Supplementary Notes for ELEN 4810 Lecture 2
Introduction to Linear Time Invariant Systems

John Wright
Columbia University

September 15, 2025

Disclaimer: These notes are intended to be an accessible introduction to the subject, with no pretense at completeness. In general, you can find more thorough discussions in Oppenheim's book. Please let me know if you find any typos.

Reading suggestions: Oppenheim and Schafer Sections 2.4-2.5

In this lecture, we will begin our discussion of linear systems by defining some basic categories of systems. We will spend much of our discussion on linear, time-invariant (LTI) systems, convolution and the impulse response. Our development will mirror that of the text, with some minor differences in notation.

1 Discrete-Time Systems

A discrete time system takes as its input a discrete time signal $x[n]$, and outputs another discrete time signal $y[n]$. We often write this as[1]

$$y = \mathcal{T}\left\{ x \right\}. \tag{1.1}$$

This is an extremely general abstract model. It is flexible enough to model propagation of a signal of interest through a medium (think, acoustics in the room, RF communications, optical microscopy), propagation of disturbances through a medium (think, seismic activity through the earth's crust, or vibrations from passing cars a bridge), or systems that we design and implement ourselves to achieve some desired effect (think, amplification, noise cancellation, equalization). Depending on the application, there are several different types of problems that we might wish to solve:

[1]The text uses the notation $y[n] = \mathcal{T}\left\{x[n]\right\}$. The simpler notation in (1.1) emphasizes that the system $\mathcal{T}$ acts on the signal x *as a whole* to produce the output y *as a whole*. In the classical (and widely used) notation adopted by the textbook, there is an ambiguity in the expression "$x[n]$" – depending on context it could mean either the whole sequence, or just the particular value at time n. Writing "x" when we mean the whole signal avoids this ambiguity.

- **Design**: Choose $\mathcal{T}$ that implements (or approximates) some desired input-output relationship, subject to resource constraints. Here, resources could include hardware, computation time, energy, memory, data, etc.
- **Identification**: If $\mathcal{T}$ represents some physical system, it may not be known ahead of time. The system identification problem asks us to determine $\mathcal{T}$ from input-output pairs $(x_1, y_1), \dots, (x_N, y_N)$.
- **Inversion**: Given $\mathcal{T}$ and y, determine x. This problem is important for applications in sensing, medical and scientific imaging, etc., where $\mathcal{T}$ either represents some degradation (blur, optical abberation) or the effect of the sensor itself, and our goal is to determine what physical signal x has been observed.

The general philosophy in classical signal processing is to identify classes of systems with common mathematical properties and develop tools and insights for these mathematical models. Ideally, this gives a framework for thinking clearly about basic issues (is the system stable? when can I identify the system?) while also being flexible enough to incorporate application-specific structure.

We will begin by discussing a few model discrete time systems, which are far simpler than the application examples described above, but which will serve as useful examples for illustrating basic properties such as stability and causality.

The ideal delay. The *ideal delay* $\mathcal{D}_k$ simply shifts the signal to the right by k samples:

$$y[n] = x[n-k], \qquad \forall n. \tag{1.2}$$

Here, $k \in \mathbb{Z}$ is an integer. If k is positive, then we produce y by shifting x to the right – the value that would have appeared at time n now appears at time $n+k$. If k is negative, we produce y by shifting x to the left. We can write the ideal delay in operator notation as

$$y = \mathcal{D}_k x. \tag{1.3}$$

The moving average. The *moving average* sets $y[n]$ to be the average of x over a window spanning from $n-m$ to $n+M$:

$$y[n] \quad = \quad \frac{1}{m+M+1} \sum_{k=-M}^{m} x[n-k]. \tag{1.4}$$

The accumulator. The *accumulator* sets $y[n]$ to be the summation (the accumulation!) of all of the samples $x[n']$ up to time $n' = n$: $y[n] = \sum_{k=-\infty}^{n} x[k]$.

2 Properties of Systems

We will define several useful categories of systems.

Memoryless systems. A system is *memoryless* if the n-th sample $y[n]$ of the output depends only on the n-th sample $x[n]$ of the input. More formally:

Definition 2.1 (Memoryless system). *The system $y = \mathcal{T}\{x\}$ is* memoryless *if for each $n \in \mathbb{Z}$, there exists a function $f_n : \mathbb{C} \to \mathbb{C}$ such that*

$$y[n] = f_n(x[n]). \tag{2.1}$$

For example, the system defined by the relationship

$$y[n] = \exp(x[n]) \tag{2.2}$$

is memoryless.

Causal systems. Informally speaking, a system is *causal* if the output at time n only depends on the input $x[n]$ at times $n' \le n$. *If* we want to make this a little more formal, we can let $\mathcal{S}_n$ denote the collection of complex-valued sequences $w[n']$ that are defined over the integers $n' \le n$ – an element of $\mathcal{S}_n$ is a sequence $\{\dots, w[n-2], w[n-1], w[n]\}$. The system is causal if there exist functions $g_n : \mathcal{S}_n \to \mathbb{C}$, such that for each n,

$$y[n] = g_n\Big(\{x[n'] \mid n' \le n\}\Big). \tag{2.3}$$

That is to say, the output depends only on the past and present values of the input, not on the future! This is clearly a useful property if we want to process signals in "real time" – that is to say, as they are generated.

Consider the forward difference: $y[n] = x[n+1] - x[n]$. Is it causal? What about the backward difference $y[n] = x[n] - x[n-1]$? The accumulator? Is a general memoryless system causal?

Stable systems. Informally, a system is stable if its output does not "blow up" when the input is bounded in an appropriate sense. There are many useful ways of formalizing this desirable property. We will mostly work with the notion of *bounded input bounded output* (*BIBO*) stability. We say that a signal x is *bounded* if there exists some $B \in \mathbb{R}$ such that

$$|x[n]| \le B, \qquad \forall\, n \in \mathbb{Z}. \tag{2.4}$$

In this definition, the bound B *can* depend on x, but *cannot* depend on the time n. So B provides a uniform bound on the size of the entries of x, which holds over all times n.

Definition 2.2 (BIBO stability). *A system $y = \mathcal{T}\{x\}$ is* bounded input bounded output (BIBO) *stable if for every bounded x, $y = \mathcal{T}\{x\}$ is also bounded.*

For a BIBO system, whenever the input is bounded (there exists a B such that $|x[n]| \le B\,\forall\, n$), the output is bounded (there exists a B' such that $|y[n]| \le B'\,\forall\, n$). To reiterate, the bound B' on the size of the output y *can* (and usually does) depend on the input x, but *cannot* depend on time n.

To check that a system is BIBO stable, we need to prove that this property holds for every input x, while to show that it is unstable, we just need to demonstrate one "bad" input x which is bounded, but creates an unbounded output.

Is the forward difference $y[n] = x[n+1] - x[n]$ BIBO stable? What about the accumulator $y[n] = \sum_{i=-\infty}^{n} x[i]$?

Linear systems. Linear systems are those which satisfy the principle of superposition – if we input a linear combination of signals x_1 and x_2, the system outputs the same linear combination of the corresponding outputs $y_1 = \mathcal{T}\{x_1\}$ and $y_2 = \mathcal{T}\{x_2\}$. More formally:

Definition 2.3 (Linear system). *A system $y = \mathcal{T}\{x\}$ is* linear *if for any signals x_1, x_2 and scalars $\alpha, \beta \in \mathbb{C}$,*

$$\mathcal{T}\{\alpha x_1 + \beta x_2\} = \alpha \mathcal{T}\{x_1\} + \beta \mathcal{T}\{x_2\}. \tag{2.5}$$

It is sometimes helpful conceptually to consider separately the effect of superimposing the signals x_1 and x_2 and the effect of multiplying by a scalar α. Please convince yourself that the system $y = \mathcal{T}\{x\}$ is linear if and only if the following two conditions both hold:

$$\begin{array}{lll} \textbf{Superposition} & \forall\, x_1,\ x_2, & \mathcal{T}\{x_1 + x_2\} = \mathcal{T}\{x_1\} + \mathcal{T}\{x_2\} \quad (2.6) \\ \textbf{Homogeneity} & \forall\, x,\ \alpha \in \mathbb{C}, & \mathcal{T}\{\alpha x\} = \alpha \mathcal{T}\{x\}. \quad (2.7) \end{array}$$

For example, consider the accumulator

$$y[n] = \sum_{k=-\infty}^{n} x[k]. \tag{2.8}$$

Is this a linear system? We can simply check the definition. Let

$$y_1 = \mathcal{T}\{x_1\}, \qquad y_2 = \mathcal{T}\{x_2\}, \tag{2.9}$$

so

$$y_1[n] = \sum_{k=-\infty}^{n} x_1[k] \tag{2.10}$$

$$y_2[n] = \sum_{k=-\infty}^{n} x_2[k] \tag{2.11}$$

Let $y = \mathcal{T}\{\alpha x_1 + \beta x_2\}$. Then

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{n} (\alpha x_1 + \beta x_2)[k] && (2.12) \\ &= \sum_{k=-\infty}^{n} \alpha x_1[k] + \beta x_2[k] && (2.13) \\ &= \alpha \sum_{k=-\infty}^{n} x_1[k] + \beta \sum_{k=-\infty}^{n} x_2[k] && (2.14) \\ &= \alpha\, y_1[n] + \beta\, y_2[n], && (2.15) \end{aligned}$$

and so the system is indeed linear.

To show that a system is nonlinear, we just need to show one example input for which the definition does not hold. For example, consider the memoryless system defined by $y[n] = (x[n])^2$. It should be fairly intuitive that this system is nonlinear – it involves a square! To prove that it is nonlinear, take (for example) $x_1[n] = \delta[n]$ and $x_2[n] = \delta[n]$, and verify that the property fails.

Time invariant systems. Time invariant systems are those that satisfy the following very natural property: if we delay the input x by k samples, the output y is delayed by k samples.

So, if $\mathcal{T}\{x\} = y$, and

$$\bar{x}[n] = x[n-k], \tag{2.16}$$

then $\mathcal{T}\{\bar{x}\} = \bar{y}$, where

$$\bar{y}[n] = y[n-k]. \tag{2.17}$$

This can be stated tersely as follows:

Definition 2.4 (Time invariance)**.** *A system $\mathcal{T}$ is* time-invariant *if for every integer $k \in \mathbb{Z}$ and every signal x,*

$$\mathcal{T}\{\mathcal{D}_k x\} = \mathcal{D}_k \mathcal{T}\{x\}. \tag{2.18}$$

For example, we can check that the accumulator is time invariant. Set

$$y[n] = \sum_{i=-\infty}^{n} x[i], \tag{2.19}$$

and

$$\bar{x}[n] = x[n-k]. \tag{2.20}$$

Let $\bar{y}$ be the output of the accumulator, applied to input $\bar{x}$. Then

$$\begin{aligned} \bar{y}[n] &= \sum_{i=-\infty}^{n} \bar{x}[i] \\ &= \sum_{i=-\infty}^{n} x[i-k] \\ &= \sum_{i=-\infty}^{n-k} x[i] \\ &= y[n-k], \end{aligned} \tag{2.21}$$

as desired.

On the other hand, to show that a system is *not* time invariant, it suffices to check that the definition fails for some specific input. For example, consider the system[2] defined by

$$y[n] = x[2n]. \tag{2.22}$$

Note that $y[-1] = x[-2]$. Consider an input sequence x for which $x[-2] \neq x[-1]$. Let

$$\bar{x}[n] = x[n-1] \tag{2.23}$$

be a version of x that is shifted one sample to the right. If we apply the system to $\bar{x}$, we obtain

$$\begin{aligned} \bar{y}[0] &= \bar{x}[0] \\ &= x[-1] \\ &\neq y[-1]. \end{aligned} \tag{2.24}$$

[2] Sometimes called a *compressor* or *downsampler*.

3 Linear Time Invariant Systems

A system $y = \mathcal{T}\{x\}$ is *linear, time-invariant* (*LTI*) if it is both linear and time-invariant. This class of systems is broad enough to contain many useful operations and to be capable of modeling or approximating many physical systems arising in applications. It is also very well-structured for analysis and computation.

To see why, let us recall the representation of an arbitrary signal $x[n]$ as a superposition of shifted unit impulses $\delta[n-k]$:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \tag{3.1}$$

This gives an expression for the n-th sample of x. We can rewrite this in slightly different notation as

$$x = \sum_{k=-\infty}^{\infty} x[k]\ \mathcal{D}_k\delta. \tag{3.2}$$

This expresses the signal x (as a whole) as a superposition of shifted impulses $\mathcal{D}_k\delta$, with coefficients given by the values $x[k]$ of the signal x itself.

Let $h[n]$ denote the response of the system to a unit impulse $\delta[n]$:

$$h = \mathcal{T}\{\delta\}. \tag{3.3}$$

Using linearity and time-invariance, we can write the response to an arbitrary input x as

$$\begin{aligned}
\mathcal{T}\{x\} &= \mathcal{T}\left\{\sum_{k=-\infty}^{\infty} x[k]\ \mathcal{D}_k\delta\right\} \\
&= \sum_{k=-\infty}^{\infty} x[k]\ \mathcal{T}\{\mathcal{D}_k\delta\} \qquad \text{(by linearity)} \\
&= \sum_{k=-\infty}^{\infty} x[k]\ \mathcal{D}_k\ \mathcal{T}\{\delta\} \qquad \text{(by time invariance)} \\
&= \sum_{k=-\infty}^{\infty} x[k]\ \mathcal{D}_k h.
\end{aligned} \tag{3.4}$$

Since $(\mathcal{D}_k h)[n] = h[n-k]$, if $y = \mathcal{T}\{x\}$, then

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]. \tag{3.5}$$

This summation is called a *convolution sum*. We say that y is the *convolution* of x and h, and write this

$$y = x * h. \tag{3.6}$$

The $*$ notation means that the sequences y and h are related by (3.5). To evaluate $x * h$ at the n-th sample, we simply use this definition:

$$\begin{aligned} y[n] &= (x * h)[n] \\ &= \sum_{k=-\infty}^{\infty} x[k]h[n-k]. \end{aligned} \tag{3.7}$$

The sequence $h = \mathcal{T}\{\delta\}$ is called the *impulse response* of the system. From (3.5), *the impulse response* $h[n]$ *completely characterizes the system.*

4 Properties of Discrete Convolution

The most important result thus far is the characterization of linear time invariant (LTI) systems in terms of their impulse response: every LTI system can be implemented as a convolution of the input x with the system's impulse response h. We next show how we can study the properties of a system through its impulse response. We establish a few natural properties of the convolution sum

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]. \tag{4.1}$$

Commutativity. The commutative property says that the order of x and h in the convolution sum does not matter:

Proposition 4.1. *Convolution is* commutative*: for every* x, h*,*

$$x * h = h * x. \tag{4.2}$$

Proof. We need to show that for every n, $(x * h)[n] = (h * x)[n]$. We just calculate

$$\begin{aligned} (x * h)[n] &= \sum_{k=-\infty}^{\infty} x[k]h[n-k] \\ &= \sum_{i=-\infty}^{\infty} x[n-i]h[i] \qquad (\text{set } i = n-k) \\ &= \sum_{i=-\infty}^{\infty} h[i]x[n-i] \\ &= (h * x)[n]. \end{aligned} \tag{4.3}$$

This establishes the result. $\square$

This means that a system with impulse response h *applied to input* x *produces exactly the same output as a system with impulse response* x *applied to input* h*.*

Linearity.

Proposition 4.2. *Convolution is* linear in *h: for every $x, h_1, h_2, \alpha, \beta \in \mathbb{C}$,*

$$x * (\alpha h_1 + \beta h_2) = \alpha(x * h_1) + \beta(x * h_2). \tag{4.4}$$

Proof. Again, we simply calculate:

$$\begin{aligned}
\{x * (\alpha h_1 + \beta h_2)\}[n] &= \sum_{k=-\infty}^{\infty} x[k](\alpha h_1[n-k] + \beta h_2[n-k]) \\
&= \alpha \sum_{k=-\infty}^{\infty} x[k]h_1[n-k] + \beta \sum_{k=-\infty}^{\infty} x[k]h_2[n-k] \\
&= \alpha(x * h_1)[n] + \beta(x * h_2)[n]. \qquad (4.5)
\end{aligned}$$

□

This immediately implies that convolution is also linear in x (convince yourself of this):

$$(\alpha x_1 + \beta x_2) * h = \alpha(x_1 * h) + \beta(x_2 * h). \tag{4.6}$$

It also gives a way of simplifying systems: *a system which sums the outputs of two LTI systems with impulse responses h_1 and h_2 is completely equivalent to a single LTI system with impulse response $h = h_1 + h_2$.*

Associativity.

Proposition 4.3. *Convolution is associative:*

$$(x * h_1) * h_2 = x * (h_1 * h_2). \tag{4.7}$$

Proof.

$$\begin{aligned}
\{(x * h_1) * h_2\}[n] &= \sum_{k=-\infty}^{\infty} (x * h_1)[k]h_2[n-k] \\
&= \sum_{k=-\infty}^{\infty} \left(\sum_{i=-\infty}^{\infty} x[i]h_1[k-i] \right) h_2[n-k] \\
&= \sum_{i=-\infty}^{\infty} x[i] \sum_{k=-\infty}^{\infty} h_1[k-i]h_2[n-k] \\
&= \sum_{i=-\infty}^{\infty} x[i] \sum_{k'=-\infty}^{\infty} h_1[k']h_2[n-i-k'] \qquad \text{(set } k' = k - i\text{)} \\
&= \sum_{i=-\infty}^{\infty} x[i](h_1 * h_2)[n-i]. \qquad (4.8)
\end{aligned}$$

□

This means that if we concatenate two LTI systems with impulse responses h_1 and h_2, the concatenation produces the same output as single equivalent system with impulse response $h = h_1 * h_2$.

5 Stability of LTI systems

Recall that a system is *Bounded Input Bounded Output* (BIBO) stable if every bounded x (i.e., x for which there exists B_x such that $|x[n]| \leq B_x$ for every n) produces a bounded output y (i.e., a y for which there exists B_y such that $|y[n]| \leq B_y$ for every n). There is a very clean characterization of stable LTI systems in terms of the impulse response:

Theorem 5.1. *A linear, time-invariant system is bounded input bounded output stable if and only if its impulse response h is absolute summable:*

$$\sum_{k=-\infty}^{\infty} |h[k]| < +\infty. \tag{5.1}$$

A notation: The summation $\sum_{k=-\infty}^{\infty} |h[k]|$ above is often written as $\|h\|_{\ell^1}$. The condition can be stated as $\|h\|_{\ell^1} < +\infty$.

Proof. We first show that whenever $\sum_k |h[k]| < +\infty$, the system is stable. Consider an arbitrary bounded input x, for which

$$|x[n]| \;\leq\; B_x, \qquad \forall\, n. \tag{5.2}$$

Write

$$\begin{aligned} |y[n]| &= \left| \sum_{k=-\infty}^{\infty} x[k]h[n-k] \right| \\ &\leq \sum_{k=-\infty}^{\infty} |x[k]h[n-k]| \\ &= \sum_{k=-\infty}^{\infty} |x[k]|\,|h[n-k]| \\ &\leq B_x \sum_{k=-\infty}^{\infty} |h[n-k]| \\ &= B_x \|h\|_{\ell^1} \\ &< +\infty. \end{aligned} \tag{5.3}$$

Since $y[n] \leq B_x \|h\|_{\ell^1}$ for every n, the output y is bounded, and the system is BIBO stable.

Now, suppose that $\|h\|_{\ell^1} = \sum_k |h[k]| = +\infty$. We prove that the system is not BIBO stable. To show this, it is enough to demonstrate one bounded input x which produces an unbounded output. For this purpose, take

$$x[n] = \frac{h^*[-n]}{|h[-n]|}. \tag{5.4}$$

Notice that $|x[n]| \leq 1$ for every n, and so x is indeed bounded. Let y denote the output of the system, applied to x. Then

$$\begin{aligned}
y[0] &= \sum_{k=-\infty}^{\infty} x[k]h[0-k] \\
&= \sum_{k=-\infty}^{\infty} \frac{h^*[-k]}{|h[-k]|} h[-k] \\
&= \sum_{k=-\infty}^{\infty} \frac{|h[-k]|^2}{|h[-k]|} \\
&= \sum_{k=-\infty}^{\infty} |h[k]| \\
&= +\infty. \qquad (5.5)
\end{aligned}$$

The output y is unbounded. Since there exists a bounded input x which produces an unbounded output y, the system is not BIBO stable. This completes the proof. □

6 Causal LTI systems

It is also relatively straightforward to characterize the causal LTI systems in terms of their impulse responses.

Theorem 6.1. *A linear, time-invariant system is causal if and only if its impulse response h satisfies*

$$h[n] = 0, \qquad \forall\, n < 0. \qquad (6.1)$$

Proof. Suppose that h satisfies (6.1). Then

$$\begin{aligned}
y[n] &= \sum_{k=-\infty}^{\infty} x[k]h[n-k] \\
&= \sum_{k=-\infty}^{n} x[k]h[n-k]. \qquad (h[n-k] = 0 \text{ when } k > n) \qquad (6.2)
\end{aligned}$$

So, $y[n]$ is a function of $\{x[k] \mid k \leq n\}$ only.

Conversely, if $h[\tilde{n}] \neq 0$ for some $\tilde{n} < 0$, we can generate two inputs x_1 and x_2 with $x_1[n] = x_2[n]$ for every $n \neq -\tilde{n}$, and $x_1[-\tilde{n}] \neq x_2[-\tilde{n}]$. Let y_1 and y_2 denote the outputs of the system with inputs x_1 and x_2 respectively. Then

$$\begin{aligned}
y_1[0] &= \sum_{k=-\infty}^{\infty} x_1[k]h[-k] \\
&= \sum_{k=-\infty}^{\infty} x_2[k]h[-k] + h[\tilde{n}](x_1[-\tilde{n}] - x_2[-\tilde{n}]) \\
&= y_2[0] + h[\tilde{n}](x_1[-\tilde{n}] - x_2[-\tilde{n}]) \\
&\neq y_2[0]. \qquad (6.3)
\end{aligned}$$

The inputs x_1 and x_2 agree for every $n \le 0$, but produce different outputs at time $n = 0$. Hence, the system is not causal. $\square$

Inspired by this characterization of causal LTI systems, we sometimes say that a sequence $h[n]$ *is causal if* $h[n] = 0$ *for every* $n < 0$.

For each of the following model systems, what is the impulse response h*? Is the system causal? Is it BIBO stable?* (i) *The ideal delay, with delay* k (ii) *The moving average, with parameters* M_1 *and* M_2 (iii) *the accumulator.*

7 Finite and Infinite Impulse Response Systems

A system has *finite impulse response* if its impulse response $h[n]$ is nonzero only for finitely many times n. More precisely, write

$$\mathrm{support}(h) = \{n \mid h[n] \neq 0\}\,. \tag{7.1}$$

A system with impulse response h has *finite impulse response* (*FIR*) if $|\mathrm{support}(h)| < +\infty$.[3] We say it has *infinite impulse response* (*IIR*) if $|\mathrm{support}(h)| = +\infty$.

The accumulator is an example of an IIR system. It is not stable. However, there exist IIR systems that *are* stable – we just need the impulse response to decay quickly enough. For example, if

$$h[n] = \begin{cases} \alpha^n & n \ge 0 \\ 0 & n < 0 \end{cases} \tag{7.2}$$

then if $|\alpha| < 1$,

$$\begin{aligned} \sum_{k=-\infty}^{\infty} |h[k]| &= \sum_{k=0}^{\infty} |\alpha^k| \\ &= \sum_{k=0}^{\infty} |\alpha|^k \\ &= \frac{1}{1-|\alpha|} \\ &< +\infty, \end{aligned} \tag{7.3}$$

and the system is stable.

8 Correlation and Convolution

As a brief of addendum on our discussion of discrete convolution, we discuss *correlation* and its relationship to convolution. To understand this clearly, we need the notion of an inner product between complex signals.

[3] An equivalent definition is that h has finite impulse response if and only if there exist $n_1 \le n_2$ such that $h[n] = 0$ for all $n < n_1$ and $h[n] = 0$ for all $n > n_2$.

Digression: inner products for complex vectors. For n-dimensional *real valued* vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, the standard inner product (or dot product) is

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i. \tag{8.1}$$

If we let

$$\|\boldsymbol{x}\|_{\ell^2} = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{8.2}$$

denote the length of a vector $\boldsymbol{x}$, then you may recall that the inner product between $\boldsymbol{x}$ and $\boldsymbol{y}$ can be written as $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\|_{\ell^2} \|\boldsymbol{y}\|_{\ell^2} \cos\theta$, where θ is the *angle* between the vectors. Since $|\cos\theta| \leq 1$ for all θ, this implies that

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq \|\boldsymbol{x}\|_{\ell^2} \|\boldsymbol{y}\|_{\ell^2}, \tag{8.3}$$

a fact which is sometimes referred to as the *Cauchy-Schwarz inequality*. Moreover, we can compute the length of a vector from its inner product with itself:

$$\|\boldsymbol{x}\|_2 = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}. \tag{8.4}$$

It is useful to be able to define an inner product for complex vectors, which works in essentially the same way. For a complex vector $\boldsymbol{x} \in \mathbb{C}^n$, we can define its length as

$$\|\boldsymbol{x}\|_{\ell^2} = \sqrt{\sum_{i=1}^{n} |x_i|^2}. \tag{8.5}$$

We define the inner product between two complex vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ as

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i^*. \tag{8.6}$$

Note the conjugation of y_i! This seemingly bizarre convention is important for making the inner product work in ways that conform to our intuition. In particular, with this choice, it is always true that $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ is real-valued, and that

$$\|\boldsymbol{x}\|_{\ell^2} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}. \tag{8.7}$$

The definition (8.6) implies that the inner product is conjugate symmetric: reversing the order of $\boldsymbol{x}$ and $\boldsymbol{y}$ is equivalent to taking the complex conjugate of the inner product:

$$\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*. \tag{8.8}$$

This is a defining property of inner products in complex vector spaces. Our purpose here is not to dwell on the mathematics of inner product spaces. However, the convention (8.6) is worth remembering – in particular, in the next lecture as we move into the Fourier domain, where it makes expressions for Fourier transforms much easier to remember.

If we consider (as we have above), complex-valued sequences $x[n]$ and $y[n]$ instead of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, we can still make the same notation:

$$\langle x, y \rangle = \sum_{k=-\infty}^{\infty} x[k] y^*[k]. \tag{8.9}$$

Notice that in (8.9), we have taken the complex conjugate of the sequence y. As above, this is essential to have a complex inner product that functions correctly. For the definition (8.9) to make sense, x and y need to be "nice enough" that the infinite summation exists. A sufficient condition is that these two sequences are square-summable:

$$\sum_{k=-\infty}^{\infty} |x[k]|^2 \ < \ +\infty, \qquad \sum_{k=-\infty}^{\infty} |y[k]|^2 \ < \ +\infty. \tag{8.10}$$

As with vectors, if we fix the energy $\|x\|_{\ell^2} = \sqrt{\sum_{k=-\infty}^{\infty} |x[k]|^2}$ and $\|y\|_{\ell^2} = \sqrt{\sum_{k=-\infty}^{\infty} |y[k]|^2}$, the inner product $\langle x, y \rangle$ will be largest when x and y are aligned. In fact, the inner product can be taken as a measure of the similarity of two sequences.

Discrete correlation. The correlation of two discrete time signals x and y is defined as

$$r_{xy}[n] = \sum_{k=-\infty}^{\infty} x^*[k] y[k+n]. \tag{8.11}$$

That is to say, we translate y by n samples to the left, and then take the inner product with x:

$$r_{xy}[n] = \langle \mathcal{D}_{-n} y, x \rangle \, . \tag{8.12}$$

Correlation is useful for (very) basic signal detection tasks – the output r_{xy} consists of the inner product of x with shifted versions of the input y. Often, this value is large when the shifted version of y is similar to x.[4] Correlation is also a basic tool for studying *random* signals and random processes.

Correlation can be implemented using convolution: let $\tilde{y}[n] = y[-n]$. Then

$$\begin{aligned} (x^* * \tilde{y})[-n] &= \sum_{k=-\infty}^{\infty} x^*[k] \tilde{y}[-n-k] \\ &= \sum_{k=-\infty}^{\infty} x^*[k] y[k+n] \\ &= r_{xy}[n]. \end{aligned} \tag{8.13}$$

So, to compute the correlation of two signals x and y, "just" conjugate x, reverse y, convolve them, and then reverse the result.

[4]In the homework, you will have a chance to play with this, and see its limitations, and a few more practical variants!