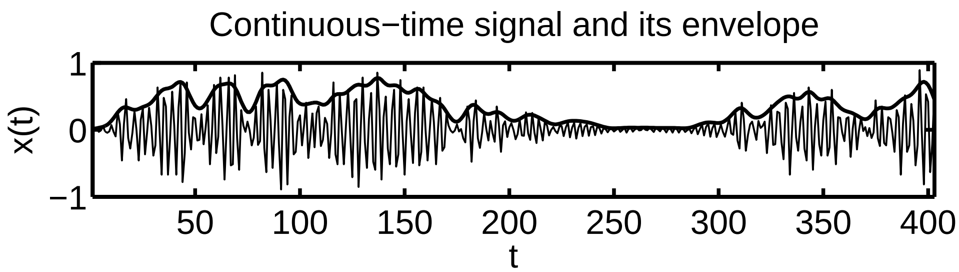
- **sinusoidal**: $x(t) = c \sin(bt + d)$, which is fluctuating if $b \neq 0$



[6] Since the value of x at time t_0 is c, strictly speaking, we should simply plot the value c at time t_0. Instead, we have drawn a vertical line from 0 up to c. This line emphasizes the change in x as it goes from $x(t) = 0$ for $t < t_0$ to $x(t) = c$ for $t > t_0$. This convention of drawing vertical lines where a function has a step change in value is quite common. In the real world, no signal can make an perfectly instantaneous step from one value to another, contrary to the formula for the step signal. Instead, a real world signal value would rise rapidly from 0 to c in the vicinity of t_0. Thus a plot of a real world step signal would have a nearly vertical line rising from 0 to c at t_0. We may think of the vertical line shown in the figure above as a reminder that, in the real world, the signal can change rapidly, but cannot actually have an ideal step change.

[7] Again notice the vertical lines, which are drawn for emphasis, and as a reminder of what happens in the real world.

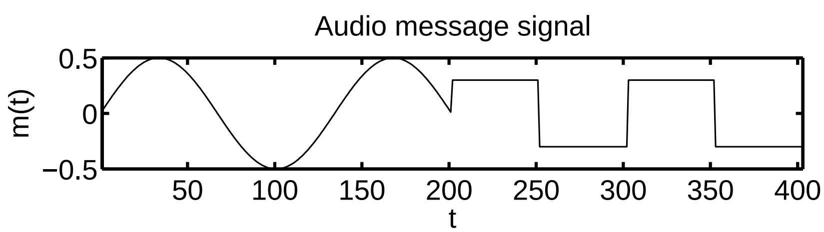
C12. **Signal Envelope**:
This is best introduced with an example. The thick black line overlaying the signal shown below is the *envelope* of the signal. That is, for a rapidly fluctuating signal $x(t)$, the envelope is a smooth curve that approximately follows the *positive* peaks of the signal. Admittedly this is not a very precise definition, and there is no universally accepted definition that can make it precise. Nevertheless, the envelope is often a useful concept.



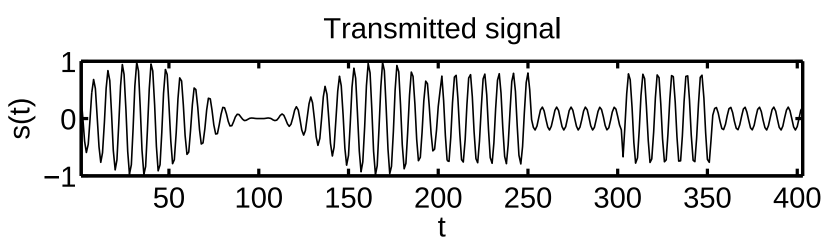
Example. An AM radio station transmits an audio signal by embedding it in the envelope of a high frequency signal. Specifically, suppose $m(t)$ is the audio signal to be transmitted. Then the radio station assigned to **carrier frequency** f_c transmits a signal of the form

$$s(t) = (m(t) + c)\cos(2\pi f_c t),$$

where c is a parameter chosen so that $m(t) + c \geq 0$ for all, or at least most, times t. Typically, f_c is a frequency much higher than the rate of fluctuation of $m(t)$. For example, if $m(t)$ is the audio signal shown below,



then the transmitted signal $s(t) = (m(t) + 0.5)\cos(2\pi f_c t)$ would be the following.

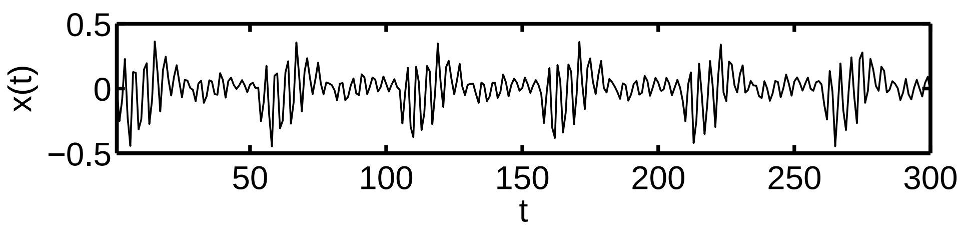


Can you see the audio signal $m(t)$ embedded in the envelope of the transmitted signal $s(t)$? Can you think of a way of recovering $m(t)$ from $s(t)$?

C13. **Periodicity**:
The signal shape characteristic known as periodicity is particularly important in signals and systems, because many signals that appear in nature are periodic, or at least nearly so, as are many human-made signals in electronic devices.
A periodic signal consists of a certain pattern that is repeated over and over, exactly the same each time.

Though many signals are **aperiodic**, *i.e.*, not periodic, it turns out that periodic signals still play a key role in their analysis.

Example. The following is a segment from a recording of someone speaking the vowel "ee."

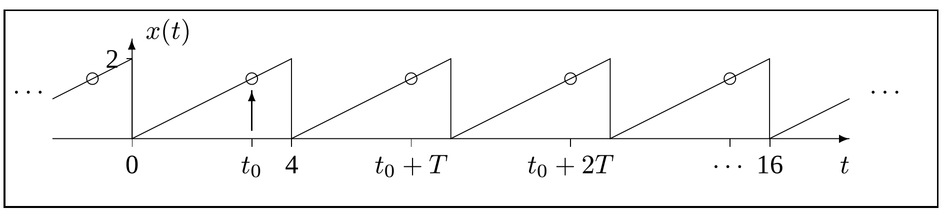


A continuous-time signal $x(t)$ is said to be **periodic** with **period** $T > 0$, or simply **T-periodic** if

$$x(t+T) = x(t) \text{ for all values of } t.$$

It is conventional to require the period T to be a positive number.

Example. The plot below shows a periodic signal called a **sawtooth wave**. Its values are marked at a particular time t_0 and also at times $t_0 + T$, $t_0 + 2T$,....



Example. Sawtooth waves occur in the scanning electronics for televisions. They are also an approximate model for the sound of a bowed violin string.

Example. Perhaps the most important periodic signals are **sinuosoidal** signals, which will be the focus of Chapter 2.
If $x(t) = \cos\left(\frac{2\pi}{T}t + \phi\right)$, then T is the (fundamental) period of $x(t)$. To see why T is a period of $x(t)$, notice that

$$x(t+T) = \cos\left(\frac{2\pi}{T}(t+T) + \phi\right) = \cos\left(\frac{2\pi}{T}t + 2\pi + \phi\right) = \cos\left(\frac{2\pi}{T}t + \phi\right) = x(t)\,,$$

because $\cos(\theta + 2\pi) = \cos(\theta)$.

Several important facts about periodic signals are given next.

Fact 1. A continuous-time signal $x(t)$ with period T is also periodic with period $2T$, because for *any* time t,

$$x(t+2T) = x((t+T)+T) = x(t+T) = x(t)\,,$$

where the last two inequalities follow from the definition of "periodic with period T."

Fact 1'. If $x(t)$ is T-periodic, then $x(t)$ is nT-periodic for all $n \in \mathbb{N} = \{1, 2, \ldots\}$. (The set $\mathbb{N}$ of positive integers is called the **natural numbers**.)

Fact 2. Though any periodic signal may be characterized as having infinitely many periods, there is always a unique smallest period, called the **fundamental period**, that is often denoted T_0. The fundamental period T_0 of a signal $x(t)$ is the smallest positive number T such that $x(t+T) = x(t)$ for every value of t. In other words, T_0 is the *smallest* period of $x(t)$.

The reciprocal of T_0 is called the **fundamental frequency** f_0 of the signal. That is, $f_0 = 1/T_0$. It is the number of fundamental periods that occur per unit time.

Warning! People often say "period" when they mean "fundamental period." (We will not be so careless in 206!) So when you hear the word "period," you need to use the context to figure out if they really mean "fundamental period."

Fact 3. If $x(t)$ has fundamental period T_0, then $x(t)$ is periodic with period nT_0 for every positive integer n.

Fact 3'. Conversely, these are the only periods of $x(t)$. That is, if $x(t)$ is T-periodic, then $T = nT_0$ for some $n \in \mathbb{N}$.

Derivation of the converse statement[8].
Suppose $x(t)$ is periodic with fundamental period T_0 and is also known to be periodic with period T. We must show that T is an integer multiple of T_0. We use proof by contradiction. Hypothetically suppose that T is not a multiple of T_0. Then $T = nT_0 + r$ where n is the integer part of T/T_0 and r is the remainder: $0 < r < T_0$. Since $x(t)$ is T_0-periodic, it must be that for any time t,

$$\begin{aligned} x(t+r) &= x((t+r)+NT_0) \qquad \text{since } x(t) \text{ is } T_0\text{-periodic and hence } NT_0\text{-periodic} \\ &= x(t+T) \qquad \text{because } T = NT_0 + r \\ &= x(t) \qquad \text{because } x(t) \text{ is periodic with period } T. \end{aligned}$$

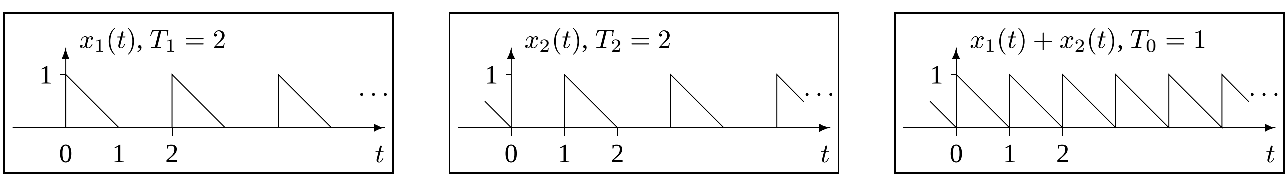
Since $x(t+r) = x(t)$, we deduce that $x(t)$ is r-periodic. But the fact that $r < T_0$ contradicts the assumption that T_0 is the fundamental period, which by definition is the smallest period of $x(t)$. Therefore, our hypothetical assumption must be false. We conclude that T must be a multiple of T_0.

Fact 4. A **constant signal**, *i.e.*, a **DC signal**, *e.g.*, $x(t) = 3$, is a special case. It satisfies $x(t+T) = x(t)$ for any choice of T. Thus it is T-periodic for every value of $T > 0$. However, it is conventional to define the fundamental period to be $T_0 = \infty$ and the fundamental frequency to be $f_0 = 0$. This somewhat arbitrary definition turns out to be more useful than other definitions.

Fact 5. If signals $x_1(t)$ and $x_2(t)$ are both periodic with period T, then the sum of these two signals, $x_1(t) + x_2(t)$ is also T-periodic.
Fact 5'. This same property holds when one sums three or more signals. (The derivation of this fact is left as an exercise.)

Fact 6. The sum of two signals with fundamental period T_0 is T_0-periodic, and *usually* T_0 is also the fundamental period of the sum. But sometimes the fundamental period of the sum can be *less* than T_0, as the following example illustrates.

Example. Below, $x_1(t)$ and $x_2(t)$ are both 2-periodic, yet $x_1(t) + x_2(t)$ is 1-periodic.



Fact 7. The sum of two signals with differing fundamental periods, T_1 and T_2, might or might not be periodic. The sum will be periodic if and only if the ratio of their fundamental periods is **rational**, *i.e.*, equals the ratio of two integers.

Example. If $T_2/T_1 = 5/3$ then the sum will be periodic. However, if $T_2/T_1 = \sqrt{2}$, then the sum will not be periodic.

To see how a rational ratio ensures periodicity, consider two signals: $x_1(t)$ with fundamental period T_1, and $x_2(t)$ with fundamental period T_2. Suppose that $T_2/T_1 = m/n$, where m and n are integers. Then $nT_2 = mT_1$. Letting $T = nT_2 = mT_1$, we see that

$$x_1(t+T) + x_2(t+T) = x_1(t+mT_1) + x_2(t+nT_2) = x_1(t) + x_2(t)$$

because $x_1(t)$ has period T_1 and $x_2(t)$ has period T_2. This shows that $x_1(t) + x_2(t)$ is T-periodic.

To complete our discussion, we should also show that if T_2/T_1 is **irrational** (not the ratio of integers), then $x_1(t) + x_2(t)$ is not periodic. However, the proof of this is beyond the scope of the course and will not be given here.

When T_2/T_1 is the ratio of two integers, the fundamental period of the sum signal is *usually* the **least common multiple** (**LCM**) of T_2 and T_1. (And this rule always works if the two signals are sinusoids of different frequencies!) To find the LCM of T_2 and T_1, we find the smallest integers m and n such that $nT_2 = mT_1$.

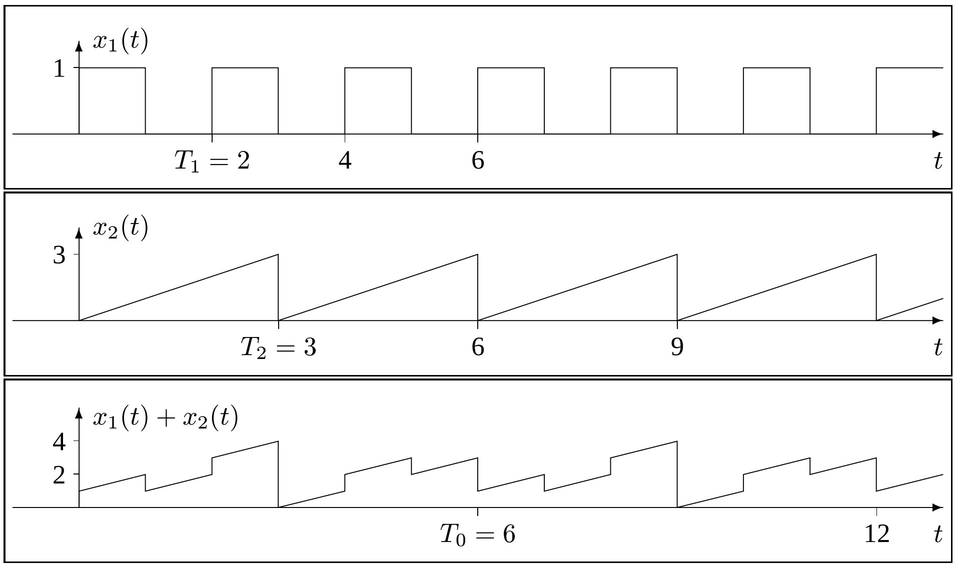
[8] This derivation is included for completeness. It is not expected that students can replicate this proof.

We say "usually" because there are also examples like that illustrated in Fact 6 where the actual fundamental period is smaller than the least common multiple[9].

Correspondingly, the fundamental frequency is *usually* the greatest common divisor of the fundamental frequencies f_2 and f_1 of the two signals.

There is one exception to the above discussion. If one of the two periodic signals is in fact simply a constant (a DC signal), then the sum will certainly be periodic and the fundamental period of the sum will equal the fundamental period of the other signal.

Example. Suppose $x_1(t)$ and $x_2(t)$ are the periodic signals shown below with fundamental periods 2 and 3, respectively. Then, their sum $x_1(t) + x_2(t)$ is periodic with fundamental period 6 = LCM(2,3).



If instead of summing two periodic signals, we sum several periodic signals, then a discussion similar to the one above shows that the sum is periodic if and only if the ratios of *each pair* of fundamental periods is rational. Moreover, the fundamental period of the sum is usually the least common multiple of all of the fundamental periods of the individual periodic signals.

Example. Consider three sinusoidal signals (*e.g.*, corresponding to the carrier signal of three radio stations): $x_1(t) = \cos(2\pi 2000 t)$, $x_2(t) = \cos(2\pi 3000 t)$, $x_3(t) = \cos(2\pi 4000 t)$. Determine the fundamental frequency of $z(t) = x_1(t) + x_2(t) + x_3(t)$. The fundamental periods are $T_1 = 1/2000$, $T_2 = 1/3000$, $T_3 = 1/4000$. The ratios of these periods are all rational, so $z(t)$ is indeed periodic. Now observe that $2T_1 = \frac{1}{1000}$, $3T_2 = \frac{1}{1000}$, $4T_3 = \frac{1}{1000}$, and the integers $2, 3, 4$ have no common divisors. Thus LCM$(T_1, T_2, T_3) = 1/1000 = T_0$. Thus the fundamental frequency of $z(t)$ is $f_0 = 1/T_0 = 1000$.

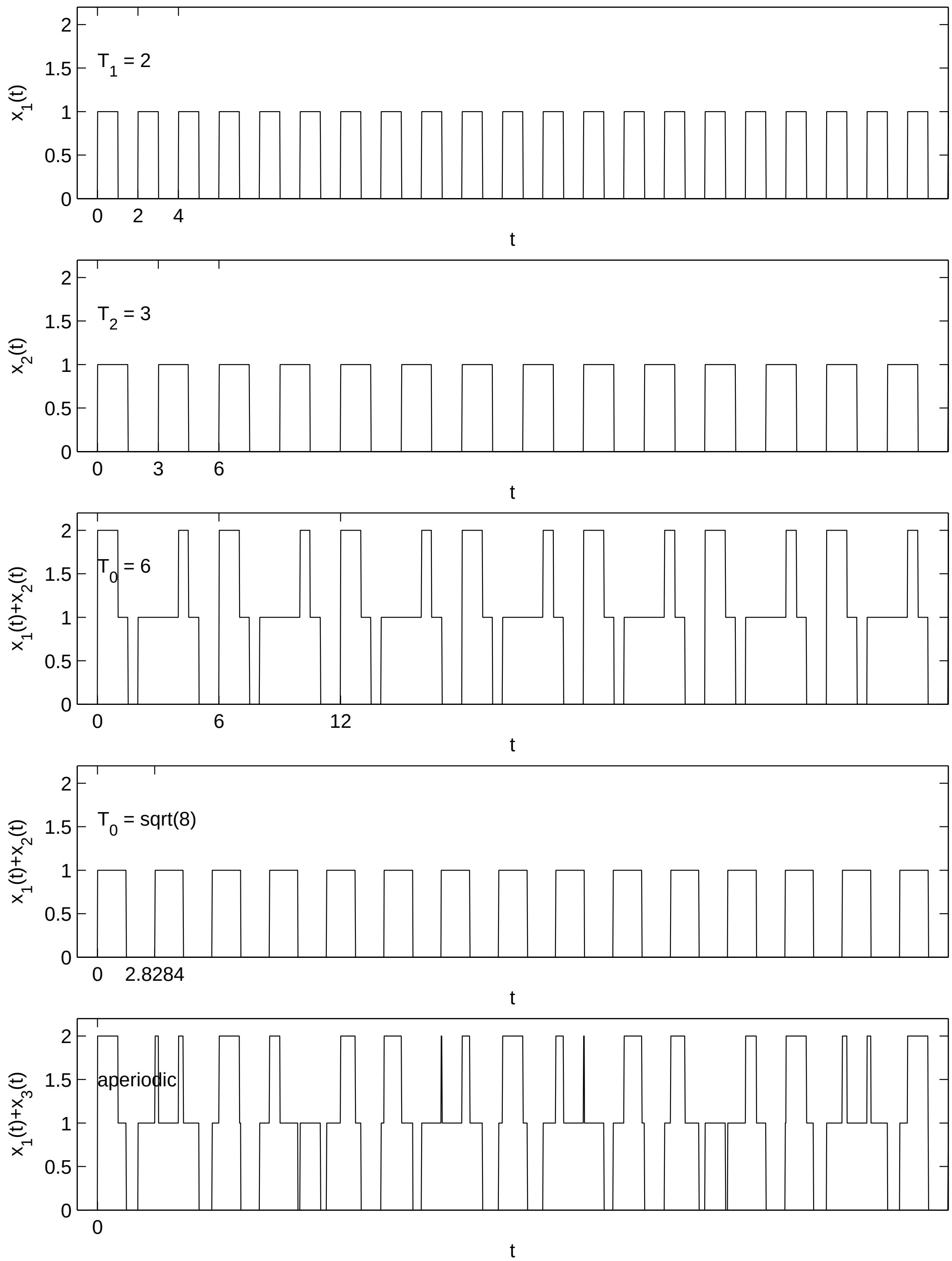
Summary of Facts 5-7

If $x_1(t)$ is T_1-periodic $x_2(t)$ is T_2-periodic, and $z(t) = x_1(t) + x_2(t)$, then
- Case 0: if $T_1 = \infty$, then $z(t)$ is T_2-periodic.
- Case 1: if T_2/T_1 is irrational, then $z(t)$ is aperiodic.
- Case 2: if $T_2/T_1 = m/n$ (*i.e.*, is rational), then $z(t)$ is T-periodic where $T = nT_2 = mT_1$.
  If T_1 and T_2 are the fundamental periods, respectively, then LCM(T_1,T_2) is usually the fundamental period of $z(t)$.

[9]There is one small exception to the "less than the LCM" rule. If the sum of two or more signals happens to be a constant, then the period of the result is ∞ regardless of what the individual periods are.
Example. $x_1(t) = 1 - \cos(t)$, and $x_2(t) = \cos(t)$.

Example. The following figure shows what happens when square waves of various periods are added.

- $x_1(t)$ is 2-periodic, $x_2(t)$ is 3-periodic, and $x_3(t)$ is $\sqrt{8}$-periodic
- $x_1(t) + x_2(t)$ is 6-periodic
- $x_1(t) + x_3(t)$ is aperiodic



Do periodic signals have finite support or infinite support? Always infinite! But the following fact spares us from using limits when computing signal statistics like mean value and average power. All of the information about the signal is contained in a single period, so we can compute all signal statistics from a single period!

Fact 8. The average of a periodic signal with period T over an interval whose length is a multiple of T equals the average over any interval of length T. The same applies to mean-squared value (equivalently, average power).

To see why, consider the average over the time interval $[t_1, t_1 + mT]$:

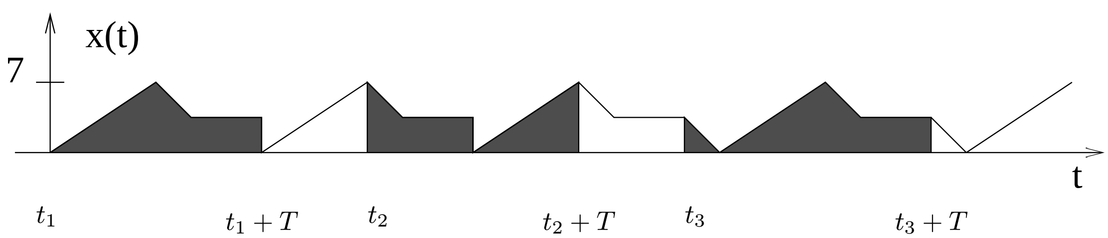
$$\begin{aligned}
\mathrm{M}(x) &= \frac{1}{mT} \int_{t_1}^{t_1+mT} x(t)\ \mathrm{d}t = \frac{1}{mT} \left[\int_{t_1}^{t_1+T} x(t)\ \mathrm{d}t + \int_{t_1+T}^{t_1+2T} x(t)\ \mathrm{d}t + \cdots + \int_{t_1+(m-1)T}^{t_1+mT} x(t)\ \mathrm{d}t \right] \\
&= \frac{1}{mT} \left[\int_{t_1}^{t_1+T} x(t)\ \mathrm{d}t + \int_{t_1}^{t_1+T} x(t)\ \mathrm{d}t + \cdots + \int_{t_1}^{t_1+T} x(t)\ \mathrm{d}t \right] \\
&\qquad \text{(because } x(t) \text{ is the same in each } t \text{ second interval)} \\
&= \frac{1}{T} \int_{t_1}^{t_1+T} x(t)\ \mathrm{d}t = \mathrm{M}(x)\,.
\end{aligned}$$

Thus we see that the average over m periods reduces to the average over just one period.

We can choose any t_1 we want. Usual choices are $t_1 = 0$ or $t_1 = -T/2$.

A related "limiting" argument shows that the average over an infinite interval of time reduces to the average over just one period.

Finally, we note that the average is the same over all intervals of length T. This follows from the fact, illustrated below, that the integral of $x(t)$ over any interval of length T is the same, because, by periodicity, the same values are being integrated, though perhaps in a different order.



Fact 8 also applies to mean-squared value, because mean-squared value is itself an average:

$$\mathrm{MS}(x) = \frac{1}{T} \int_0^T x^2(t)\ \mathrm{d}t = \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\ \mathrm{d}t.$$

(Any interval of length T will suffice.)

What about the energy $\mathrm{E}(x)$? For periodic signals, $\mathrm{E}(x) = \infty$ unless $x(t) = 0$.
So we always focus on average power rather than on energy for periodic signals.

Signal Shape Characteristics of Discrete-Time Signals

Discrete-time signals can have all the same shape characteristics as continuous-time signals. For example, they can be increasing, decreasing or fluctuating. Common signal shapes include all of those mentioned previously: constant, step, rectangular pulse, ramp, exponential and sinusoidal. Envelope is again a useful concept, as is periodicity. Because periodicity is such an important concept, we repeat the discussion of it here, this time for discrete-time signals.

Periodicity of discrete-time signals

A discrete-time signal $x[n]$ is called **periodic** with **period** N, or simply **N-periodic**, where $N \in \mathbb{N}$ (a natural number), iff

$$x[n+N] = x[n] \qquad \text{for all integers } n.$$

This definition is the same as the definition for continuous-time signals, except that instead of the equality holding for all continuous times t, it holds for all integer times $n \in \mathbb{Z}$. It is conventionally required that $N > 0$.

Example. The signal $x[n] = \cos(\pi n) = (-1)^n$ is 2-periodic.

We now revisit the various facts about periodicity.
Facts 1-6 are essentially identical as those for continuous-time signals, except that n and N replace t and T.
However, Facts 7 and 8 are different, and this difference is quite important!

Fact 1. A discrete-time signal with period N is also periodic with period mN for any positive integer m.

Fact 2. The **fundamental period**, denoted N_0, is the smallest positive integer N such that $x[n+N] = x[n]$ for all integers n. In other words, N_0 is the smallest period of $x[n]$.

The reciprocal of N_0 is called the **fundamental frequency** f_0 of the signal. That is, $f_0 = 1/N_0$. It is the number of fundamental periods occurring per sample. (It is always less than or equal to one.)

Warning! People often say "period" when they mean "fundamental period."

Fact 3. If $x[n]$ has fundamental period N_0, then $x[n]$ is periodic with period mN_0 for every positive integer m. Conversely, these are the only periods of $x[n]$. That is, if $x[n]$ is periodic with period N, then $N = mN_0$ for some integer m.

Fact 4. A constant signal, *e.g.*, $x[n]$ = 3, is a special case. It satisfies $x[n+N] = x[n]$ for any choice of $N \in \mathbb{Z}$. Thus $x[n]$ is periodic with period N for every value of $N > 0$. However, it is conventionally defined to have fundamental period $N_0 = \infty$ and fundamental frequency $f_0 = 0$. This somewhat arbitrary definition turns out to be more useful than other definitions.

Fact 5. If signals $x_1[n]$ and $x_2[n]$ are both periodic with period N, then the sum of these two signals, $z[n] = x_1[n] + x_2[n]$ is also periodic with period N. This same property holds when one sums three or more signals.

Fact 6. The sum of two signals with fundamental period N_0 is periodic with period N_0, but the fundamental period of the sum might be less than N_0.

Fact 7. The sum of two discrete-time signals $x_1[n]$ and $x_2[n]$ with (possibly) *differing* fundamental periods, N_1 and N_2, is always periodic!
This statement differs from the continuous-time case because the fundamental periods of discrete-time periodic signals are always rational. Therefore N_1/N_2 is also always rational, so the sum is always periodic.

The least common multiple of N_1 and N_2 is always a period of the sum, and *usually* equals the fundamental period of the sum. (The LCM is always the fundamental period if the signals are sinusoids.)

Similarly, the fundamental frequency usually equals the greatest common divisor of the fundamental frequencies f_1 and f_2.

Similarly, the sum of several discrete-time periodic signals is always periodic. Usually the fundamental period of the sum is the least common multiple of the fundamental periods of the individual signals.

Fact 8. The average of a periodic signal with period N over an interval whose length is a multiple of N equals the average over any interval of length N. The same applies to mean-squared value (equivalently, power).

$$\mathrm{M}(x) = \frac{1}{N} \sum_{n=n_1}^{n_1+N-1} x[n], \qquad \mathrm{MS}(x) = \frac{1}{N} \sum_{n=n_1}^{n_1+N-1} x^2[n].$$

(Usually we use $n_1 = 0$.) Again, a single period contains all the information needed to compute any signal statistics.

Example. To determine the average power of the signal $x[n] = \cos(\frac{2\pi}{6}n)$, we note by plotting that $x[n]$ cycles through the six values $\{1, 1/2, 0, -1/2, -1, -1/2, 1/2\}$ so $\mathrm{MS}(x) = \frac{1}{6}[1^2 + (1/2)^2 + 0^2 + (-1/2)^2 + (-1)^2 + (-1/2)^2 + (1/2)^2] = 1/2$.

Period of discrete-time sinusoidal signals

To find the fundamental period of a discrete-time sinusoidal signal $x[n] = \cos(\hat{\omega}n + \phi)$, one must work a bit harder than in the continuous-time case.

- First, express $\hat{\omega}$ in the form $\hat{\omega} = 2\pi\frac{M}{N}$ where N is a positive integer.
- If (and only if) M is also an integer, then $x[n]$ is periodic.
- In addition, after M and N have been simplified to have no common divisors, then N is the fundamental period of $x[n]$. (Unless $N = 1$, in which case $x[n]$ is a constant signal so the fundamental period is said to be infinity.)

To help understand these claims, observe that if $x[n] = \cos\left(2\pi\frac{M}{N}n + \phi\right)$, where M and N are integers, then

$$x[n+N] = \cos\left(2\pi\frac{M}{N}(n+N) + \phi\right) = \cos\left(2\pi\frac{M}{N}n + \phi + 2\pi M\right) = \cos\left(2\pi\frac{M}{N}n + \phi\right) = x[n],$$

because $\cos(\theta + M2\pi) = \cos(\theta)$. This confirms that N is a period of $x[n]$, but more work is needed to show that N is the fundamental period when M/N is simplified.

Since N is a period, we know that $N = LN_0$ where is L is a positive integer and N_0 is the fundamental period of $x[n]$. To show that N is in fact the fundamental period, we must show that $L = 1$ is the only choice. We do this using a "proof by contradiction." Suppose $L > 1$. Then

$$x[n+N_0] = \cos(2\pi\frac{M}{N}(n+N_0) + \phi) = \cos(2\pi\frac{M}{N}n + 2\pi\frac{MN_0}{N} + \phi) = \cos(2\pi\frac{M}{N}n + 2\pi\frac{M}{L} + \phi) \neq x[n]$$

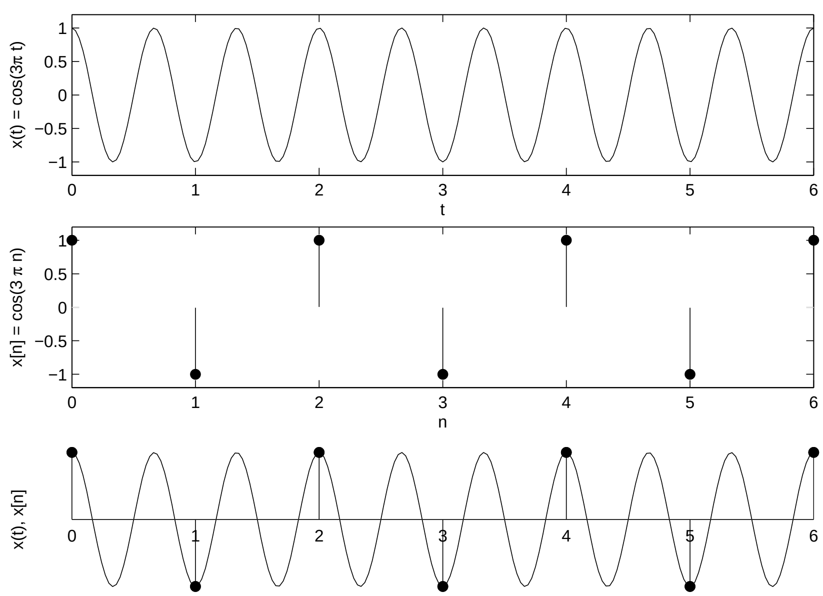
for some n because L is a divisor of N but N and M have no common divisors (except unity), so $2\pi\frac{M}{L}$ is not an integer multiple of 2π. So we must have $L = 1$ and hence N is the fundamental period when the ratio M/N is simplified to have no common factors in the numerator and denominator.

Example. Consider the signal $x[n] = \cos(\frac{7\pi}{9}n + 1)$. We write $\hat{\omega} = \frac{7\pi}{9} = 2\pi\frac{7}{18}$, so the fundamental period is 18.

Example. Consider the signal $x[n] = \cos(3\pi n)$. We write $3\pi = 2\pi\frac{3}{2}$, so $x[n]$ is periodic with fundamental period $N_0 = 2$.

In constrast, the continuous-time signal $x(t) = \cos(3\pi t) = \cos(2\pi\frac{3}{2}t)$ has fundamental frequency $f_0 = 3/2$ and hence fundamental period $T_0 = 2/3$. This is a very importance difference between continuous-time and discrete-time signals!

The following figure helps explain the situation. This phenomena is called **aliasing** and will be a main topic in Part 4.



I.C. Two-Dimensional Signals

A picture or **image**, as we will usually say, can also be modeled as a signal. However, in this case, it must be modeled as a *two-dimensional* signal $x(t, s)$. That is, instead of single independent parameter t representing time, there are two independent parameters t and s, representing vertical and horizontal position respectively. That is, $x(t, s)$ represents the intensity or brightness of the image at the position specified by horizontal position t and vertical position s, relative to some coordinate system. All of the previously mentioned concepts and characteristics can be extended to apply to two-dimensional signals. But we will not discuss them here.

Two-dimensional images can be "discrete-time" as well as "continuous-time" (**discrete-space** and **continuous-space** are better terms). In the discrete-space case, the signal (image) is denoted $x[m, n]$ where m and n are integers representing vertical and horizontal positions, respectively.

To display such a signal (image) graphically, we usually use a color scheme in which the value 0 is shown as a small black square, called a **pixel**, and the maximum signal value, often 255, is shown as a small white square, and intermediate values are shown in various shades of gray. The `colormap` command in MATLAB controls this behavior when displaying images using the `image` command. The `colorbar` command shows the mapping between signal values and grayscale intensities.

Example.

What is the dimensionality of television? Three dimensional: $s(x, y, t)$, two space dimensions plus time.

What about digital video? $s[n, m, k]$, where n, m denote spatial coordinates (pixel locations) and k denotes the time frame.

II. Elementary Signal Operations

Engineers design systems that manipulate signals in useful ways, such as combining multiple signals (*e.g.*, audio mixer) or modifying a signal, *e.g.*, an audio amplifier. In courses like EECS 215, you learn how to build such systems from physical components. In 206/306, you learn to design, analyze, and compare such systems independently of the physical devices.

Perhaps the most important such operation that will be discussed in this course is **filtering**, an operation so powerful that we will devote the last half of the course to it. Here we start with basic operations.

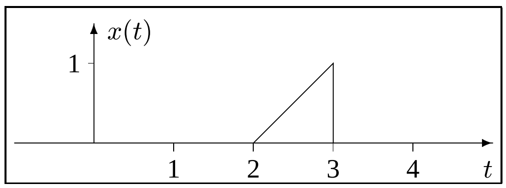
A. Elementary Operations on One Signal.

When discussing signals and systems, we routinely use a number of elementary **operations** that, when applied to one signal, result in another closely related signal. In the following we introduce these operations using continuous-time notation. With two exceptions, to be noted, they apply equally well to discrete-time signals.

Each of these operations is defined mathematically, but each also corresponds to one or more physical situations.

It is important to be able to apply these operations both graphically *and* mathematically, so we illustrate all of them using the following signal.

$$x(t) = \begin{cases} t-2, & 2 < t < 3 \\ 0, & \text{otherwise} \end{cases}$$



We deliberately chose a signal defined piecewise since such signals are often of interest.

Operations that modify signal values

O1. **Adding a constant**:
This is the operation of adding a constant to the signal. More specifically, there is a number c that is added to the signal value at every time t. If the original signal is $x(t)$, then the result is a new signal

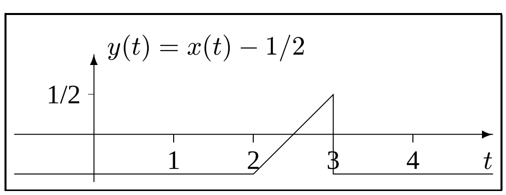
$$y(t) = x(t) + c.$$

It is easy to see that this operation has the effect of increasing the average value of x by c. That is, $\mathrm{M}(y) = \mathrm{M}(x) + c$.

Example. Connecting a waveform generator in series with a 9V battery would yield an overall voltage of $x(t) + 9$.

Example.

$$y(t) = x(t) - 1/2 = \begin{cases} t-2.5, & 2 < t < 3 \\ -1/2, & \text{otherwise} \end{cases}$$



O2. **Amplitude scaling**:
Amplitude scaling is the operation of multiplying a signal by a constant. That is, there is a constant c, called a **scale factor** or **gain**, and the value of the signal at every time t is multiplied by c. If the signal being scaled is $x(t)$, then the result of the scaling is

$$y(t) = cx(t)\,.$$

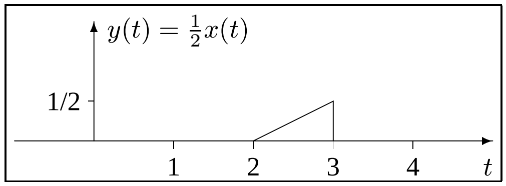
This has the effect of scaling both the average and the mean-squared values. Specifically,

$$\mathrm{M}(y) = c\mathrm{M}(x)\,, \qquad \mathrm{MS}(y) = c^2\mathrm{MS}(x).$$

Example. An ideal audio amplifier will scale the input signal (*e.g.*, from a phonograph) to a larger values suitable for driving speakers.

Example.

$$y(t) = \frac{1}{2}x(t) = \begin{cases} \frac{1}{2}(t-2), & 2 < t < 3 \\ \frac{1}{2}(0), & \text{otherwise} \end{cases} = \begin{cases} t/2 - 1, & 2 < t < 3 \\ 0, & \text{otherwise.} \end{cases}$$

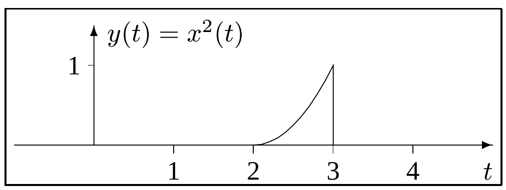


O3. **Squaring**:
Here we simply square the value of the signal at each time, yielding

$$y(t) = x^2(t)\,.$$

Example.

$$y(t) = x^2(t) = \begin{cases} (t-2)^2, & 2 < t < 3 \\ (0)^2, & \text{otherwise} \end{cases} = \begin{cases} (t-2)^2, & 2 < t < 3 \\ 0, & \text{otherwise.} \end{cases}$$



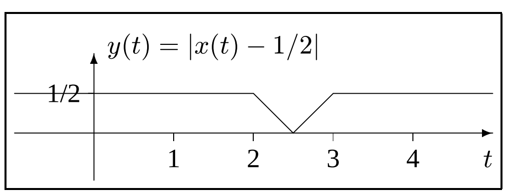
O4. **Absolute value**:
As the name suggests,

$$y(t) = |x(t)|\,.$$

What is a practical situation where we need an absolute value operation? AC to DC rectification.

Example. (combining absolute value and amplitude shift)

$$y(t) = |x(t) - 1/2| = \begin{cases} |t - 2.5|, & 2 < t < 3 \\ |-1/2|\,, & \text{otherwise} \end{cases} = \begin{cases} |t - 2.5|, & 2 < t < 3 \\ 1/2, & \text{otherwise} \end{cases}$$



Operations that modify the time axis

O5. **Time shifting**:
If $x(t)$ is a signal and T is some number, then the following signal is a **time-shifted** version of $x(t)$:

$$y(t) = x(t - T)\,.$$

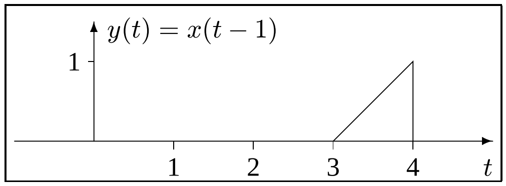
That is, the value of y at time t is precisely the value of x at time $t - T$. This means that if $T > 0$, then as illustrated below, anything that "happens" in the signal x also happens in the signal y, but it happens T time units later in y than in x. Similarly, if $T < 0$, it happens T time units earlier in y. It is useful to remember the rule that a positive value of T leads to a right shift of the plot of $x(t)$ and a negative value of T leads to a left shift.

This is our first exception where the discrete-time case is slightly different. Specifically, the shift T must be an integer for discrete-time signals, *e.g.*, $x[n - 5]$.

Example. Signal propagation times can be modeled by a time shift.

Example.

$$y(t) = x(t-1) = \begin{cases} (t-1)-2, & 2 < (t-1) < 3 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} t-3, & 3 < t < 4 \\ 0, & \text{otherwise} \end{cases}$$



O6. **Time reflection/reversal**:
The **time reflected** or **time reversed** version of a signal $x(t)$ is

$$y(t) = x(-t)\,.$$

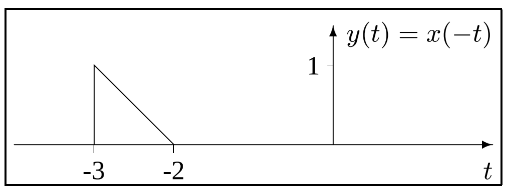
That is, whatever happens in x also happens in y but at the negative of the time it happens in y.

Example. Playing a recording backwards.

Example. A mirror. (2D case)

Example.

$$y(t) = x(-t) = \begin{cases} (-t)-2, & 2 < (-t) < 3 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} -t-2, & -3 < t < -2 \\ 0, & \text{otherwise.} \end{cases}$$



O7. **Time scaling**:
The operation of **time-scaling** a signal $x(t)$ produces a signal

$$y(t) = x(ct),$$

where c is some positive constant. If $c > 1$, this has the effect of "speeding up time" in the sense that the value of y at time t is the value of x at time ct, which is a later time. Alternatively, whatever happens in x in the time interval $[t_1, t_2]$, for $t_1 \geq 0$, now happens in y in the earlier and shorter time interval $[t_1/c, t_2/c]$.

This is the second property for which the discrete-time case includes an extra wrinkle. Specifically, in discrete-time, the time values must be integers. Therefore, if we take

$$y[n] = x[cn]\,,$$

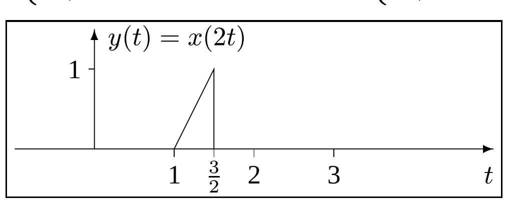
then c must be an integer.

Example. Playing back a recording faster or slower than the original recording speed.

Example. Doppler shift. (It is called a "shift" because motion of the source (*e.g.*, a train horn) causes a shift of the *frequency* of the signal, which in this context is nearly equivalent to time scaling.)

Example.

$$y(t) = x(2t) = \begin{cases} (2t)-2, & 2 < (2t) < 3 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 2t-2, & 1 < t < 3/2 \\ 0, & \text{otherwise.} \end{cases}$$



Combinations of the above operations

We will frequently encounter signals obtained by combining several of the operations introduced above. In particular, we will need combinations of time shift and time scale operations, possibly in conjunction with amplitude shift and/or amplitude scaling.

For example, given the signal $x(t)$ described in the previous examples, we may need to determine the signal $y(t) = 2x\left(3 - \frac{t}{2}\right)$. Again, it is important to be able to perform such operations both mathematically and graphically. The mathematical approach consists of applying the amplitude scaling and replacing each "t" in $x(t)$ with the appropriate argument, in this case $3 - \frac{t}{2}$, as follows:

$$y(t) = 2x(3 - t/2) = \begin{cases} 2\left(3 - \frac{t}{2} - 2\right), & 2 < 3 - \frac{t}{2} < 3 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 2\left(1 - \frac{t}{2}\right), & -1 < -t/2 < 0 \\ 0, & \text{otherwise} \end{cases} = \begin{cases} 2 - t, & 0 < t < 2 \\ 0, & \text{otherwise.} \end{cases}$$

For the graphical approach, there are two right ways and two wrong ways to do it. We will describe one correct way, which is all that is needed. Amplitude shift and amplitude scale are easy, but combinations of time shift and time scale must be done carefully. The following three-step recipe is a safe approach.

- First express the combined time shift and time scale in the following form: $y(t) = x(c(t - t_0))$.
- Time scale $x(t)$ by c.
  (Remember that if $c < 0$ then there is also a time reversal, and remember the difference between $c > 1$ and $c < 1$.)
- Time shift the resulting signal by t_0.
  (Remember that when $t_0 > 0$, the graph shifts to the right.)

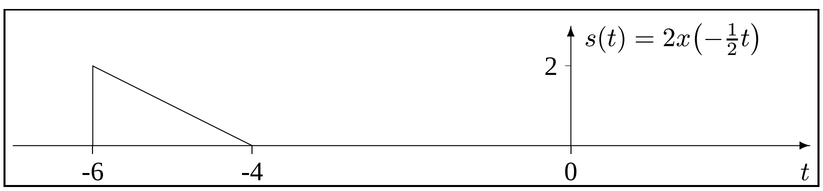
To help remember the order, note that "scale" comes before "shift" alphabetically.

To see that this approach is mathematically correct, consider first defining an intermediate signal $s(t) = x(ct)$, which is $x(t)$ time-scaled by c. Then we have $y(t) = s(t - t_0) = x(c(t - t_0))$, so $y(t)$ corresponds to $s(t)$ time-shifted by t_0.

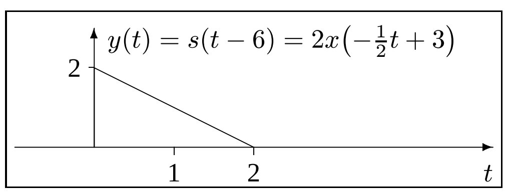
Example. Continuing to use the $x(t)$ described earlier, find $y(t) = 2x\left(3 - \frac{t}{2}\right)$.
First we write this in the recommended form:

$$y(t) = 2x\left(-\frac{1}{2}(t - 6)\right).$$

Next we time-scale $x(t)$ by $c = -1/2$ and we will also apply the amplitude scaling here to form $s(t) = 2x\left(-\frac{1}{2}t\right)$.



Finally we time-shift $s(t)$ by 6 (to the right in this case) to form $y(t)$.



Of course this final graph agrees with the mathematical formula.

If instead you used the form $x(at + b)$ then you could apply the time-shift first and then apply the time-scale second. But it is probably safer to just practice one way to do it and stick with that way.

Even with experience, trying to combine time shift and time scale operations into one step is quite prone to errors. Stick with the three-step approach!

B. Elementary Operations on Two or More Signals

O8. **Summing**:
As its name suggests, this is simply the operation of creating a new signal as the sum of two or more signals, as in

$$z(t) = x(t) + y(t) \,.$$

More specifically, the value of z at time each time t is the sum of x at time t and y at time t.

O9. **Linear combining**:
Linear combining is like summing except that we allow amplitude scaling (*i.e.*, multiply the signals by constants) in addition to summing, as in

$$y(t) = 3x_1(t) + 4x_2(t) - 2x_3(t) \,.$$

In this case, $y(t)$ is said to be a **linear combination** of $x_1(t)$, $x_2(t)$, and $x_3(t)$. The scale factors multiplying the signals are often called **coefficients**.

Example. Audio mixer in recording studio.

Linear combinations arise in a several ways. As one example, sometimes we are given a collection of signals, say $x_1(t)$, $x_2(t)$ and $x_3(t)$ and are asked to **synthesize** another signal $y(t)$ as a linear combination of the signals in the collection. For example, suppose we need to create the signal $y(t)$, but our hardware can only produce signals $x_1(t)$, $x_2(t)$ and $x_3(t)$ and perform linear combinations. Often, it is not possible to exactly synthesize $y(t)$ from the given collection so the synthesis must necessarily be approximate.

As another example, sometimes we are given a signal $z(t)$ that is known to be a linear combination of $x_1(t)$, $x_2(t)$ and $x_3(t)$, and we are asked to find the scale factors. This task, which is called **analysis**, happens for example in communications systems, where the scale factors determine the information carried by the signal $y(t)$. It also happens in **Fourier analysis**, to be discussed considerably throughout the course, where we consider a signal $y(t)$ to be the linear combination of sinusoidal signals with different fundamental frequencies.

O10. **Multiplying**:
As its name suggests, this is simply the operation of creating a new signal as the product of two or more signals, as in

$$z(t) = x(t)\, y(t) \,.$$

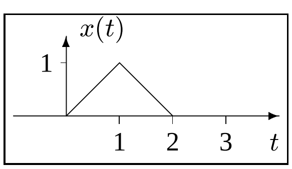
More specifically, the value of z at time each time t is the product of x at time t and y at time t.

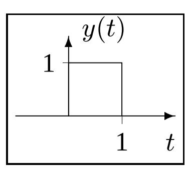
Example. Signal multiplication is a basic operation of most radio transmitters which, as in the example of AM radio described earlier, typically multiply a sinusoidal signal by some information-bearing signal.

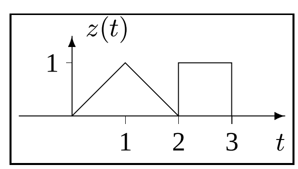
O11. **Concatenating**:
Appending one signal to the end of another is called **concatenation**.

Example. If $x(t)$ is a signal with support interval $(0, t_1)$ and $y(t)$ is a signal with support $(0, t_2)$, then as illustrated below their concatenation is the signal

$$z(t) = \begin{cases} x(t)\,, & t \le t_1 \\ y(t - t_1)\,, & t > t_1. \end{cases}$$

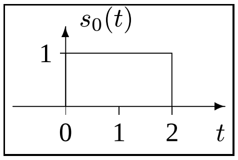


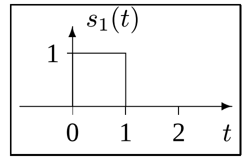




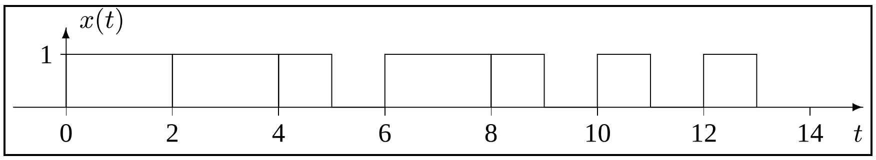
Example. Concatenation occurs in digital communications where, to transmit a sequence at the rate of one bit every T seconds, there is a signal $s_0(t)$ with support within $(0, T)$ used to send 0's, and a signal $s_1(t)$ also with support within $(0, T)$ used to send 1's. The transmitted signal is the concatenation of these.

Example. When the signals shown below





are used to send the binary sequence 0,0,1,0,1,1,1,..., with $T = 2$, the transmitted signal[10] is the following.



(This is an example of pulse-width modulation, which is described in more detail in EECS 353 and EECS 455.)

Concluding Remarks

The signal operations discussed in this section are elementary operations that are used in a variety of situations. One may view them as basic tools or building blocks. The signal operations considered later in the course (*e.g.*, Chapters 5-8 of the text) are more sophisticated operations, which are developed with some specific task in mind. These operations can be thought of as **systems**; that is, when the operation is applied to a signal $x(t)$, the signal $x(t)$ is viewed as the input to a **system** that performs the operation and produces at its output another signal $y(t)$, which is the result of the operation. In such cases, we often draw a block diagram like the one shown below. Much of the course will be devoted to designing systems to perform the tasks described in the Section IV.

$$x(t) \rightarrow \boxed{\text{System}} \rightarrow y(t)$$

Example. Telephone system.

[10] As usual, vertical lines shown just emphasize the transitions between transmitted bits, as well as the jumps from 0 to 1 and 1 to 0.

Effects of signal operations on signal characteristics

So far we have covered the following
- Signal characteristics (duration, energy, etc., about 13)
- Signal operations (amplitude scaling, etc., about 11)

It is logical to expect that performing **operations** (even simple ones!) on signals will change their **characteristics**.
So we could now derive about 13×11 "properties" that describe the effect of operation X on characteristic Y.
Instead, we will work a couple examples that illustrate the *methods* one can use to derive such properties when needed.

Example. Operation: time scaling. Characteristic: duration.

Suppose $x(t)$ is a finite-duration signal with support interval $[t_1, t_2]$.
Let $y(t) = x(2t)$. Find the duration of $y(t)$.

Answer.
Since $x(t)$ has support interval $[t_1, t_2]$, its nonzero values occur when $t_1 \leq t \leq t_2$. (Use what is given about "input signal.")
Since $y(t) = x(2t)$, its nonzero values occur when $t_1 \leq 2t \leq t_2$. (Use what is known about relationship between y and x.)
Rearranging, we see that the nonzero values of $y(t)$ occur when $t_1/2 \leq t \leq t_2/2$. (Use math.)
Thus, the **duration** of $y(t)$ is $t_2/2 - t_1/2 = \frac{1}{2}(t_2 - t_1)$, so $\boxed{\text{duration(y) = duration(x) / 2.}}$

Example. Operation: time scaling. Characteristic: energy.

If $y(t) = x(-2t)$, relate the energy of $y(t)$ to the energy of $x(t)$.
For simplicity we consider a finite-support signal $x(t)$, with support $[t_1, t_2]$.
By similar argument as above, the support interval of $y(t)$ is $[-t_2/2, -t_1/2]$. So the energy $\mathrm{E}(y)$ is given by:

$$
\begin{aligned}
\mathrm{E}(y) &= \int_{-t_2/2}^{-t_1/2} y^2(t)\, \mathrm{d}t \qquad \text{(Definition)} \\
&= \int_{-t_2/2}^{-t_1/2} x^2(-2t)\, \mathrm{d}t \qquad \text{(Substitute given relationship)} \\
&= \int_{t_2}^{t_1} x^2(t')\, \frac{\mathrm{d}t'}{-2} \qquad \text{(Calculus: let } t' = -2t\text{)} \\
&\quad \text{Why? To make the integrand match the formula for } \mathrm{E}(x). \\
&= \int_{t_1}^{t_2} x^2(t')\, \frac{\mathrm{d}t'}{2} \qquad \text{(Calculus: exchanging limits)} \\
&= \frac{1}{2} \int_{t_1}^{t_2} x^2(t')\, \mathrm{d}t' = \frac{1}{2}\mathrm{E}(x) \qquad \text{(Using energy definition again)}
\end{aligned}
$$

Exercise: show the following. $\boxed{\text{If } y(t) = x(ct) \text{ for } c \neq 0, \text{ then } \mathrm{E}(y) = \frac{1}{|c|}\mathrm{E}(x).}$

Can one memorize all 100+ such properties? Can one "cram" them in before an exam? Unlikely.
Instead, one must learn these methods by working problems, paying attention to the tools used to find the solutions, so as to be able to apply those tools when needed for future problems.

Example. Operation: signal addition. Characteristic: energy. (Relate $\mathrm{E}(x + y)$ to $\mathrm{E}(x)$ and $\mathrm{E}(y)$ and ?.)

$$\mathrm{E}(x+y) = \sum_n (x[n] + y[n])^2 = \sum_n x^2[n] + 2\sum_n x[n]\, y[n] + \sum_n y^2[n] = \mathrm{E}(x) + 2\sum_n x[n]\, y[n] + \mathrm{E}(y)\,.$$

The term $\sum_n x[n]\, y[n]$ is called the **correlation** of $x[n]$ and $y[n]$, and is central to the next topic!