# Digital Signal Processing

**By:**
Mark A. Davenport

# Digital Signal Processing

**By:**
Mark A. Davenport

**Online:**
< http://cnx.org/content/col11172/1.4/ >

**C O N N E X I O N S**

**Rice University, Houston, Texas**

This selection and arrangement of content as a collection is copyrighted by Mark A. Davenport. It is licensed under the Creative Commons Attribution 3.0 license (http://creativecommons.org/licenses/by/3.0/).
Collection structure revised: December 16, 2011
PDF generated: December 16, 2011
For copyright and attribution information for the modules contained in this collection, see p. 62.

# Table of Contents

# Introduction to Digital Signal Processing[1]

## Information, Signals and Systems

Signal processing concerns primarily with signals and systems that operate on signals to extract useful information. In this course our concept of a "signal" will be very broad, encompassing virtually any data that can be represented as an organized "collection" of data.

**Example**

- A continuous function $f(t)$
- A sequence of discrete data points $f[n]$
- A multi-dimensional array of data
- Audio, images, video, voltage of antenna
- Stock prices, potassium concentration in a neuron

Our concept of a "system" will be a black box that takes a signal as input and provides another signal as output.

**Example**

- Analog-to-digital converters (ADCs)
- Filters
- Decimators/Interpolators
- Matched filters
- Face recognition systems

In this course we will approach signal processing from the point of view that signals are *vectors* living in an appropriate *vector space*, and systems are *operators* that map signal from one vector space to another. This allows us to use a common mathematical framework to talk about how to:

- represent signals
- measure similarity/distance between signals
- transform signals from one representation to another
- understand the operation of linear systems on the signals

Since the ficus of this course in on digital signal processing, this will also allow us to use tools from linear algebra to facilitate this understanding.

## Digital Signal Processing

DSP is often presented as an alternative to analog signal processing, i.e., instead of a purely analog system as in Figure 1, we can build a digital implementation of an analog system as in Figure 2. This can

[1]This content is available online at <http://cnx.org/content/m33588/1.2/>.

be advantageous since high-precision analog components are expensive (even compared to the cost of an ADC/DAC).

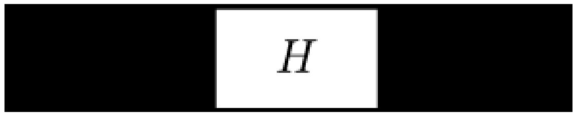



**Figure 1:** An analog system.

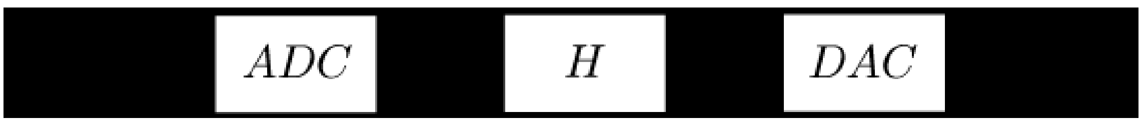
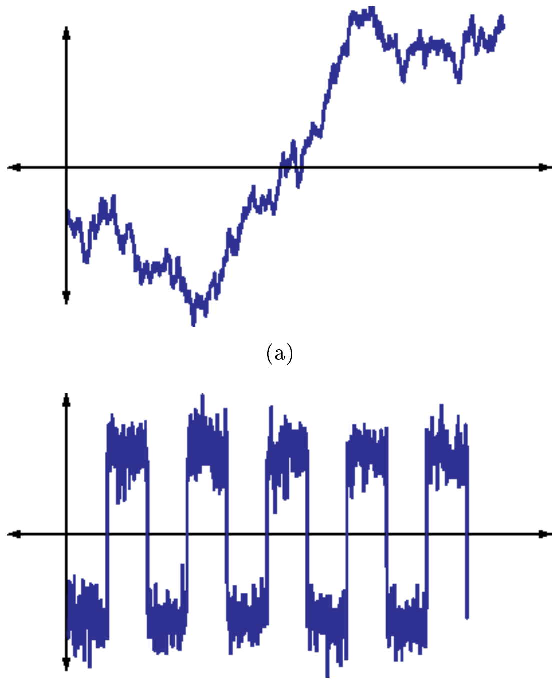



**Figure 2:** A digital implementation of an analog system.

However, the success of DSP derives to a much greater extent from the facts that:

1. Discrete-valued signals can be more robust to noise, as illustrated in Figure 3. In Figure 3(a), noise may be impossible to eliminate, but in Figure 3(b) noise can be eliminated entirely by exploiting the discrete structure of the signal.
2. Once we have a digital, discrete-time signal, we can store it in memory and perform highly complex processing.



(a)

(b)

**Figure 3:** (a) An analog signal corrupted with noise; (b) A discrete-valued signal corrupted with noise.

In this course we will consider signal processing systems beyond simple LTI filters. Themes of the course include:

- Signals as vectors, vector space geometry
- Signal representations and bases
- Linear systems analysis and linear algebra
- "Optimality" in signal processing (e.g., optimal filter design)

# Chapter 1

# Signal Representation and Approximation in Vector Spaces

## 1.1 Metric Spaces[1]

We will view signals as elements of certain mathematical spaces. The spaces have a common structure, so it will be useful to think of them in the abstract.

### 1.1.1 Metric Spaces

**Definition 1**
A *set* is a (possibly infinite) collection of distinct objects.

**Example 1.1**

- The empty set: $\varnothing = \{\}$ (plays a role akin to zero)
- Binary numbers: $\{0, 1\}$
- Natural numbers: $\mathbb{N} = \{1, 2, 3, ...\}$
- Integers: $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ ($Z$ is short for "Zahlen", German for "numbers")
- Rational numbers: $\mathbb{Q}$ ($Q$ for "quotient")
- Real numbers: $\mathbb{R}$
- Complex numbers: $\mathbb{C}$

In this course we will assume familiarity with a number of common set operations. In particular, for the sets $A = \{0, 1\}$, $B = \{1\}$, $C = \{2\}$, we have the operations of:

**Union:** $A \cup B = \{0, 1\}$, $B \cup C = \{1, 2\}$
**Intersection:** $A \cap B = \{1\}$, $B \cap C = \varnothing$
**Exclusion:** $A \setminus B = \{0\}$
**Complement:** $A^c = U \setminus A$, $A^c = \{2\}$
**Cartesian Product:** $A^2 = A \times A = \{(0,0), (0,1), (1,0), (1,1)\}$

In order to be useful a set must typically satisfy some additional structure. We begin by defining a notion of distance.
**Definition 2**
A *metric space* is a set $M$ together with a *metric* (distance function) $d : M \times M \to \mathbb{R}$ such that for all $x, y, z \in M$

[1] This content is available online at <http://cnx.org/content/m33568/1.3/>.

**M1.** $d(x,y) = d(y,x)$ (symmetry)
**M2.** $d(x,y) \geq 0$ (non-negative)
**M3.** $d(x,y) = 0$ iff $x = y$ (positive semi-definite)
**M4.** $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality).

**Example 1.2**

**Trivial metric:** ($M$ is arbitrary) $d(x,y) = \{ \begin{array}{l} 0 \text{ if } x = y, \\ 1 \text{ if } x \neq y. \end{array}$
**Standard metric:** ($M = \mathbb{R}$) $d(x,y) = |x - y|$
**Euclidean ($\ell_2$) metric:** ($M = \mathbb{R}^N$) $d(x,y) = \sqrt{\sum_{i=1}^{N} |x_i - y_i|^2}$
**$\ell_1$ metric:** ($M = \mathbb{R}^N$) $d(x,y) = \sum_{i=1}^{N} |x_i - y_i|$
**$\ell_p$ metric, $1 \leq p < \infty$:** ($M = \mathbb{R}^N$) $d(x,y) = \left(\sum_{i=1}^{N} |x_i - y_i|^p\right)^{1/p}$
**$\ell_\infty$ metric:** ($M = \mathbb{R}^N$) $d(x,y) = \max_{i=1,\dots,N} |x_i - y_i|$
**$L_p$ metric:** ($M$ = real (or complex) valued functions defined on $[a,b]$) $d_p(x,y) = \left(\int_a^b |x(t) - y(t)|^p dt\right)^{1/p}$

## 1.2 Completeness[2]

Distance functions allow us to talk concretely about limits and convergence of sequences.

**Definition 1**
Let $(M, d(x,y))$ be a metric space and $\{x_i\}_{i=1}^{\infty}$ be a sequence of elements in $M$. We say that $\{x_i\}_{i=1}^{\infty}$ *converges* to $x^*$ if and only if for every $\epsilon > 0$ there is an $N$ such that $d(x_i, x^*) < \epsilon$ for all $i > N$. In this case we say that $x^*$ is the *limit* of $\{x_i\}_{i=1}^{\infty}$.

**Figure 1.1:** A sequence of points $\{x_i\}$ converging to $x^*$.

**Definition 2**
A sequence $\{x_i\}_{i=1}^{\infty}$ is said to be a *Cauchy sequence* if for any $\epsilon > 0$ there is an $N$ such that $d(x_i, x_j) < \epsilon$ for every $i, j > N$.

It can be shown that any convergent sequence is a Cauchy sequence. However, it is possible for a Cauchy sequence to not be convergent!

[2] This content is available online at <http://cnx.org/content/m33586/1.2/>.

**Example 1.3**
Suppose that $M = (0, 2)$, i.e., the open interval from 0 to 2 on the real line, and let $d(x, y) = |x - y|$. Consider the sequence defined by $x_i = \frac{1}{i}$. $\{x_i\}$ is Cauchy since for any $\epsilon$ we can set $N$ such that $\frac{1}{N} < \frac{\epsilon}{2}$, so that $|x_i - x_j| \leq |x_i| + |x_j| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. However, $x_i \to 0$, but $0 \notin M$, i.e., the sequence converges to something that lives outside of our space.

**Example 1.4**
Suppose that $M = C[-1, 1]$ (the set of continuous functions defined on $[-1, 1]$) and let $d_2$ denote the $L_2$ metric. Consider the sequence of functions defined by

$$f_i(t) = \begin{cases} 0 & \text{if } t \leq -\frac{1}{i} \\ \frac{it}{2} + \frac{1}{2} & \text{if } -\frac{1}{i} < t < \frac{1}{i} \\ 1 & \text{if } t \geq \frac{1}{i}. \end{cases} \tag{1.1}$$

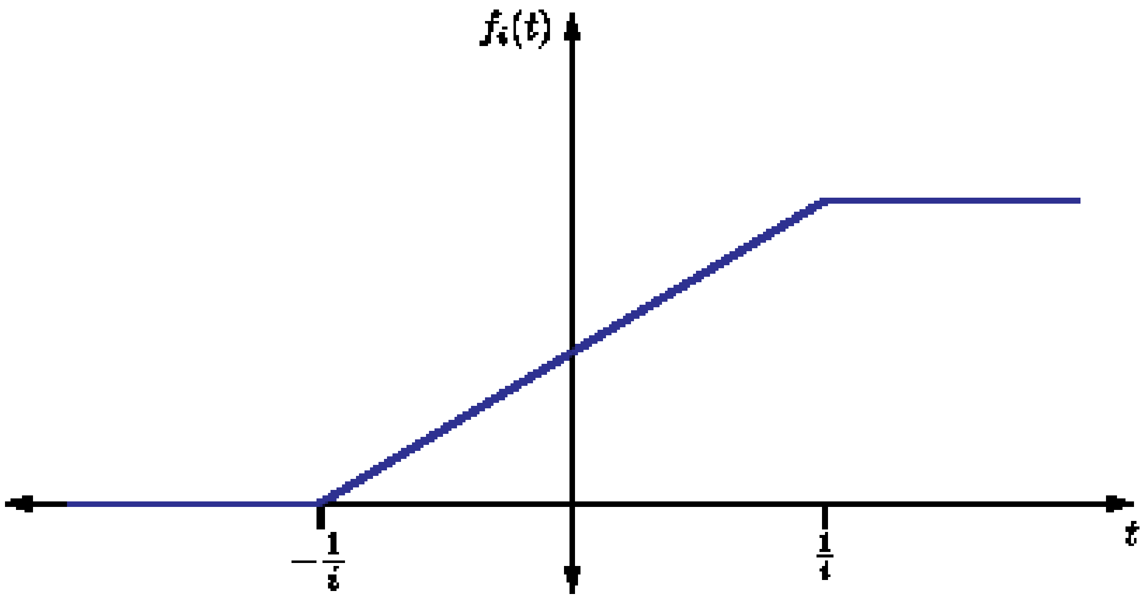



**Figure 1.2**

For $j > i$ we have that

$$d_2(f_i, f_j) = \frac{(j - i)^2}{6j^3 i}. \tag{1.2}$$

This goes to 0 for $j, i$ sufficiently large. Thus, the sequence $\{f_i\}_{i=1}^{\infty}$ is Cauchy, but it converges to a discontinuous function, and thus it is not convergent in $M$.

**Definition 3**
A metric space $(M, d(x, y))$ is *complete* if every Cauchy sequence in $M$ is convergent in $M$.

**Example 1.5**

- $M = [0,1]\,, d(x,y) = |x-y|$ is complete.
- $(C[-1,1]\,, d_2)$ is not complete, but one can check that $(C[-1,1]\,, d_\infty)$ is complete. (This space works because using $d_\infty$, the above example is no longer Cauchy.)
- $\mathbb{Q}$ is not complete, but $\mathbb{R}$ is.

## 1.3 Vector Spaces[3]

Metric spaces impose no requirements on the structure of the set $M$. We will now consider more structured $M$, beginning by generalizing the familiar concept of a vector.

**Definition 1**

Let $K$ be a field of *scalars*, i.e., $K = \mathbb{R}$ or $\mathbb{C}$. Let $V$ be a set of *vectors* equipped with two binary operations:

1. vector addition: $+ : V \times V \to V$
2. scalar multiplication: $\cdot : K \times V \to V$

We say that $V$ is a *vector space* (or *linear space*) over $K$ if

**VS1:** $V$ forms a group under addition, i.e.,

- $(x+y)+z = x+(y+z)$ (associativity)
- $x+y = y+x$ (commutativity)
- $\exists\ 0 \in V$ such that $\forall\ x \in V$, $x+0 = 0+x = x$
- $\forall x \in V$, $\exists\ y$ such that $x+y = 0$

**VS2:** For any $\alpha, \beta \in K$ and $x, y \in V$

- $\alpha(\beta x) = (\alpha\beta)x$ (compatibility)
- $(\alpha+\beta)(x+y) = \alpha x + \alpha y + \beta x + \beta y$ (distributivity)
- $\exists 1 \in K$ such that $1x = x$

**Example 1.6**

- $\mathbb{R}^N$ over $\mathbb{R}$ (not $\mathbb{R}^N$ over $\mathbb{C}$)
- $\mathbb{C}^N$ over $\mathbb{C}$ or $\mathbb{C}^N$ over $\mathbb{R}$
- Set of polynomials of degree $N$ with rational coefficients over $\mathbb{Q}$
- The set of all infinitely-long sequences of real numbers over $\mathbb{R}$
- $GF(2)^N : \{0,1\}^N$ over $\{0,1\}$ with mod 2 arithmetic (Galois field)
- $C[a,b]$ over $\mathbb{R}$

## 1.4 Normed Vector Spaces[4]

While vector spaces have additional structure compared to a metric space, a general vector space has no notion of “length” or “distance.”

**Definition 1**

Let $V$ be a vector space over $K$. A *norm* is a function $\|\cdot\| : V \to \mathbb{R}$ such that

**N1.** $\|x\| \geq 0 \forall x \in V$
**N2.** $\|x\| = 0$ iff $x = 0$

---

[3] This content is available online at <http://cnx.org/content/m33583/1.2/>.
[4] This content is available online at <http://cnx.org/content/m33561/1.2/>.

**N3.** $\|\alpha x\| = |\alpha|\|x\| \forall x \in V,\ \alpha \in K$
**N4.** $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in V$

A vector space together with a norm is called a *normed vector space* (or *normed linear space*).

**Example 1.7**

- $V = \mathbb{R}^N$: $\|x\|_2 = \sqrt{\sum_{i=1}^{N} |x_i|^2}$

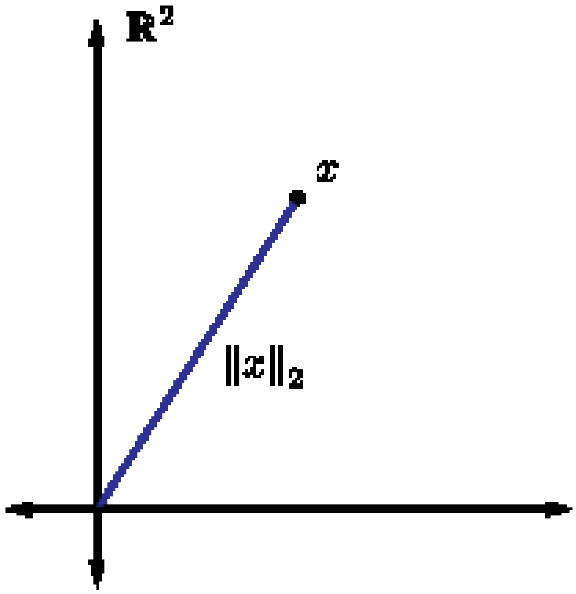



**Figure 1.3**

- $V = \mathbb{R}^N$: $\|x\|_1 = \sum_{i=1}^{N} |x_i|$ (“Taxicab”/“Manhattan” norm)

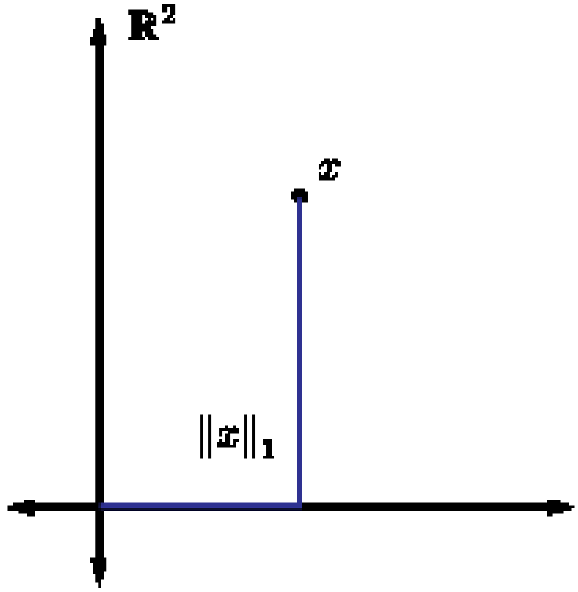



**Figure 1.4**

- $V = \mathbb{R}^N$: $\|x\|_\infty = \underset{i=1,\dots,N}{max} |x_i|$

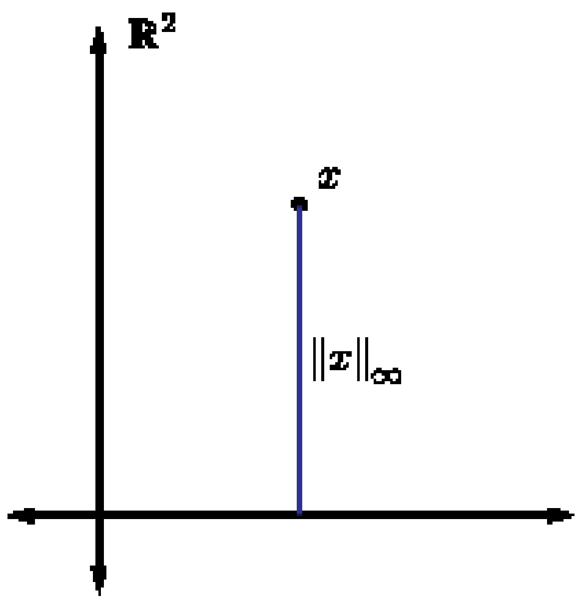



**Figure 1.5**

- $V = L_p[a,b]$, $p \in [1,\infty)$: $\|x(t)\|_p = \left(\int_a^b |x(t)|^p dt\right)^{1/p}$ (The notation $L_p[a,b]$ denotes the set of all functions defined on the interval $[a,b]$ such that this norm exists, i.e., $\| x(t) \|_p < \infty$.)

Note that any normed vector space is a metric space with induced metric $d(x,y) = \|x - y\|$. (This follows since $\|x - y\| = \|x - z + z - y\| \in \|x - z\| + \|y - z\|$.) While a normed vector space "feels like" a metric space, it is important to remember that it actually satisfies a great deal of additional structure.

**Technical Note:** In a normed vector space we must have (from N2) that $x = y$ if $\|x - y\| = 0$. This can lead to a curious phenomenon when dealing with continuous-time functions. For example, in $L_2([a,b])$, we can consider a pair of functions like $x(t)$ and $y(t)$ illustrated below. These functions differ only at a single point, and thus $\| x(t) - y(t) \|_2 = 0$ (since a single point cannot contribute anything to the value of the integral.) Thus, in order for our norm to be consistent with the axioms of a norm, we must say that $x = y$ whenever $x(t)$ and $y(t)$ differ only on a set of measure zero. To reiterate $x = y$[U+21CE]$x(t) = y(t) \, \forall t \in [a,b]$, i.e., when we treat functions as vectors, we will *not* interpret $x = y$ as pointwise equality, but rather as equality *almost everywhere*.

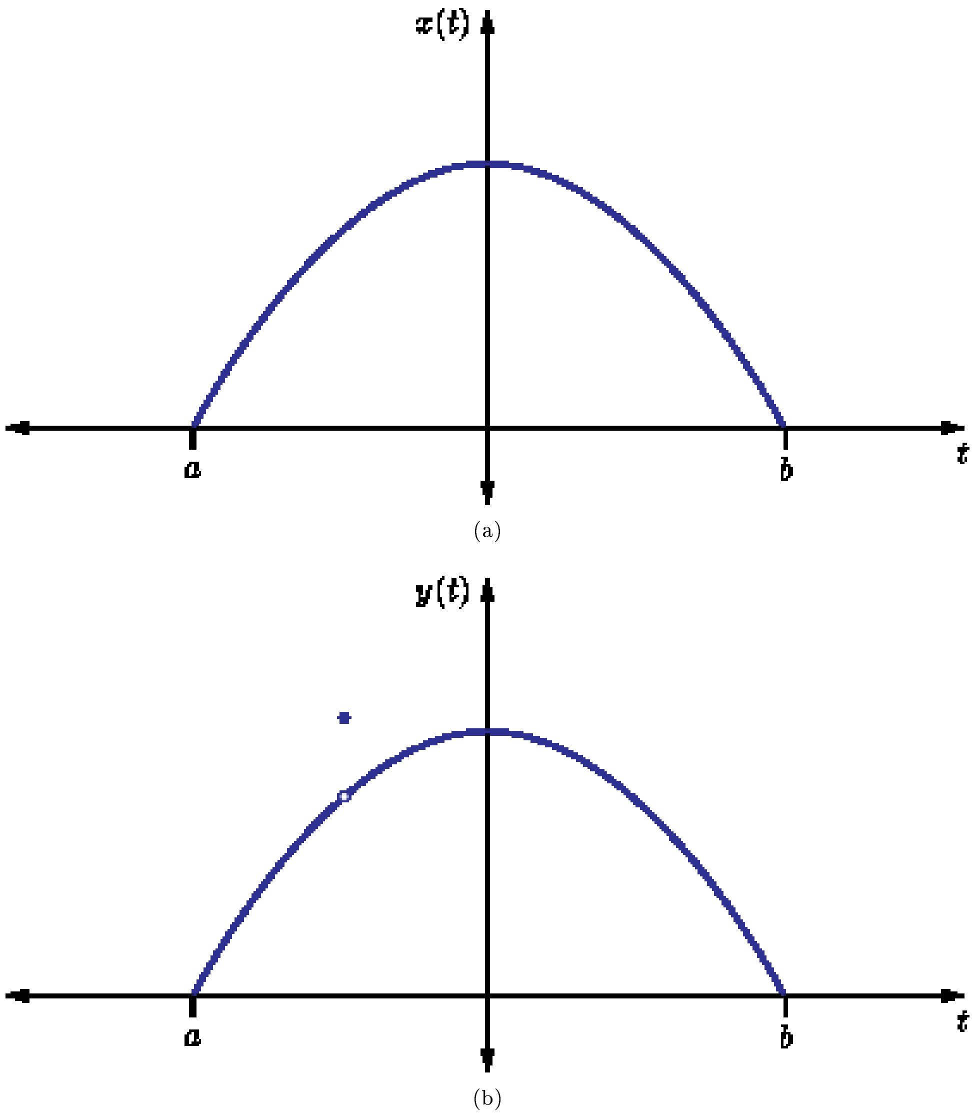



**Figure 1.6**

## 1.5 Inner Product Spaces[5]

Where normed vector spaces incorporate the concept of length into a vector space, inner product spaces incorporate the concept of angle.

**Definition 1**

Let $V$ be a vector space over $K$. An *inner product* is a function $< \cdot, \cdot >: V \times V \rightarrow K$ such that for all $x, y, z \in V, \alpha \in K$

[5]This content is available online at <http://cnx.org/content/m33556/1.2/>.

**IP1.** $< x, y >= \overline{< y, x >}$
**IP2.** $< \alpha x, y >= \alpha < x, y >$
**IP3.** $< x + y, z >=< x, z > + < y, z >$
**IP4.** $< x, x > \geq 0$ with equality iff $x = 0$.

A vector space together with an inner product is called an *inner product space*.

**Example 1.8**

- $V = \mathbb{C}^N$, $< x, y >:= \sum_{i=1}^{N} x_i \overline{y_i} = y^* x$
- $V = C[a, b]$, $< x, y >:= \int_a^b x(t) \overline{y(t)} dt$

Note that a valid inner product space induces a normed vector space with norm $\|x\| = \sqrt{< x, x >}$. (Proof relies on Cauchy-Schwartz inequality.) In $\mathbb{R}^N$ or $\mathbb{C}^N$, the standard inner product induces the $\ell_2$-norm. We summarize the relationships between the various spaces introduced over the last few lectures in Figure 1.7.

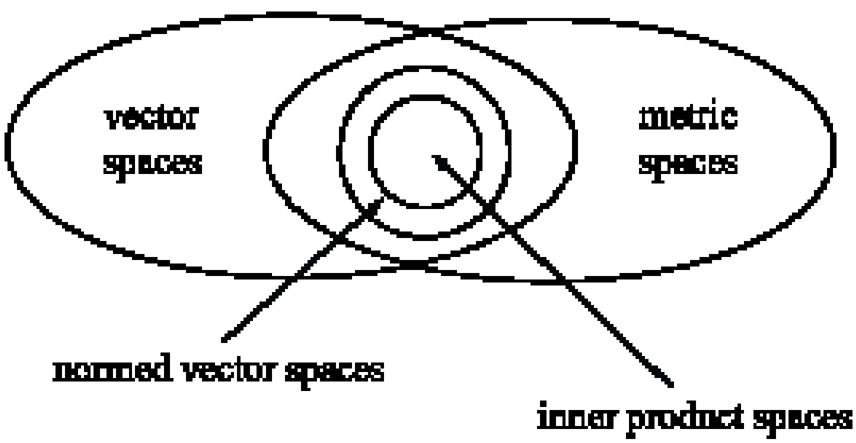



**Figure 1.7:** Venn diagram illustrating the relationship between vector and metric spaces.

## 1.6 Properties of Inner Products[6]

Inner products and their induced norms have some very useful properties:

**Cauchy-Schwartz Inequality:** $| < x, y > | \leq \|x\| \|y\|$ with equality iff $\exists\ \alpha \in \mathbb{C}$ such that $y = \alpha x$
**Pythagorean Theorem:** $< x, y >= 0 \Rightarrow \|x + y\|^2 = \|x - y\|^2 = \|x\|^2 + \|y\|^2$
**Parallelogram Law:** $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$
**Polarization Identity:** $\mathrm{Re}[< x, y >] = \frac{\|x+y\|^2 - \|x-y\|^2}{4}$

In $\mathbb{R}^2$ and $\mathbb{R}^3$, we are very familiar with the geometric notion of an angle between two vectors. For example, if $x, y \in \mathbb{R}^2$, then from the law of cosines, $< x, y >= \|x\| \|y\| cos\theta$. This relationship depends only on norms and inner products, so it can easily be extended to any inner product space.

[6] This content is available online at <http://cnx.org/content/m33573/1.2/>.

**Figure 1.8**

**Definition 1**
The *angle*$\theta$ between two vectors $x, y$ in an inner product space is defined by $cos\theta = \frac{<x,y>}{\|x\|\|y\|}$
**Definition 2**
Vectors $x, y$ in an inner product space are said to be *orthogonal* if $< x, y >= 0$.

## 1.7 Complete Vector Spaces[7]

**Definition1**
A complete normed vector space is called a *Banach space*.

**Example 1.9**

- $\cdot C[a,b]$ with $L_\infty$ norm, i.e., $\|f\|_\infty = \underset{t\in[a,b]}{\text{ess sup}} |f(t)|$ is a Banach space.
- $L_p[a,b] = \{f : \| f \|_p < \infty\}$ for $p \in [1,\infty]$ and $-\infty \leq a < b \leq \infty$ is a Banach space.
- $\ell_p(\mathbb{N}) = \{\text{sequences } x : \| x \|_p < \infty\}$ for $p \in [1,\infty]$ is a Banach space.
- Any finite-dimensional normed vector space is Banach, e.g., $\mathbb{R}^N$ or $\mathbb{C}^N$ with any norm.
- $C[a,b]$ with $L_p$ norm for $p < \infty$ is not Banach.

**Definition 2**
A complete inner product space is called a *Hilbert space*.

**Example 1.10**

- $L_2[a,b]$ is a Hilbert space.
- $\ell_2(\mathbb{N})$ is a Hilbert space.
- Any finite-dimensional inner product space is a Hilbert space.

Note that every Hilbert space is Banach, but the converse is not true. Hilbert spaces will be extremely important in this course.

## 1.8 Hilbert Spaces in Signal Processing[8]

What makes Hilbert spaces so useful in signal processing? In modern signal processing, we often represent a signal as a point in high-dimensional space. Hilbert spaces are spaces in which our geometry intuition from $\mathbb{R}^3$ is most trustworthy. As an example, we will consider the approximation problem.

[7] This content is available online at <http://cnx.org/content/m33548/1.2/>.
[8] This content is available online at <http://cnx.org/content/m34036/1.2/>.

**Definition 1.**
A subset $W$ of a vector space $V$ is *convex* if for all $x, y \in W$ and $\lambda \in (0,1)$, $\lambda x + (1-\lambda)\, y \in W$.

**Theorem 1.1:** The Fundamental Theorem of Approximation
Let $A$ be a nonempty, closed (complete), convex set in a Hilbert space $H$. For any $x \in H$ there is a unique point in $A$ that is closest to $x$, i.e., $x$ has a unique "best approximation" in $A$.

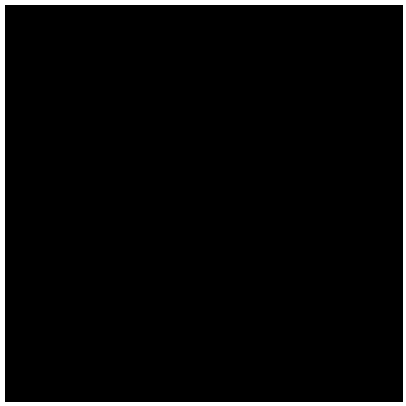


**Figure 1.9:** The best approximation to $x$ in convex set $A$.

Note that in non-Hilbert spaces, this may not be true! The proof is rather technical. See Young Chapter 3 or Moon and Stirling Chapter 2. Also known as the "closest point property", this is very useful in compression and denoising.

## 1.9 Linear Combinations of Vectors[9]

Suppose we have a set of vectors $v_1, v_2, ..., v_N$ that lie in a vector space $V$. Given scalars $\alpha_1, \alpha_2, ..., \alpha_N$, observe that the linear combination

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_N v_N \tag{1.3}$$

is also a vector in $V$.

**Definition 1**
Let $M \subset V$ be a set of vectors in $V$. The *span* of $M$, written $\mathrm{span}\,(M)$, is the set of all linear combinations of the vectors in $M$.

**Example 1.11:** $V = \mathbb{R}^3$

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{1.4}$$

[9]This content is available online at <http://cnx.org/content/m34030/1.2/>.

span $(\{v_1, v_2\})$ = the $x_1x_2$-plane, i.e., for any $x_1, x_2$ we can write $x_1 = \alpha_1$ and $x_2 = \alpha_1 + \alpha_2$ for some $\alpha_1, \alpha_2 \in \mathbb{R}$.

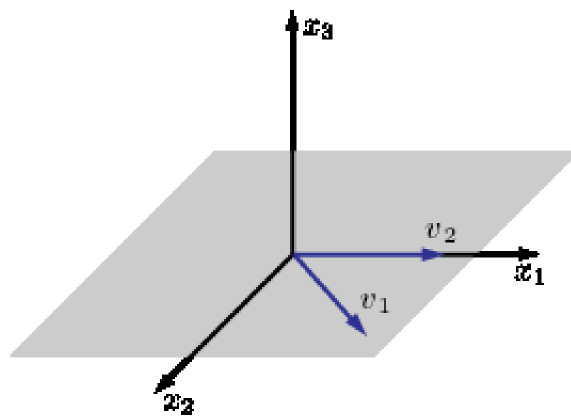



**Figure 1.10:** Illustration of the set of all linear combinations of $v_1$ and $v_2$, i.e., the $x_1x_2$-plane.

**Example 1.12**

$V = \{f : f(t) \text{ is periodic with period } 2\pi\}$, $M = \{e^{jkt}\}_{k=-B}^{B}$

span $(M)$ = periodic, bandlimited (to $B$) functions, i.e., $f(t)$ such that $f(t) = \sum_{k=-B}^{B} c_K e^{jkt}$ for some $c_{-B}, c_{-B+1}, ..., c_0, c_1, ..., c_B \in \mathbb{C}$.

## 1.10 Vector Subspaces[10]

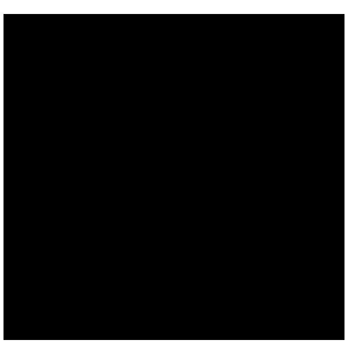
**Defintition 1**

A (non-empty) subset $W$ of $V$ is called a *supspace* of $V$ if for any $x, y \in W$, span $(\{x, y\}) \subseteq W$.

Note that this definition easily implies that:

- $0 \in W$
- $W$ is itself a vector space

**Example 1.13: Which of these are subspaces?**

-  [No]

[10]This content is available online at <http://cnx.org/content/m34073/1.2/>.

-  [Yes]
- $V = \mathbb{R}^5$, $W = \{x : x_4 = 0, x_5 = 0\}$ [Yes]
- $V = \mathbb{R}^5$, $W = \{x : x_4 = 1, x_5 = 1\}$ [No]
- $V = C[0,1]$, $W =$ {polynomials of degree $N$} [Yes]
- $V = C[0,1]$, $W = \{f : f$ is bandlimited to $B\}$ [Yes]
- $V = \mathbb{R}^N$, $W = \{x : x \text{ hasnomorethan5nonzerocomponents}, i.e., \|x\|_0 \leq 5\}$ [No]

## 1.11 Signal Approximation in a Hilbert Space[11]

We will now revisit "The Fundamental Theorem of Approximation" for the extremely important case where our set $A$ is a subspace. Specifically, suppose that $H$ is a Hilbert space, and let $A$ be a (closed) subspace of $H$. From before, we have that for any $x \in H$ there is a unique $\hat{x} \in A$ such that $\hat{x}$ is the closest point in $A$ to $x$. When $A$ is also a subspace, we also have:

**Theorem 1.2:** The Orthogonality Principle

$\hat{x} \in A$ is the minimizer of $\|x - \hat{x}\|$ if any only if $\hat{x} - x \perp A$ i.e., $< \hat{x} - x, y >= 0$ for all $y \in A$.

**Proof:**

(a) Suppose that $\hat{x} - x \perp A$. Then for any $y \in A$ with $y \neq \hat{x}$, $\|y - x\|^2 = \|y - \hat{x} + \hat{x} - x\|^2$. Note that $y - \hat{x} \in A$, but $\hat{x} - x \perp A$, so that $< y - \hat{x}, \hat{x} - x >= 0$, and we can apply Pythagoras to obtain $\|y - x\|^2 = \|y - \hat{x}\|^2 + \|\hat{x} - x\|$. Since $y \neq \hat{x}$, we thus have that $\|y - x\|^2 > \|\hat{x} - x\|^2$. Thus $\hat{x}$ must be the closest point in $A$ to $x$.

[11] This content is available online at <http://cnx.org/content/m34068/1.2/>.

**Figure 1.11:** Illustration of the orthogonality principle.

(b) Suppose that $\hat{x}$ minimizes $\|x-\hat{x}\|$. Suppose for the sake of a contradiction that $\exists y \in A$ such that $\|y\| = 1$ and $< x-\hat{x}, y >= \delta \neq 0$.

Let $z = \hat{x} + \delta y$.

$$\begin{aligned} \|x - z\|^2 &= \|x-\hat{x} - \delta y\|^2 \\ &= < x-\hat{x}, x-\hat{x} > - < x-\hat{x}, \delta y > - < \delta y, x-\hat{x} > + < \delta y, \delta y > \\ &= \|x-\hat{x}\|^2 - \overline{\delta}\delta - \delta\overline{\delta} + \delta\overline{\delta} \\ &= \|x-\hat{x}\|^2 - |\delta|^2. \end{aligned} \quad (1.5)$$

Thus $\|x - z\| \leq \|x-\hat{x}\|$, contradicting the assumption that $\hat{x}$ minimizes $\|x-\hat{x}\|$.

This result suggests a that a possible method for finding the best approximation to a signal $x$ from a vector space $V$ is to simply look for a vector $\hat{x}$ such that $\hat{x} - x \perp V$. In the coming lectures we will show how to do this, but it will require a brief review of some concepts from linear algebra.

## 1.12 Linear Operators[12]

**Definition 1**
A transformation (mapping) $L : X \rightarrow Y$ from a vector space $X$ to a vector space $Y$ (with the same scalar field $K$) is a *linear* transformation if:

1. $L(\alpha x) = \alpha L(x) \forall x \in X,\ \alpha \in K$
2. $L(x_1 + x_2) = L(x_1) + L(x_2) \forall x_1, x_2 \in X$.

[12]This content is available online at <http://cnx.org/content/m34058/1.2/>.

We call such transformations *linear operators*.

**Example 1.14**

- $X = \mathbb{R}^N$, $Y = \mathbb{R}^M L : \mathbb{R}^N \to \mathbb{R}^M$ is an $M \times N$ matrix
- Fourier transform: $F(x(t)) = \int_{-\infty}^{\infty} x(t) e^{-jwt} dt F : L_2(\mathbb{R}) \to L_2(\mathbb{R})$

Let $L : X \to Y$ be an operator (linear or otherwise). The *range space* $\mathcal{R}(L)$ is

$$\mathcal{R}(L) = \{L(x) \in Y : x \in X\}. \tag{1.6}$$

The *null space* $\mathcal{N}(L)$, also known as "kernel", is

$$\mathcal{N}(L) = \{x \in X : L(x) = 0\}. \tag{1.7}$$

If $L$ is linear, then both $\mathcal{R}(L)$ and $\mathcal{N}(L)$ are subspaces.

## 1.13 Projections[13]

**Definition 1**
A linear transformation $P : X \to X$ is called a *projection* if $P(x) = x \forall x \in \mathcal{R}(P)$, i.e, $P(P(x)) = P(x) \forall x \in X$.

**Example 1.15**
$P : \mathbb{R}^3 \to \mathbb{R}^3$, $P(x_1, x_2, x_3) = (x_1, x_2, 0)$

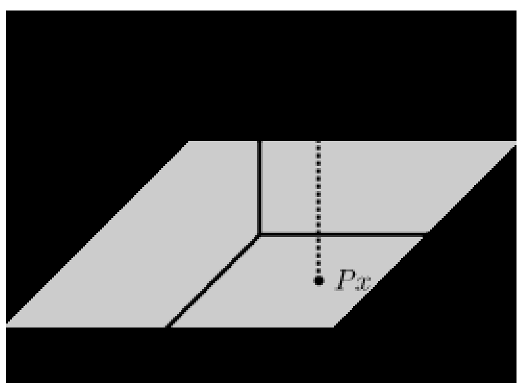



**Figure 1.12**

**Definition 2**
If $P$ is a projection operator on an inner product space $V$, we say that $P$ is an *orthogonal projection* if $\mathcal{R}(P) \perp \mathcal{N}(P)$ , i.e., $< x, y >= 0 \forall x \in \mathcal{R}(P), y \in \mathcal{N}(P)$ .

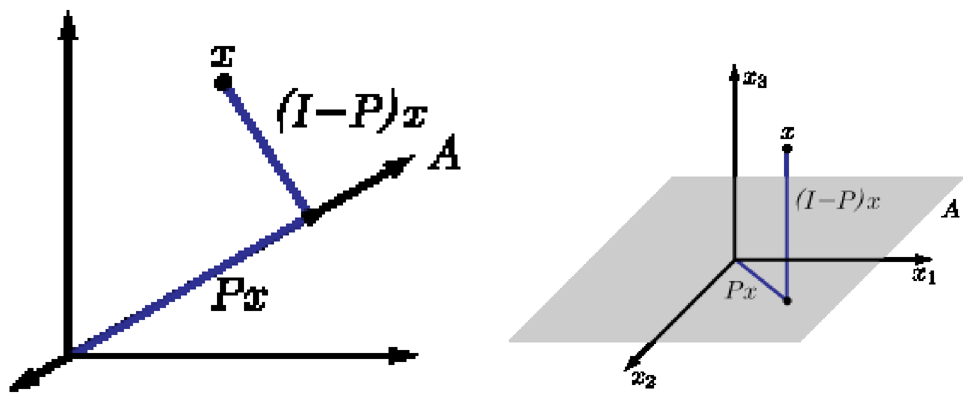
If $P$ is an orthogonal projection, then for any $x \in V$ we can write:

$$x = Px + (I - P) x \tag{1.8}$$

where $Px \in \mathcal{R}(P)$ and $(I - P) x \in \mathcal{N}(P)$ (since $P(I - P) x = Px - P(Px) = Px - Px = 0$.)

[13] This content is available online at <http://cnx.org/content/m34064/1.2/>.

Now we see that the solution to our "best approximation in a linear subspace" problem is an orthogonal projection: we wish to find a $P$ such that $R(P) = A$.




(a) (b)

**Figure 1.13**

The question is now, how can we design such a projection operator?

## 1.14 Linear Independence[14]

**Definition 1**
A set of vectors $\{v_j\}_{j=1}^{N}$ is said to be *linearly dependent* is there exists a set of scalars $\alpha_1, ..., \alpha_N$ (not all 0) such that

$$\sum_{j=1}^{N} \alpha_j v_j = 0. \tag{1.9}$$

Likewise if $\sum_{j=1}^{N} \alpha_j v_j = 0$ only when $\alpha_j = 0 \forall j$, then $\{v_j\}_{j=1}^{N}$ is said to be *linearly independent*.

**Example 1.16:** $V = \mathbb{R}^3$

$$v_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}. \tag{1.10}$$

Find $\alpha_1, \alpha_2, \alpha_3$ such that $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$. $[\alpha_1 = 1, \alpha_2 = -3, \alpha_3 = 1.]$ Note that any two vectors are linearly independent.

Note that if a set of vectors $\{v_j\}_{j=1}^{N}$ are linearly dependent then we can remove vectors from the set without changing the span of the set.

[14]This content is available online at <http://cnx.org/content/m34042/1.2/>.

## 1.15 Bases[15]

**Definition 1**
A *basis* of a vector space $V$ is a set of vectors $B$ such that

- $\text{span}(B) = V$.
- $B$ is linearly independent.

The second condition ensures that all bases of $V$ will have the same size. In fact, the dimension of a vector space $V$ is defined as the number of elements required in a basis for $V$. (Could easily be in infinite.)

**Example 1.17**

- $\mathbb{R}^N$ with $B$ the "standard basis" for $\mathbb{R}^N$

$$\{b_1, b_2, ..., b_N\} = \{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}\} \tag{1.11}$$

  Note that this easily extends to $\ell_p(\mathbb{Z})$.
- $\mathbb{R}^N$ with any set of $N$ linearly independent vectors
- $V = \{\text{polynomialsofdegreeatmost } p\} B = \{1, t, t^2, ..., t^p\}$ (Note that the dimension of $V$ is $p+1$)
- $V = \{f(t) : f(t) \text{ isperiodicwithperiod } T\} B = \{e^{jkt}\}_{k=-\infty}^{\infty}$ (Fourier series, infinite dimensional)

## 1.16 Orthogonal Bases[16]

**Definition 1**
A collection of vectors $B$ in an inner product space $V$ is called an *orthogonal basis* if

1. $\text{span}(B) = V$
2. $v_i \perp v_j$ (i.e., $< v_i, v_j >= 0$) $\forall\ i \neq j$

If, in addition, the vectors are normalized under the induced norm, i.e., $\| v_i \| = 1\ \forall\ i$ , then we call $V$ an *orthonormal basis* (or "*orthobasis*"). If $V$ is infinite dimensional, we need to be a bit more careful with 1. Specifically, we really only need the closure of $\text{span}(B)$ to equal $V$. In this case any $x \in V$ can be written as

$$x = \sum_{i=1}^{\infty} c_i v_i \tag{1.12}$$

for some sequence of coefficients $\{c_i\}_{i=1}^{\infty}$.

(This last point is a technical one since the span is typically defined as the set of linear combinations of a finite number of vectors. See Young Ch 3 and 4 for the details. This won't affect too much so we will gloss over the details.)

**Example 1.18**

[15] This content is available online at <http://cnx.org/content/m34017/1.2/>.
[16] This content is available online at <http://cnx.org/content/m34067/1.2/>.

- $V = \mathbb{R}^2$, standard basis

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{1.13}$$

**Example 1.19**

- Suppose $V = \{$ piecewise constant functions on $\left[0, \frac{1}{4}\right), \left[\frac{1}{4}, \frac{1}{2}\right), \left[\frac{1}{2}, \frac{3}{4}\right), \left[\frac{3}{4}, 1\right]\}$. An example of such a function is illustrated below.

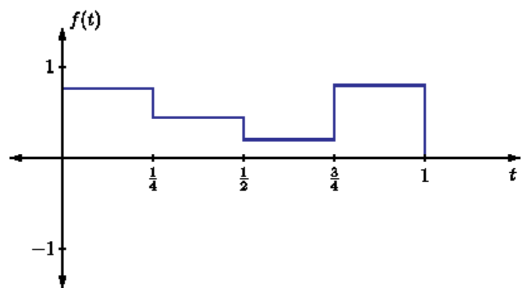



**Figure 1.14**

Consider the set

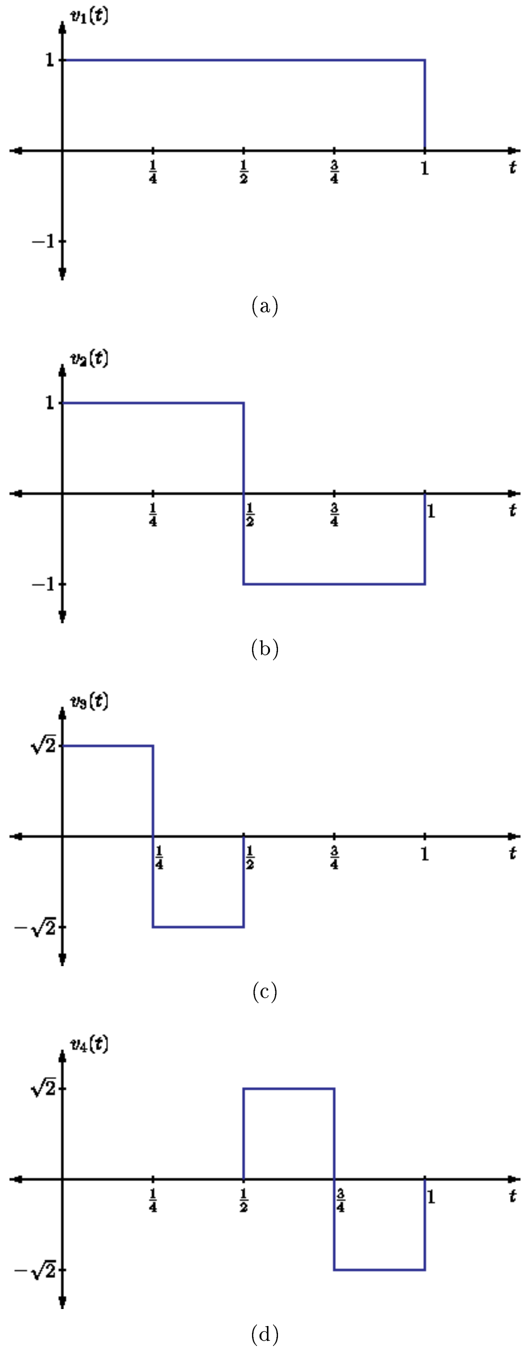



**Figure 1.15**

The vectors $\{v_1, v_2, v_3, v_4\}$ form an orthobasis for $V$.

- Suppose $V = L_2[-\pi, \pi]$. $B = \{\frac{1}{\sqrt{2\pi}} e^{jkt}\}_{k=-\infty}^{\infty}$, i.e, the Fourier series basis vectors, form an orthobasis for $V$. To verify the orthogonality of the vectors, note that:

$$\begin{aligned} < \frac{1}{\sqrt{2\pi}} e^{jkt}, \frac{1}{\sqrt{2\pi}} e^{jkt} > \quad & = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(k_1-k_2)t} \\ & = \frac{1}{2\pi} \frac{e^{j(k_1-k_2)t}}{j(k_1-k_2)} \Big|_{-\pi}^{\pi} \\ & = \frac{1}{2\pi} \cdot \frac{-1+1}{j(k_1-k_2)} = 0 \ \ (k_1 \neq k_2) \end{aligned} \tag{1.14}$$

See Young for proof that the closure of $B$ is $L_2[-\pi,\pi]$, i.e., the fact that *any* $f \in L_2[-\pi,\pi]$ has a Fourier series representation.

## 1.17 Computing the Best Approximation[17]

Recall that if $P$ is an orthogonal projection onto a subspace $A$, we can write any $x$ as

$$x = Px + (I-P)x \tag{1.15}$$

where $Px \in A$ and $(I-P)x \perp A$. We now turn to how to actually find $P$.

We begin with the finite-dimensional case, assuming that $\{v_1, ..., v_N\}$ is a basis for $A$. If $(I-P)x \perp A$ then we have that for any $x$

$$< (I-P)x, v_j >= 0 \quad \text{for } j = 1, ..., N \tag{1.16}$$

We also note that since $Px \in A$, we can write $Px = \sum_{k=1}^{N} c_k v_k$. Thus we obtain

$$< x - \sum_{k=1}^{N} c_k v_k, v_j >= 0 \quad \text{for } j = 1, ..., N \tag{1.17}$$

from which we obtain

$$< x, v_j >= \sum_{k=1}^{N} c_k < v_k, v_j > \quad \text{for } j = 1, ..., N \tag{1.18}$$

We know $x$ and $v_1, ..., v_N$. Our goal is to find $c_1, ..., c_N$. Note that a procedure for calculating $c_1, ..., c_k$ for any given $x$ is equivalent to one that computes $Px$.

To find $c_1, ..., c_N$, observe that (1.18) represents a set of $N$ equations with $N$ unknowns.

$$\begin{bmatrix} < v_1, v_1 > & < v_2, v_1 > & \cdots & < v_N, v_1 > \\ < v_1, v_2 > & < v_2, v_2 > & & < v_N, v_2 > \\ \vdots & & \ddots & \vdots \\ < v_1, v_N > & < v_2, v_N > & \cdots & < v_N, v_N > \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} < x, v_1 > \\ < x, v_2 > \\ \vdots \\ < x, v_N > \end{bmatrix} \tag{1.19}$$

More compactly, we want to find a vector $c \in \mathbb{C}^N$ such that $Gc = b$ where

$$b = \begin{bmatrix} < x, v_1 > \\ < x, v_2 > \\ \vdots \\ < x, v_N > \end{bmatrix} \tag{1.20}$$

[17]This content is available online at <http://cnx.org/content/m34021/1.2/>.

NOTE:

- $G$ is called the "Grammian" or "Gram matrix" of $\{v_j\}$
- One can show since $v_1, ..., v_N$ are linearly independent that $G$ is positive definite, and hence inevitable.
- Also note that by construction, $G$ is *conjugate symmetric*, or *"Hermitian"*, i.e., $G = G^H$, where $^H$ denotes the conjugate transpose of $G$.

Thus, since $G^{-1}$ exists, we can write $c = G^{-1}b$ to calculate $c$.

As a special case, suppose now that $\{v_j\}$ is an orthobasis for $A$? What is $G$? It is just the identity matrix $I$! Computing $c$ just got much easier, since now $c = b$. Plugging this $c$ back into out formula for $Px$ we obtain

$$Px = \sum_{k=1}^{N} < x, v_k > v_k \tag{1.21}$$

Just to verify, note that $P$ is indeed a projection matrix:

$$\begin{aligned} P(Px) &= \sum_{k=1}^{N} < \sum_{j=1}^{N} < x, v_j > v_j, v_k > v_k \\ &= \sum_{k=1}^{N} \sum_{j=1}^{N} < x, v_j >< v_j, v_k > v_k \\ &= \sum_{j=1}^{N} < x, v_j > v_j = Px. \end{aligned} \tag{1.22}$$

**Example** Suppose $f \in L_2([0,4])$ is given by

**Example 1.20**

Suppose $f \in L_2([0,4])$ is given by

$$f(t) = \{ \begin{array}{cc} t & \text{if } t \in \left[0, \frac{1}{2}\right] \\ 1-t & \text{if } t \in \left[\frac{1}{2}, 1\right]. \end{array} \tag{1.23}$$

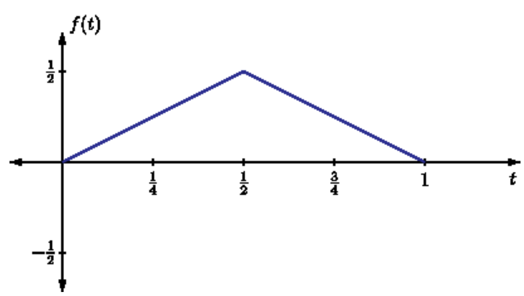



**Figure 1.16**

Let $A = \{$ piecewise constant functions on $\left[0, \frac{1}{4}\right), \left[\frac{1}{4}, \frac{1}{2}\right), \left[\frac{1}{2}, \frac{3}{4}\right), \left[\frac{3}{4}, 1\right]\}$. Our goal is to find the closest (in $L_2$) function in $A$ to $f(t)$. Using $v_1, ..., v_4$ from before, we can calculate $c_1 = \frac{1}{4}$, $c_2 = 0$, $c_3 = \frac{-\sqrt{2}}{16}$, $c_4 = \frac{\sqrt{2}}{16}$. Thus, we have that

$$\hat{f}(t) = \frac{1}{4} v_1 - \frac{\sqrt{2}}{16} v_3 + \frac{\sqrt{2}}{16} v_4. \tag{1.24}$$

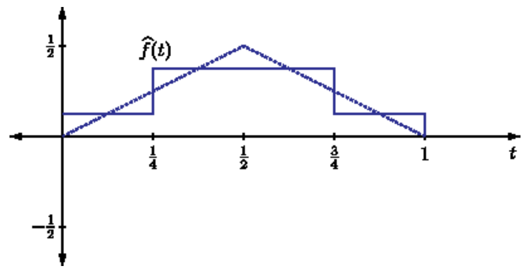



**Figure 1.17**

## 1.18 Matrix Representation of the Approximation Problem[18]

Suppose our inner product space $V = \mathbb{R}^M$ or $\mathbb{C}^M$ with the standard inner product (which induces the $\ell_2$-norm).

Re-examining what we have just derived, we can write our approximation $\hat{x} = Px = Vc$, where $V$ is an $M \times N$ matrix given by

$$V = \begin{bmatrix} \vdots & \vdots & & \vdots \\ v_1 & v_2 & \cdots & v_N \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{1.25}$$

and $c$ is an $N \times 1$ vector given by

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}. \tag{1.26}$$

Given $x \in \mathbb{R}^M$ (or $\mathbb{C}^M$), our search for the closest approximation can be written as

$$\underset{c}{min} \|x - V_c\|_2 \tag{1.27}$$

or as

$$\underset{c,e}{min} \|e\|_2^2 \quad \text{subjectto} \quad x = Vc + e \tag{1.28}$$

Using $V$, we can replace $G = V^H V$ and $b = V^H x$. Thus, our solution can be written as

$$c = \left(V^H V\right)^{-1} V^H x, \tag{1.29}$$

[18]This content is available online at <http://cnx.org/content/m34050/1.2/>.

which yields the formula

$$\hat{x} = V\left(V^H V\right)^{-1} V^H x. \tag{1.30}$$

The matrix $V^{\dagger} = \left(V^H V\right)^{-1} V^H$ is known as the "pseudo-inverse." Why the name "pseudo-inverse"? Observe that

$$V^{\dagger} V = \left(V^H V\right)^{-1} V^H V = I. \tag{1.31}$$

Note that $\hat{x} = V V^{\dagger} x$. We can verify that $V V^{\dagger}$ is a projection matrix since

$$\begin{aligned} V V^{\dagger} V V^{\dagger} &= V\left(V^H V\right)^{-1} V^H V \left(V^H V\right)^{-1} V^H \\ &= V\left(V^H V\right)^{-1} V^H \\ &= V V^{\dagger} \end{aligned} \tag{1.32}$$

Thus, given a set of $N$ linearly independent vectors in $\mathbb{R}^M$ or $\mathbb{C}^M$ $(N < M)$, we can use the pseudo-inverse to project any vector onto the subspace defined by those vectors. This can be useful any time we have a problem of the form:

$$x = Vc + e \tag{1.33}$$

where $x$ denotes a set of known "observations", $V$ is a set of known "expansion vectors", $c$ are the unknown coefficients, and $e$ represents an unknown "noise" vector. In this case, the least-squares estimate is given by

$$c = V^{\dagger} x, \qquad \hat{x} = V V^{\dagger} x. \tag{1.34}$$

## 1.19 Orthobasis Expansions[19]

Suppose that the $\{v_j\}_{j=1}^N$ are a finite-dimensional orthobasis. In this case we have

$$\hat{x} = \sum_{j=1}^{N} < x, v_j > v_j. \tag{1.35}$$

But what if $x \in \text{span}\left(\{v_j\}\right) = V$ already? Then we simply have

$$x = \sum_{j=1}^{N} < x, v_j > v_j \tag{1.36}$$

for all $x \in V$. This is often called the "reproducing formula". In infinite dimensions, if $V$ has an orthobasis $\{v_j\}_{j=1}^{\infty}$ and $x \in V$ has

$$\sum_{j=1}^{\infty} |< x, v_j >|^2 < \infty \tag{1.37}$$

then we can write

$$x = \sum_{j=1}^{\infty} < x, v_j > v_j. \tag{1.38}$$

[19] This content is available online at <http://cnx.org/content/m34055/1.2/>.

In other words, $x$ is perfectly captured by the list of numbers $< x, v_1 >, < x, v_2 >, ...$

Sound familiar?

**Example 1.21**

- $V = \mathbb{C}^n$, $\{v_k\}$ is the standard basis.

$$x_k = < x, v_k > v_k. \tag{1.39}$$

- $V = L_2[-\pi, \pi]$, $v_k(t) = \frac{1}{\sqrt{2\pi}} e^{jkt}$ For any $f \in V$ we have

$$f(t) = \sum_{k=-\infty}^{\infty} c_x v_x \tag{1.40}$$

where

$$c_k = < f, v_k > = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-jkt} dt. \tag{1.41}$$

The general lesson is that we can recreate a vector $x$ in an inner product space from the coefficients $\{< x, v_k >\}$. We can think of $\{< x, v_k >\}$ as "transform coefficients."

## 1.20 Parseval's and Plancherel's Theorems[20]

When dealing with transform coefficients, we will see that our notions of *distance* and *angle* carry over to the coefficient space.

Let $x, y \in V$ and suppose that $\{v_k\}_{k\in\Gamma}$ is an orthobasis. ($\Gamma$ denotes the index set, which could be finite or infinite.) Then $x = \sum_{k\in\Gamma} \alpha_k v_k$ and $y = \sum_{k\in\Gamma} \beta_k v_k$, and

$$< x, y >_V = \sum_{k\in\Gamma} \alpha_k \overline{\beta_k}. \tag{1.42}$$

So

$$< x, y >_V = < \alpha, \beta >_{\ell_2} \tag{1.43}$$

This is Plancherel's theorem. Parseval's theorem follows since $< x, x >_V = < \alpha, \alpha >_{\ell_2}$ which implies that $\|x\|_V^2 = \|x\|_{\ell_2}^2$. Thus, an orthobasis makes every inner product space equivalent to $\ell_2$!

## 1.21 Error of the Best Approximation in an Orthobasis[21]

As an application of Parseval's Theorem, say $\{v_k\}_{k=1}^{\infty}$ is an orthobasis for an inner product space of $V$.

Let $A$ be the subspace spanned by the first 10 elements of $\{v_k\}$, i.e., $A = \text{span}(\{v_1, ..., v_{10}\})$

1. Given $x \in v$, what is the closest point in $A$ (call it $\hat{x}$) to $x$? We have seen that it is $\hat{x} = \sum_{k=1}^{10} < x, v_k > v_k$
2. How good of an approximation is $\hat{x}$ to $x$? Measured with $\| \cdot \|_V$:

$$\begin{aligned} \| x - \hat{x} \|_V^2 &= \| \sum_{k>10} < x, v_k > v_k \|_V^2 \\ &= \sum_{k>10} | < x, v_k > |^2 \end{aligned} \tag{1.44}$$

[20] This content is available online at <http://cnx.org/content/m34062/1.2/>.
[21] This content is available online at <http://cnx.org/content/m34038/1.3/>.

Since we also have that $\| x \|_V^2 = \sum_{k=1}^{\infty} |< x, v_k >|^2$, the approximation $\hat{x}$ will be "good" if the first 10 transform coefficients contain "most" of the total energy. Constructing these types of approximations is exactly what is done in image compression.

## 1.22 Approximation in $\ell$_p Norms[22]

So far, our approximation problem has been posed in an inner product space, and we have thus measured our approximation error using norms that are induced by an inner product such as the $L_2/\ell_2$ norms (or weighted $L_2/\ell_2$ norms). Sometimes this is a natural choice – it can be interpreted as the "energy" in the error and arises often in the case of signals corrupted by Gaussian noise. However, more often than not, it is used simply because it is easy to deal with.

In some cases we might be interested in approximating with respect to other norms – in particular we will consider approximation with respect to $\ell_p$-norms for $p \neq 0$. First, we introduce the concept of a "unit ball". Any norm gives us rise to a unit ball, i.e., $\{x : \|x\| = 1\}$. Some important examples of unit balls for the $\ell_p$ norms in $\mathbb{R}^2$ are depicted below.

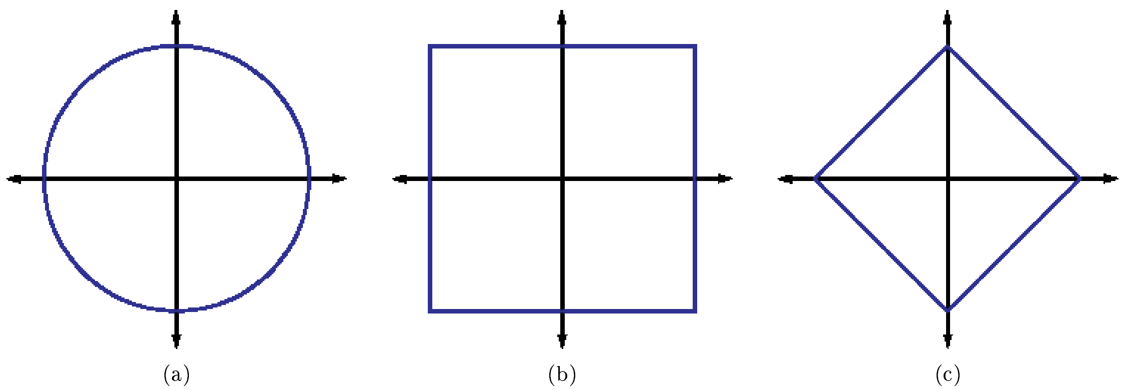



**Figure 1.18**

We now consider an example of approximating a point in $\mathbb{R}^2$ with a point in a 1-D subspace while measuring error using the $\ell_p$ norm for $p = 1, 2, \infty$.

**Example 1.22**

Suppose $V = \mathbb{R}^2$,

$$A = \text{span}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right), \qquad \text{and} \qquad x = \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{1.45}$$

We will want to find $\hat{x} \in A$ that minimizes $\|x - \hat{x}\|_p$. Since $\hat{x} \in A$, we can write

$$\hat{x} = \begin{bmatrix} 2\alpha \\ -\alpha \end{bmatrix} \tag{1.46}$$

[22] This content is available online at <http://cnx.org/content/m34019/1.3/>.

and thus

$$e = x - \hat{x} = \begin{bmatrix} 2 - 2\alpha \\ 1 + \alpha \end{bmatrix}. \tag{1.47}$$

While we can solve for $\alpha \in \mathbb{R}$ to minimize $\|e\|_p$ directly in some cases, a geometric interpretation is also useful. In each case, on can imagine growing an $\ell_p$ ball centered on $x$ until the ball intersects with $A$. This will be the point $\hat{x} \in A$. that is closest to $x$ in the $\ell_p$ norm. We first illustrate this for the $\ell_2$ norm below:

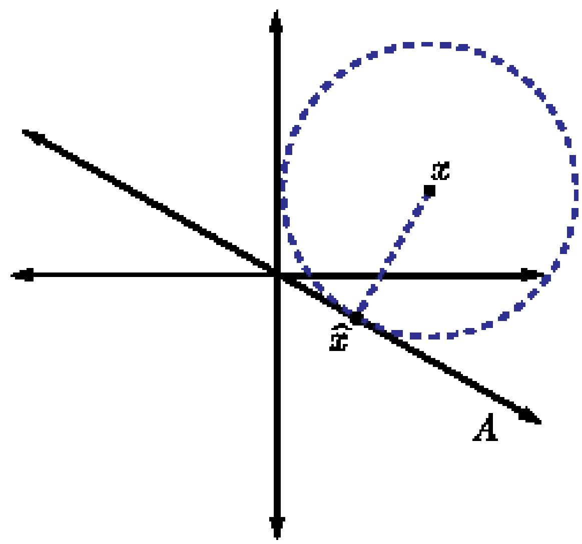



**Figure 1.19**

In order to calculate $\hat{x}$ we can apply the orthogonality principle. Since $< e, [2\ 1]^T >= 0$ we obtain a solution defined by $\alpha = \frac{3}{5}$.

We now observe that in the case of the $\ell_\infty$ norm the picture changes somewhat. The closest point in $\ell_\infty$ is illustrated below:

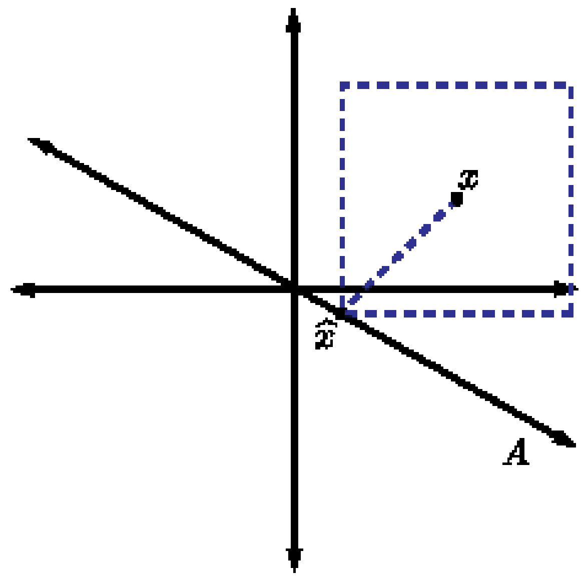



**Figure 1.20**

Note that the error is no longer orthogonal to the subspace $A$. In this case we can still calculate $\hat{x}$ from the observation that the two terms in the error should be equal, which yields $\alpha = \frac{1}{3}$.

The situation is even more different for the case of the $\ell_1$ norm, which is illustrated below:

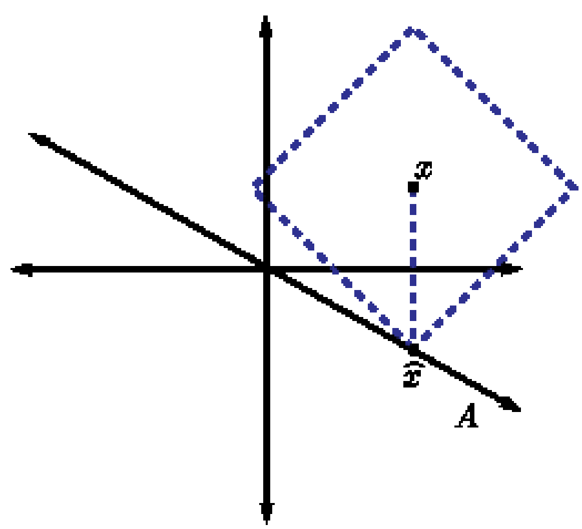



**Figure 1.21**

We now observe that $\hat{x}$ corresponds to $\alpha = 1$. Note that in this case the error term is $[0\ 2]^T$. This punctuates a general trend: for large values of $p$, the $\ell_p$ norm tends to spread error evenly across all terms, while for small values of $p$ the error is more highly concentrated.

When is it useful to approximate in $\ell_p$ or $L_p$ norms for $p \neq 0$?

**Example 1.23**

**Filter Design:** In some cases we will want the best fit to a specified frequency response in an $L_\infty$ sense rather than the $L_2$ sense. This minimizes the *maximum* error rather than total energy in the error. In the figure below we illustrate a desired frequency response. If the $L_\infty$ norm of the error is small, then we are guaranteed that the approximation to our desired frequency response will lie within the illustrated bounds.

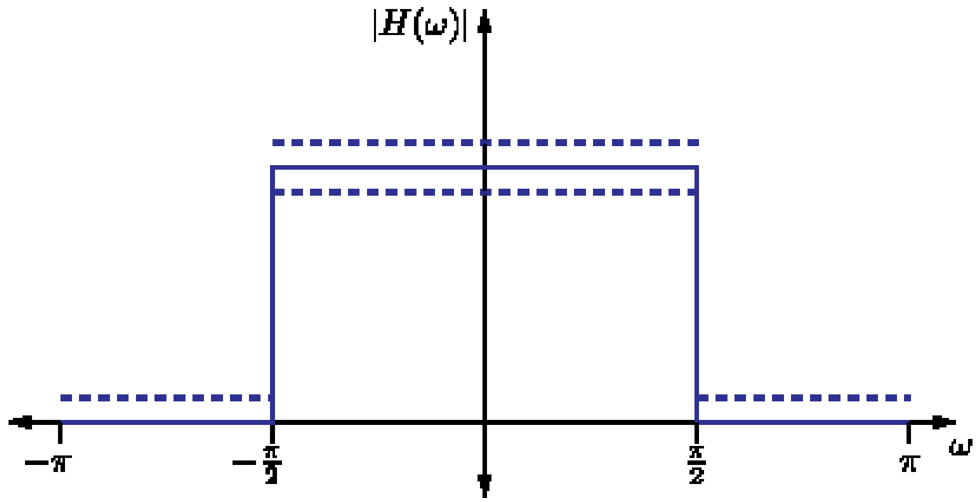



**Figure 1.22**

**Geometry representation:** In compressing 3D geometry, can be useful to bound the $L_\infty$ error to ensure that basic shapes of narrow features (like poles, power lines, etc.) are preserved.

**Sparsity:** In the case where the error is known to be sparse (i.e., zero on most indices) it can be useful to measure the error in the $\ell_1$ norm.

# Chapter 2

# Representation and Analysis of Systems

## 2.1 Linear Systems[1]

In this course we will focus much of our attention on *linear* systems. When our input and output signals are vectors, then the system is a *linear operator*.

Suppose that $L : X \rightarrow Y$ is a linear operator from a vector space $X$ to a vector space $Y$. If $X$ and $Y$ are normed vector spaces, then we can also define a norm on $L$. Specifically, we can let

$$\begin{aligned} \|L\|_{\mathcal{L}(X,Y)} &= \underset{x \in X}{max} \frac{\|L_x\|_Y}{\|x\|_X} \\ &= \underset{x \in X : \|x\|_X = 1}{max} \|L_x\|_Y \end{aligned} \tag{2.1}$$

An operator for which $\|L\|_{\mathcal{L}(X,Y)} < \infty$ is called a *bounded operator*.

**Example 2.1**
BIBO (bounded-input, bounded-output) stable systems are systems for which

$$\|x\|_\infty < A \, \texttt{[U+27F9]} \, \|Lx\|_\infty < B. \tag{2.2}$$

Such a system satisfies $\|L\|_\infty < \frac{B}{A}$.

One can show that $\| \cdot \|_{\mathcal{L}(X,Y)}$ satisfies the requirements of a valid norm. In fact $\mathcal{L}(X,Y) = \{\text{bounded linear operators from } X \text{ to } Y\}$ is itself a normed vector space! If $Y$ is a Banach space, then so is $\mathcal{L}(X,Y)$!

Bounded linear operators are common in DSP—they are "safe" in that "normal" inputs are guaranteed to not make your system explode.

Are there any common systems that are unbounded? Not in finite dimensions, but in infinite dimensions there are plenty of examples!

**Example 2.2**
Consider $L_2[-\pi, \pi]$. For any $k$, $f_k(t) = \frac{1}{\sqrt{2\pi}} e^{-jkt}$ is an element of $L_2[-\pi, \pi]$ with $\|f_k(t)\|_2 = 1$. Consider the system $D = \frac{d}{dt}$, and note that

$$\frac{d}{dt} f_k(t) = \frac{-jk}{\sqrt{2\pi}} e^{-jkt} \, \texttt{[U+27F9]} \, \|Df_k(t)\|_2 = |k|. \tag{2.3}$$

Since $f_k(t) \in L_2[-\pi, \pi]$ for all $k$, we can set $k$ to be as large as we want, so $D$ cannot be bounded.

A very important class of linear operators are those for which $X = Y$. In this case we have the following important definition.

[1] This content is available online at <http://cnx.org/content/m34053/1.2/>.

**Definition 1**
Suppose that $L = X \to X$ is a linear operator. An *eigenvector* is a vector $x$ for which $Lx = \alpha x$ for some $\alpha \in K$ (i.e. $\alpha \in \mathbb{R}$ or $\alpha \in \mathbb{C}$). In this case, $\alpha$ is called the corresponding *eigenvalue*.

Eigenvalues and eigenvectors tell you a lot about a system (more on this later!). While they can sometimes be tricky to calculate (unless you know the eig command in Matlab), we will see that as engineers we can usually get away with the time-honored method of "guess and check".

## 2.2 Discrete-Time Systems[2]

We begin with the simplest of discrete-time systems, where $X = \mathbb{C}^N$ and $Y = \mathbb{C}^M$. In this case a linear operator is just an $M \times N$ matrix. We can generalize this concept by letting $M$ and $N$ go to $\infty$, in which case we can think of a linear operator $L : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ as an infinite matrix.

**Example 2.3**
Consider the shift operator $\Delta_k : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ that takes a sequence and shifts it by $k$. As an example, $\Delta_1$ can be viewed as the infinite matrix given by

$$\begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ y_{-1} \\ y_0 \\ y_1 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & & & & & \cdots & 0 \\ \ddots & \ddots & & & & & & & \vdots \\ \ddots & \ddots & \ddots & & & & & & \\ & 0 & 1 & 0 & & & & & \\ & & 0 & 1 & 0 & & & & \\ & & & 0 & 1 & 0 & & & \\ \vdots & & & & \ddots & \ddots & \ddots & & \\ 0 & \cdots & & & & \ddots & \ddots & \ddots & \end{bmatrix} \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ x_{-1} \\ x_0 \\ x_1 \\ \vdots \\ \vdots \end{bmatrix} \tag{2.4}$$

Note that $\|\Delta_k\|_{\ell_2} = 1$ (for any $k$ and $p$) since the delay doesn't change the norm of $x$. The delay operator is also an example of a linear shift-invariant (LSI) system.

**Definition 1**
An operator $L : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ is called shift-invariant if $L(\Delta_k(x)) = \Delta_k(L(x))$ for all $x \in \ell_2(\mathbb{Z})$ and for any $k \in \mathbb{Z}$.

Observe that $\Delta_{k_1}(\Delta_{k_2}(x)) = \Delta_{k_1+k_2}(x)$ so that $\Delta_k$ itself is an LSI operator.

Lets take a closer look at the structure of an LSI system by viewing it as an infinite matrix. In this case we write $y = Hx$ to denote

$$\begin{bmatrix} \vdots \\ y_{-1} \\ y_0 \\ y_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} & \vdots & \vdots & \vdots & \\ & | & | & | & \\ \cdots & h^{-1} & h^0 & h^1 & \cdots \\ & | & | & | & \\ & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \vdots \\ x_{-1} \\ x_0 \\ x_1 \\ \vdots \end{bmatrix} \tag{2.5}$$

[2]This content is available online at <http://cnx.org/content/m34041/1.2/>.

Suppose we want to figure out the column of $H$ corresponding to $h^0$. What input $x$ could help us determine $h^0$? Consider the vector

$$x = \begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}, \tag{2.6}$$

i.e., $x = \delta[n]$. For this input $y = Hx = h^0$. What about $h^1$? $\Delta_1(x) = \delta[n-1]$ would yield $h^1$. In general $\Delta_k(x) = \delta[n-k]$ tell us the column $h^k$. *But*, if $H$ is LSI, then

$$\begin{aligned} h^k &= H(\Delta_k(\delta[n])) \\ &= \Delta_k(H(\delta[n])) \\ &= \Delta_k\left(h^0\right) \end{aligned} \tag{2.7}$$

This means that each column is just a shifted version of $h^0$, which is usually called the impulse response.

Now just to keep notation clean, let $h = h^0$ denote the impulse response. Can we get a simple formula for the output $y$ in terms of $h$ and $x$? Observe that we can write

$$\begin{bmatrix} \vdots \\ y_{-1} \\ y_0 \\ y_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} & \vdots & \vdots & \vdots & \\ & h_0 & h_{-1} & h_{-2} & \\ \cdots & h_1 & h_0 & h_{-1} & \cdots \\ & h_2 & h_1 & h_0 & \\ & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} \vdots \\ x_{-1} \\ x_0 \\ x_1 \\ \vdots \end{bmatrix} \tag{2.8}$$

Each column is just shifted down one. (Each successive row is also shifted right one.) Looking at $y_{-1}$, $y_0$ and $y_1$, we can rewrite this formula as

$$\begin{pmatrix} y[-1] \\ y[0] \\ y[1] \end{pmatrix} = \cdots + x[-1]\begin{pmatrix} h[0] \\ h[1] \\ h[2] \end{pmatrix} + x[0]\begin{pmatrix} h[-1] \\ h[0] \\ h[1] \end{pmatrix} + x[1]\begin{pmatrix} h[-2] \\ h[-1] \\ h[0] \end{pmatrix} + \cdots \tag{2.9}$$

From this we can observe the general pattern

$$y[n] = \cdots + x[-1]\,h[n+1] + x[0]\,h[n+0] + x[1]\,h[n-1] + \cdots \tag{2.10}$$

or more concisely

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]\,h[n-k]. \tag{2.11}$$

Does this look familiar? It is simply the formula for the discrete-time *convolution* of $x$ and $h$, i.e.,

$$y = x * h. \tag{2.12}$$

## 2.3 Eigenvectors of LSI Systems[3]

Suppose that $h$ is the impulse response of an LSI system. Consider an input $x[n] = z^n$ where $z$ is a complex number. What is the output of the system? Recall that $x * h = h * x$. In this case, it is easier to use the formula:

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k]\, x[n-k] \\ &= \sum_{k=-\infty}^{\infty} h[k]\, z^{n-k} \\ &= z^n \sum_{k=-\infty}^{\infty} h[k]\, z^{-k} \\ &= x[n]\, H(z) \end{aligned} \tag{2.13}$$

where

$$H(z) = \sum_{k=-\infty}^{\infty} h[k]\, z^{-k}. \tag{2.14}$$

In the event that $H(z)$ converges, we see that $y[n]$ is just a re-scaled version of $x[n]$. Thus, $x[n]$ is an eigenvector of the system $H$, right? Not exactly, but almost... technically, since $z^n \notin \ell_2(\mathbb{Z})$ it isn't really an eigenvector. However, most DSP texts ignore this subtlety. The intuition provided by thinking of $z^n$ as an eigenvector is worth the slight abuse of terminology.

Next time we will analyze the function $H(z)$ in greater detail. $H(z)$ is called the *$z$-transform* of $h$, and provides an extremely useful characterization of a discrete-time system.

## 2.4 The z-Transform[4]

### 2.4.1 The $z$-transform

We introduced the $z$-transform before as

$$H(z) = \sum_{k=-\infty}^{\infty} h[k]\, z^{-k} \tag{2.15}$$

where $z$ is a complex number. When $H(z)$ exists (the sum converges), it can be interpreted as the "response" of an LSI system with impulse response $h[n]$ to the input of $z^n$. The $z$-transform is useful mostly due to its ability to simplify system analysis via the following result.

**Theorem**

If $y = h * x$, then $Y(z) = H(z)\, X(z)$.

**Proof**

First observe that

$$\begin{aligned} \sum_{n=-\infty}^{\infty} y[n]\, z^{-n} &= \sum_{n=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} x[k]\, h[n-k] \right) z^{-n} \\ &= \sum_{k=-\infty}^{\infty} x[k] \left( \sum_{n=-\infty}^{\infty} h[n-k]\, z^{-n} \right) \end{aligned} \tag{2.16}$$

[3] This content is available online at <http://cnx.org/content/m34026/1.2/>.

[4] This content is available online at <http://cnx.org/content/m34821/1.2/>.

Let $m = n - k$, and note that $z^{-n} = z^{-m} \cdot z^{-k}$. Thus we have

$$\begin{aligned}\sum_{n=-\infty}^{\infty} y[n] z^{-n} &= \sum_{k=-\infty}^{\infty} x[k] \left( \sum_{n=-\infty}^{\infty} h[m] z^{-m} \right) z^{-k} \\ &= \sum_{k=-\infty}^{\infty} x[k] H(z) z^{-k} \\ &= H(z) \left( \sum_{k=-\infty}^{\infty} x[k] z^{-k} \right) \\ &= H(z) X(z)\end{aligned} \tag{2.17}$$

This yields the "transfer function"

$$H(z) = \frac{Y(z)}{X(z)}. \tag{2.18}$$

## 2.5 The Discrete-Time Fourier Transform[5]

### 2.5.1 The discrete-time Fourier transform

The (non-normalized) DTFT is simply a special case of the $z$-transform for the case $|z| = 1$, i.e., $z = e^{j\omega}$ for some value $\omega \in [-\pi, \pi]$

$$X\left(e^{j\omega}\right) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}. \tag{2.19}$$

The picture you should have in mind is the complex plane. The $z$-transform is defined on the whole plane, and the DTFT is simply the value of the $z$-transform on the unit circle, as illustrated below.

[5]This content is available online at <http://cnx.org/content/m34819/1.2/>.

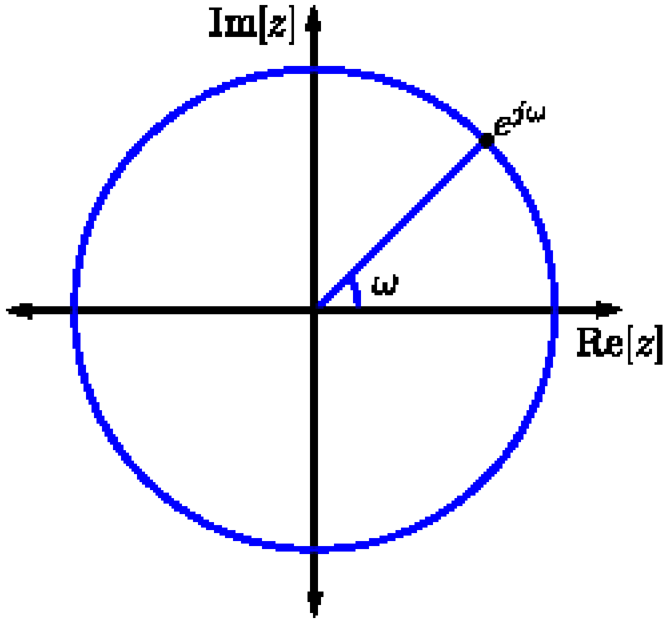



**Figure 2.1**

This picture should make it clear why the DTFT is defined only for $\omega \in [-\pi, \pi]$ (or why it is periodic). Using the normalization above, we also have the inverse DTFT formula:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X\left(e^{j\omega}\right) e^{j\omega n} d\omega. \tag{2.20}$$

# 2.6 z-Transform Examples[6]

## 2.6.1 $z$-transform examples

**Example 2.4**

Consider the $z$-transform given by $H(z) = z$, as illustrated below.

[6]This content is available online at <http://cnx.org/content/m34823/1.2/>.

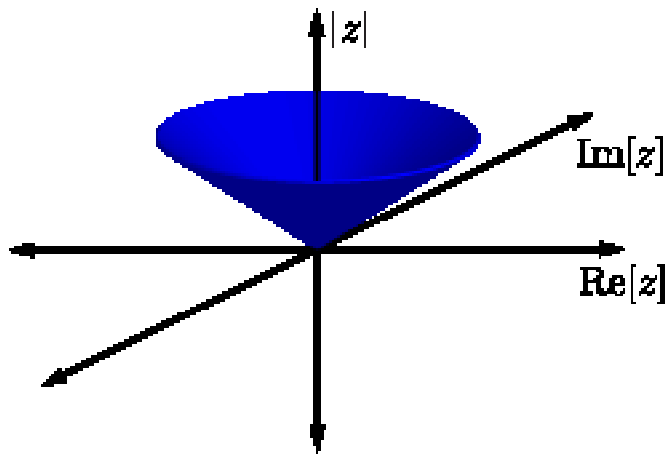



**Figure 2.2**

The corresponding DTFT has magnitude and phase given below.

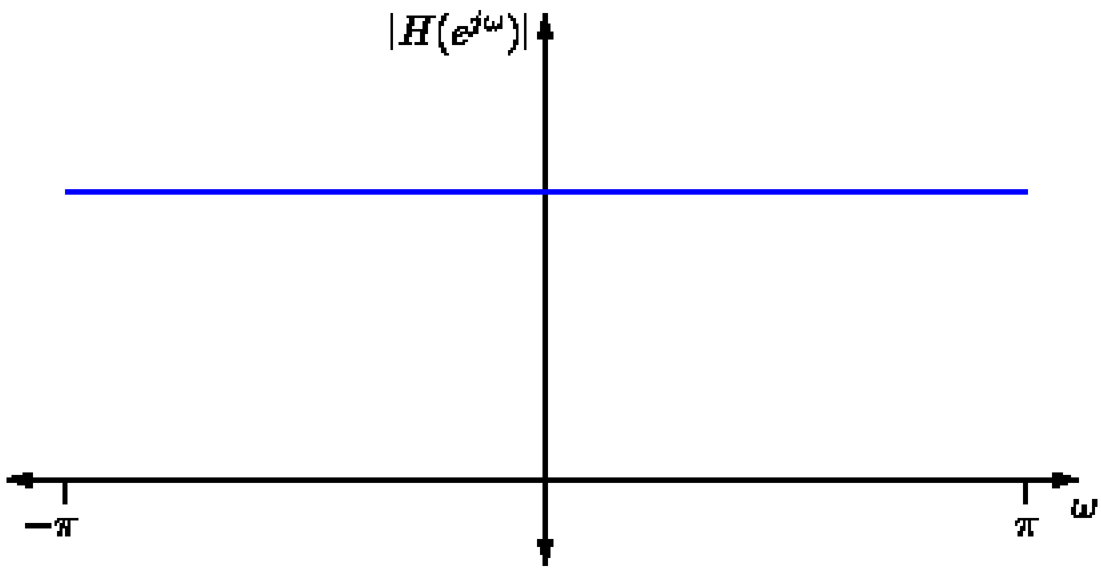



**Figure 2.3**

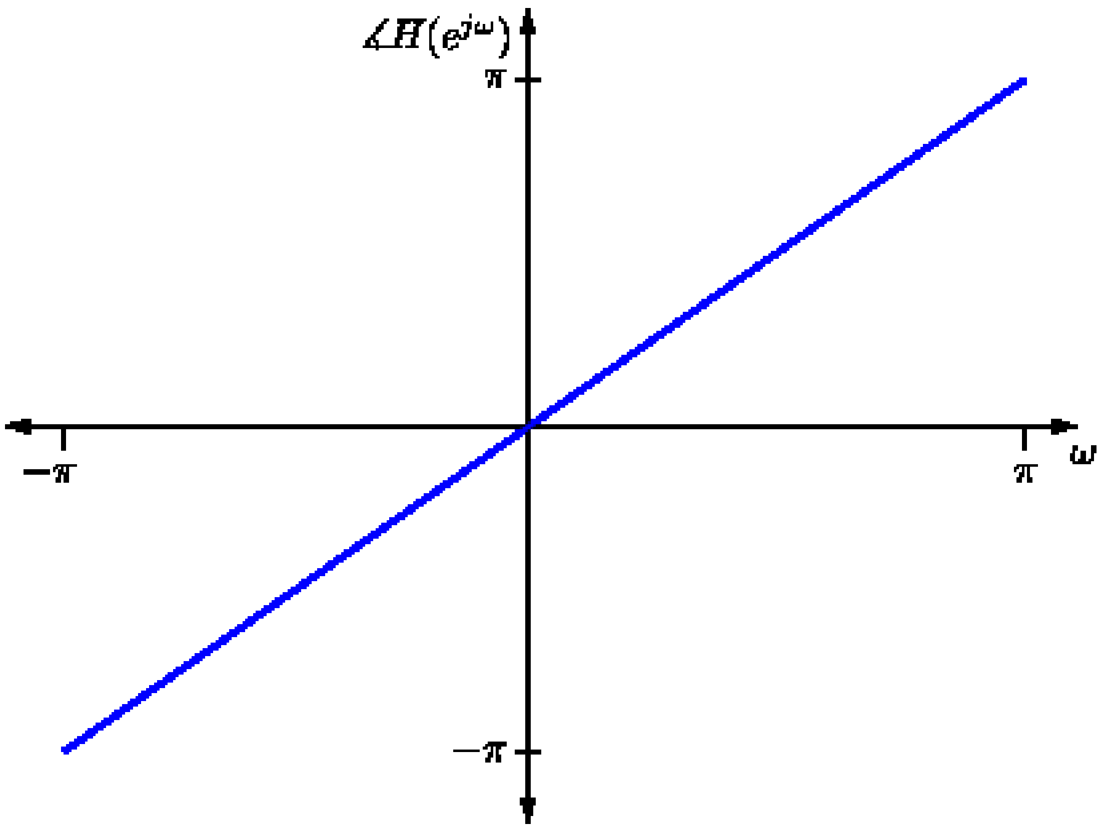



**Figure 2.4**

What could the system $H$ be doing? It is a perfect all-pass, linear-phase system. But what does this mean?

Suppose $h[n] = \delta[n - n_0]$. Then

$$\begin{aligned} H(z) &= \sum_{n=-\infty}^{\infty} h[n] z^{-n} \\ &= \sum_{n=-\infty}^{\infty} \delta[n - n_0] z^{-n} \\ &= z^{-n_0}. \end{aligned} \tag{2.21}$$

Thus, $H(z) = z^{-n_0}$ is the $z$-transform of a system that simply delays the input by $n_0$. $H(z) = z^{-1}$ is the $z$-transform of a unit-delay.

**Example 2.5**

Now consider $x[n] = \alpha^n u[n]$

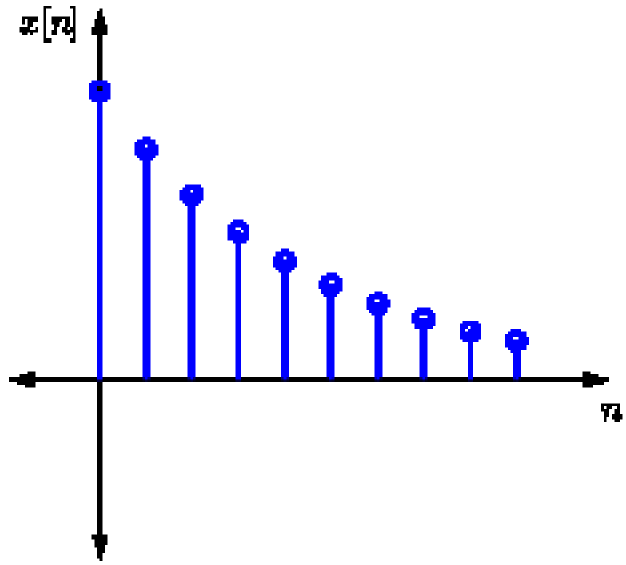



**Figure 2.5**

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^{\infty} x[n]\, z^{-n} = \sum_{n=0}^{\infty} \alpha^n z^{-n} \\ &= \sum_{n=0}^{\infty} \left(\frac{\alpha}{z}\right)^n \\ &= \frac{1}{1-\frac{\alpha}{z}} \quad (\text{if } |\alpha/z| < 1) \ \ (\text{Geometric Series}) \\ &= \frac{z}{z-\alpha} \end{aligned} \tag{2.22}$$

What if $\left|\frac{a}{z}\right| \geq 1$? Then $\sum_{n=0}^{\infty} \left(\frac{\alpha}{n}\right)^n$ does not converge! Therefore, whenever we compute a $z$-transform, we must also specify the set of $z$'s for which the $z$-transform exists. This is called the *region of convergence* (ROC). In the above example, the ROC$=\{z : |z| > |\alpha|\}$.

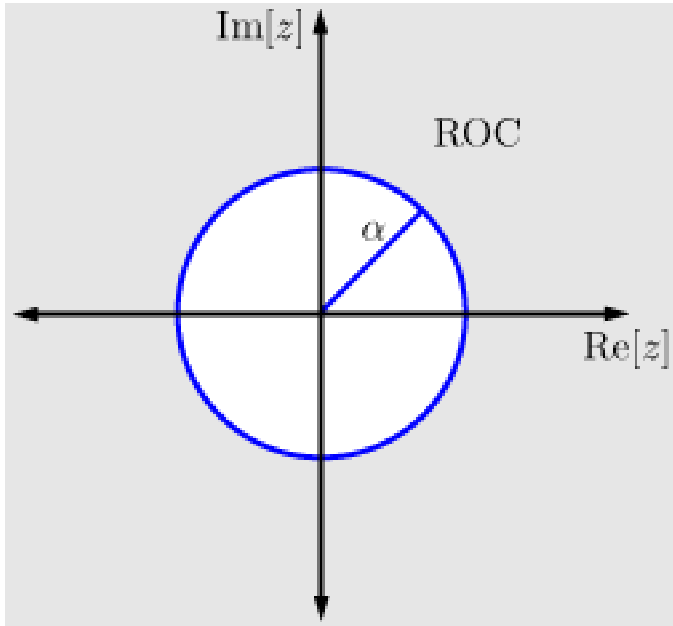



**Figure 2.6**

**Example 2.6**
What about the "evil twin" $x[n] = -\alpha^n u[-1-n]$?

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^{\infty} -\alpha^n u[-1-n] z^{-n} = \sum_{n=-\infty}^{-1} -\alpha^n z^{-n} \\ &= -\sum_{n=-\infty}^{-1} \left(\frac{z}{\alpha}\right)^{-n} \\ &= -\sum_{n=1}^{\infty} \left(\frac{z}{\alpha}\right)^{n} \\ &= 1 - \sum_{n=0}^{\infty} \left(\frac{z}{\alpha}\right)^{n} \quad \text{(converges if } |z/\alpha| < 1) \\ &= 1 - \frac{1}{1-\frac{z}{\alpha}} = \frac{\alpha - z - \alpha}{\alpha - z} = \frac{z}{z-\alpha} \end{aligned} \tag{2.23}$$

We get the exact same result but with ROC=$\{z : |z| < |\alpha|\}$.

# 2.7 z-Transform Analysis of Discrete-Time Filters[7]

## 2.7.1 $z$-transform analysis of discrete-time filters

The $z$-transform might seem slightly ugly. We have to worry about the region of convergence, and we haven't even talked about how to invert it yet (it isn't pretty). However, in the end it is worth it because it is extremely useful in analyzing digital filters with feedback. For example, consider the system illustrated below

[7]This content is available online at <http://cnx.org/content/m34822/1.2/>.

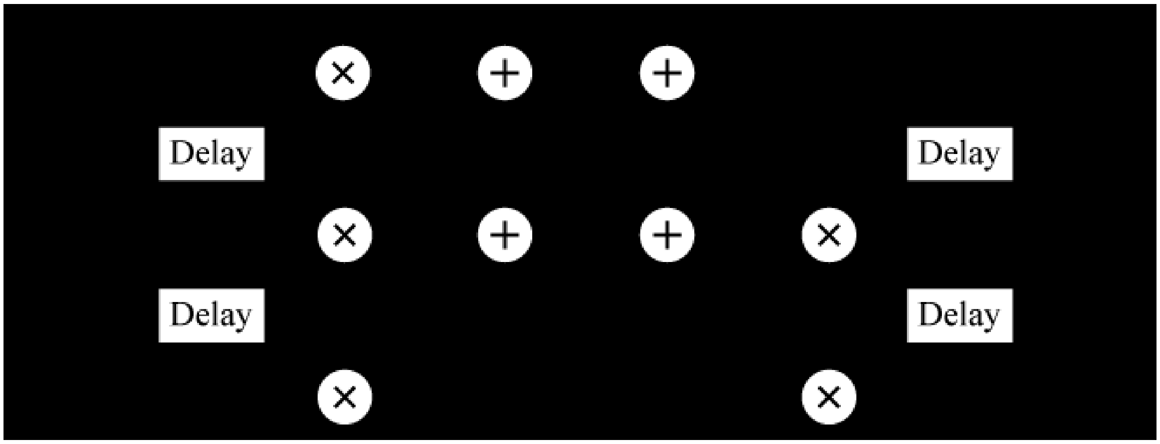



**Figure 2.7**

We can analyze this system via the equations

$$v[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] \tag{2.24}$$

and

$$y[n] = v[n] + a_1 y[n-1] + a_2 y[n-2]. \tag{2.25}$$

More generally,

$$v[n] = \sum_{k=0}^{N} b_k x[n-k] \tag{2.26}$$

and

$$y[n] = \sum_{k=1}^{M} a_k y[n-k] + v[n] \tag{2.27}$$

or equivalently

$$\sum_{k=0}^{N} b_k x[n-k] = y[n] - \sum_{k=1}^{M} a_k y[n-k]. \tag{2.28}$$

In general, many LSI systems satisfy linear difference equations of the form:

$$\sum_{k=0}^{M} a_k y[n-k] = \sum_{k=0}^{N} b_k x[n-k]. \tag{2.29}$$

What does the $z$-transform of this relationship look like?

$$\begin{array}{ll} Z\{\sum_{k=0}^{M} a_k y[n-k]\} & = Z\{\sum_{k=0}^{M} b_k x[n-k]\} \\ \sum_{k=0}^{M} a_k Z\{y[n-k]\} & = \sum_{k=0}^{N} b_k Z\{x[n-k]\}. \end{array} \tag{2.30}$$

Note that

$$\begin{aligned} Z\{y[n-k]\} &= \sum_{n=-\infty}^{\infty} y[n-k]\, z^{-n} \\ &= \sum_{m=-\infty}^{\infty} y[m]\, z^{-m} \cdot z^{-k} \\ &= Y(z)\, z^{-k}. \end{aligned} \tag{2.31}$$

Thus the relationship above reduces to

$$\begin{aligned} \sum_{k=0}^{M} a_k Y(z)\, z^{-k} &= \sum_{k=0}^{N} b_k X(z)\, z^{-k} \\ Y(z)\left(\sum_{k=0}^{M} a_k z^{-k}\right) &= X(z)\left(\sum_{k=0}^{N} b_k z^{-k}\right) \\ \frac{Y(z)}{X(z)} &= \frac{\left(\sum_{k=0}^{N} b_k z^{-k}\right)}{\left(\sum_{k=0}^{M} a_k z^{-k}\right)} \end{aligned} \tag{2.32}$$

Hence, given a system like the one above, we can pretty much immediately write down the system's transfer function, and we end up with a rational function, i.e., a ratio of two polynomials in $z$. Similarly, given a rational function, it is easy to realize this function in a simple hardware architecture. We will focus exclusively on such rational functions in this course.

# 2.8 Poles and Zeros[8]

## 2.8.1 Poles and zeros

Suppose that $X(z)$ is a rational function, i.e.,

$$X(z) = \frac{P(z)}{Q(z)} \tag{2.33}$$

where $P(z)$ and $Q(z)$ are both *polynomials* in $z$. The roots of $P(z)$ and $Q(z)$ are very important.

**Definition 2.1: zero**

A *zero* of $X(z)$ is a value of $z$ for which $X(z) = 0$ (or $P(z) = 0$). A *pole* of $X(z)$ is a value of $z$ for which $X(z) = \infty$ (or $Q(z) = 0$).

For finite values of $z$, poles are the roots of $Q(z)$, but poles can also occur at $z = \infty$. We denote poles in a $z$-plane plot by "×" we denote zeros by "○". Note that the ROC clearly cannot contain any poles since by definition the ROC only contains $z$ for which the $z$-transform converges, and it does not converge at poles.

**Example 2.7**

Consider

$$x_1[n] = \alpha^n u[n] \overset{Z}{\texttt{[U+27F7]}} X_1(z) = \frac{z}{z-\alpha}, \quad |z| > |\alpha| \tag{2.34}$$

and

$$x_2[n] = -\alpha^n u[-1-n] \overset{Z}{\texttt{[U+27F7]}} X_2(z) = \frac{z}{z-\alpha}, \quad |z| < |\alpha| \tag{2.35}$$

[8]This content is available online at <http://cnx.org/content/m34817/1.2/>.

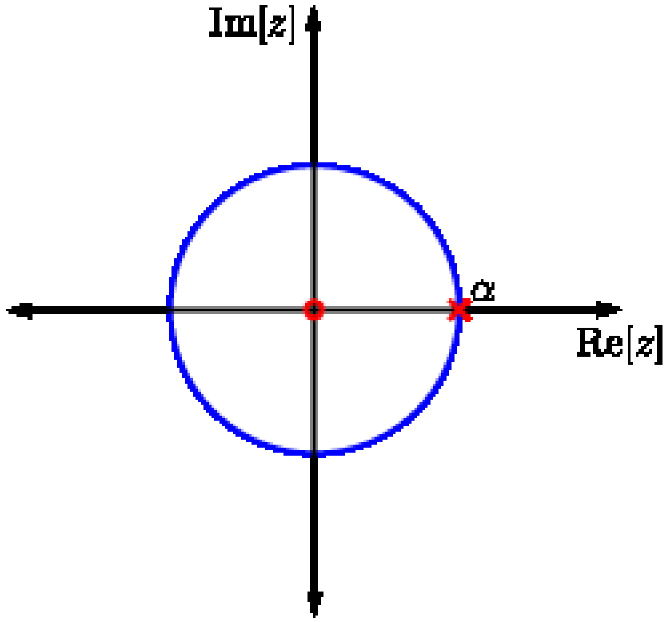



**Figure 2.8**

Note that the poles and zeros of $X_1(z)$ and $X_2(z)$ are identical, but with opposite ROCs. Note also that neither ROC contains the point $\alpha$.

**Example 2.8**

Consider

$$x_3[n] = \left(\frac{1}{2}\right)^n u[n] + \left(-\frac{1}{3}\right)^n u[n]. \tag{2.36}$$

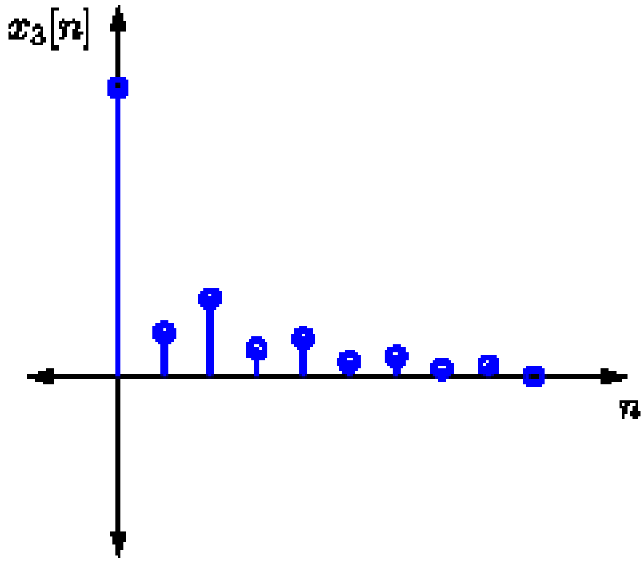



**Figure 2.9**

We can compute the $z$-transform of $x_3[n]$ by simply adding the $z$-transforms of the two different terms in the sum, which are given by

$$\left(\frac{1}{2}\right)^n u[n] \overset{Z}{\texttt{[U+27F7]}} \frac{z}{z-\frac{1}{2}} \quad \text{ROC: } |z| > \frac{1}{2} \tag{2.37}$$

and

$$\left(-\frac{1}{3}\right)^n u[n] \overset{Z}{\texttt{[U+27F7]}} \frac{z}{z+\frac{1}{3}} \quad \text{ROC: } |z| > \frac{1}{3}. \tag{2.38}$$

The poles and zeros for these $z$-transforms are illustrated below.

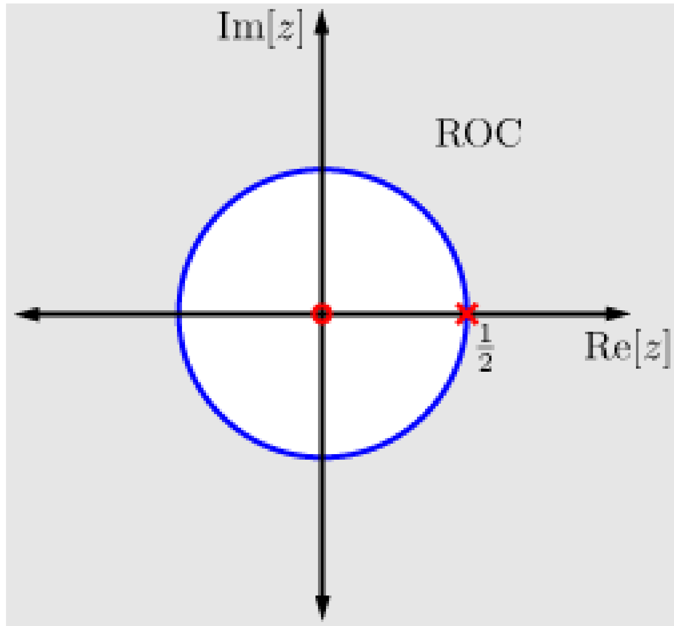



**Figure 2.10**

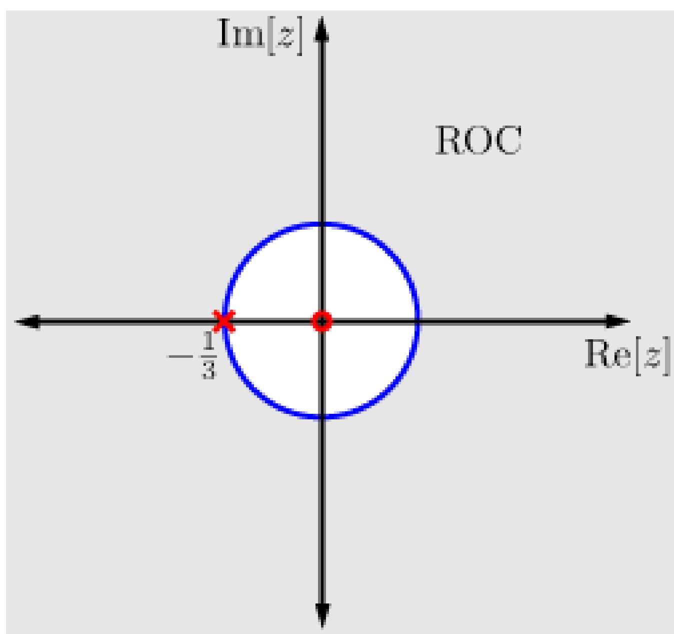



**Figure 2.11**

$X_3(z)$ is given by

$$\begin{aligned} X_3(z) &= \frac{z}{z-\frac{1}{2}} + \frac{z}{z+\frac{1}{3}} \\ &= \frac{z\left(z+\frac{1}{3}\right)+z\left(z-\frac{1}{2}\right)}{\left(z+\frac{1}{3}\right)\left(z-\frac{1}{2}\right)} \\ &= \frac{z\left(2z-\frac{1}{6}\right)}{\left(z+\frac{1}{3}\right)\left(z-\frac{1}{2}\right)} \quad \text{ROC: } |z| > \frac{1}{2} \end{aligned} \tag{2.39}$$

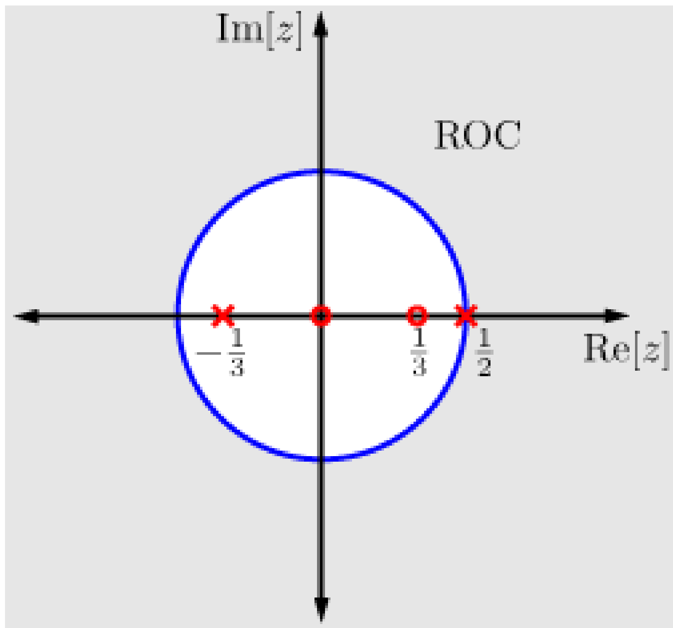



**Figure 2.12**

Note that the poles do not change, but the zeros do, as illustrated above.

**Example 2.9**
Now consider the finite-length sequence

$$x_4[n] = \begin{cases} \alpha^n & 0 \le n \le N-1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.40}$$

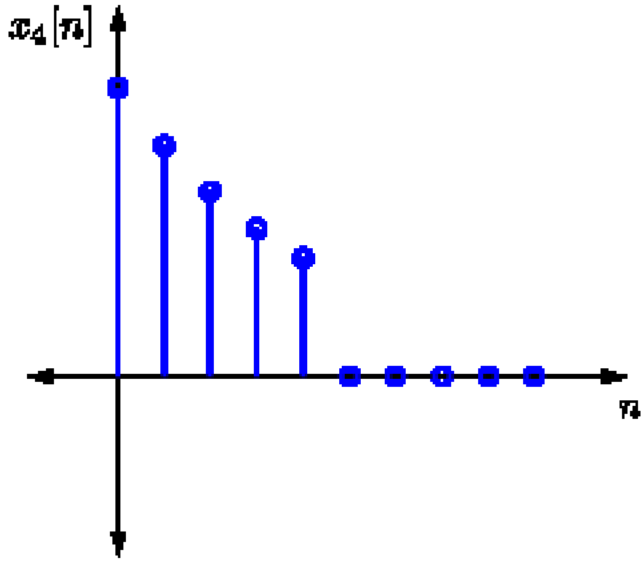



**Figure 2.13**

The $z$-transform for this sequence is

$$X_4(z) = \sum_{n=0}^{N-1} x_4[n] z^{-n} = \sum_{n=0}^{N-1} \alpha^n z^{-n} = \frac{1-\left(\frac{\alpha}{z}\right)^N}{1-\frac{\alpha}{z}} = \frac{z^N-\alpha^N}{z^{N-1}(z-\alpha)} \quad \text{ROC: } z \neq 0 \tag{2.41}$$

We can immediately see that the zeros of $X_4(z)$ occur when $z^N = \alpha^N$. Recalling the "N$^{\text{th}}$ roots of unity", we see that the zeros are given by

$$z_k = \alpha e^{j\frac{2\pi}{N}k}, \; k = 0, 1, ..., N-1. \tag{2.42}$$

At first glance, it might appear that there are $N-1$ poles at zero and 1 pole at $\alpha$, but the pole at $\alpha$ is cancelled by the zero $(z_0)$ at $\alpha$. Thus, $X_4(z)$ actually has only $N-1$ poles at zero and $N-1$ zeros around a circle of radius $\alpha$ as illustrated below.

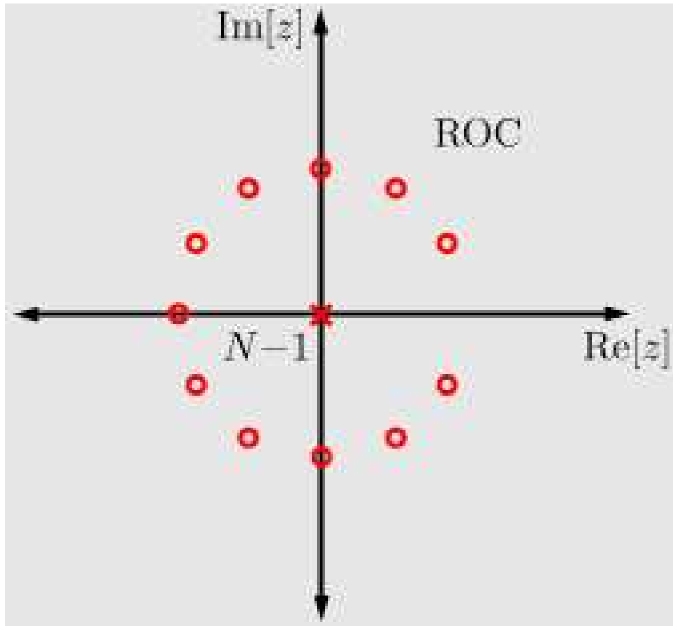



**Figure 2.14**

So, provided that $|\alpha| < \infty$, the ROC is the entire $z$-plane except for the origin. This actually holds for all finite-length sequences.

## 2.9 Stability, Causality, and the z-Transform[9]

### 2.9.1 Stability, causality, and the $z$-transform

In going from

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{m} b_k x[n-k] \tag{2.43}$$

to

$$H(z) = \frac{Y(z)}{X(z)} \tag{2.44}$$

we did not specify an ROC. If we factor $H(z)$, we can plot the poles and zeros in the $z$-plane as below.

[9]This content is available online at <http://cnx.org/content/m34818/1.2/>.

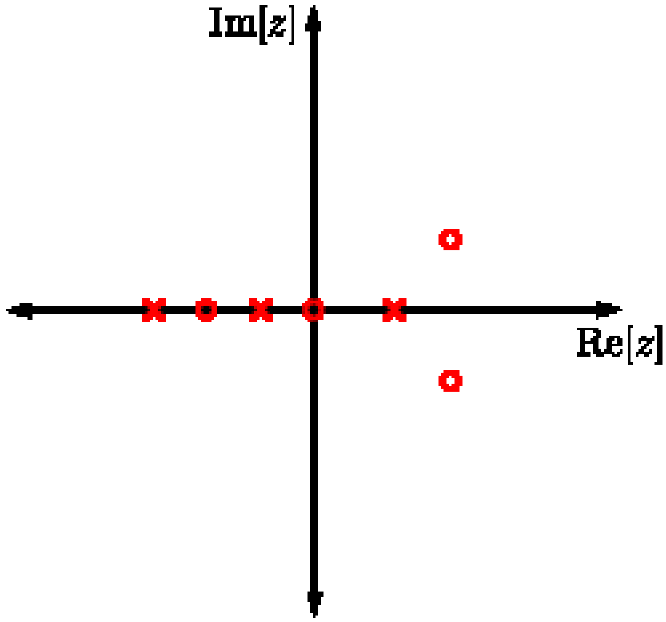



**Figure 2.15**

Several ROCs may be possible. Each ROC corresponds to a different impulse response, so which one should we choose? In general, there is no "right" choice, however, there are some choices that make sense in practice.

In particular, if $h[n]$ is *causal*, i.e., if $h[n] = 0$, $n < 0$, then the ROC extends *outward from the outermost pole*. This can be seen in the examples up to this point. Moreover, recall that a system is *BIBO stable* if the impulse response $h \in \ell_1(\mathbb{Z})$. In this case,

$$\begin{aligned} |H(z)| &= \left|\sum_{n=-\infty}^{\infty} h[n] z^{-n}\right| \\ &\leq \sum_{n=-\infty}^{\infty} |h[n]| \left|z^{-n}\right| \end{aligned} \tag{2.45}$$

Consider the unit circle $z = e^{j\omega}$. In this case we have $|z^{-n}| = |e^{-j\omega n}| = 1$, so that

$$\left|H\left(e^{j\omega}\right)\right| \leq \sum_{n=-\infty}^{\infty} |h[n]| < \infty \tag{2.46}$$

for all $\omega$. Thus, if a system is BIBO stable, the ROC of $H(z)$ must include the unit circle. In general, *any ROC containing the unit circle will be BIBO stable.*

This leads to a key question – are stability and causality always compatible? The answer is *no*. For example, consider

$$H(z) = \frac{z^2}{(z-2)\left(z+\frac{1}{2}\right)} = \frac{\frac{4}{5}z}{z-2} + \frac{\frac{1}{5}z}{z+\frac{1}{2}} \tag{2.47}$$

and its various ROC's and corresponding inverses. If the ROC contains the unit-circle (so that the corresponding system is stable) and is not to contain any poles, then it must extend inward towards the origin, and hence it cannot be causal. Alternatively, if the ROC is to extend outward, it will not contain the unit-circle so that the corresponding system will not be BIBO stable.

## 2.10 Inverse Systems[10]

### 2.10.1 Inverse systems

Many signal processing problems can be interpreted as trying to undo the action of some system. For example, echo cancellation, channel obvolution, etc. The problem is illustrated below.

**Image not finished**

**Figure 2.16**

If our goal is to design a system $H_I$ that reverses the action of $H$, then we clearly need $H(z) H_I(z) = 1$. In the case where

$$H(z) = \frac{P(z)}{Q(z)} \tag{2.48}$$

then this can be achieved via

$$H_I(z) = \frac{Q(z)}{P(z)}. \tag{2.49}$$

Thus, the zeros of $H(z)$ become poles of $H_I(z)$, and the poles of $H(z)$ become zeros of $H_I(z)$. Recall that $H(z)$ being stable and causal implies that all poles are inside the unit circle. If we want $H(z)$ to have a stable, causal inverse $H_I(z)$, then we must have all zeros inside the unit circle, (since they become the poles of $H_I(z)$.) Combining these, $H(z)$ is stable and causal with a stable and causal inverse if and only if all poles and zeros of $H(z)$ are inside the unit circle. This type of system is called a *minimum phase* system.

## 2.11 Inverse z-Transform[11]

### 2.11.1 Inverse $z$-transform

Up to this point, we have ignored how to actually invert a z-transform to find $x[n]$ from $X(z)$. Doing so is very different from inverting a DTFT. We will consider three main techniques:

1. Inspection (look it up in a table)
2. Partial fraction expansion
3. Power series expansion

One can also use contour integration combined with the Cauchy Residue Theorem. See Oppenheim and Schafer for details.

[10]This content is available online at <http://cnx.org/content/m34814/1.2/>.
[11]This content is available online at <http://cnx.org/content/m34815/1.2/>.

#### 2.11.1.1 Inspection

Basically, become familiar with the $z$-transform pairs listed in tables, and "reverse engineer"

**Example 2.10**
Suppose that

$$X\left(z\right)=\frac{z}{z-a},\quad \left|z\right|>\left|a\right|. \tag{2.50}$$

By now you should be able to recognize that $x\left[n\right]=a^n u\left[n\right]$.

#### 2.11.1.2 Partial fraction expansion

If $X\left(z\right)$ is rational, break it up into a sum of elementary forms, each of which can be inverted by inspection.

**Example 2.11**
Suppose that

$$X\left(z\right)=\frac{1+2z^{-1}+z^{-2}}{1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2}},\quad \left|z\right|>1. \tag{2.51}$$

By computing a partial fraction expansion we can decompose $X\left(z\right)$ into

$$X\left(z\right)=\frac{8}{1-z^{-1}}-\frac{9}{1-\frac{1}{2}z^{-1}}+2, \tag{2.52}$$

where each term in the sum can be inverted by inspection.

#### 2.11.1.3 Power Series Expansion

Recall that

$$\begin{array}{rl} X\left(z\right) & =\sum_{n=-\infty}^{\infty}x\left[n\right]z^{-n} \\ & =...x\left[-2\right]z^2+x\left[-1\right]z+x\left[0\right]+x\left[1\right]z^{-1}+x\left[2\right]z^{-2}+.... \end{array} \tag{2.53}$$

If we know the coefficients for the Laurent series expansion of $X\left(z\right)$, then these coefficients give us the inverse $z$-transform.

**Example 2.12**
Suppose

$$\begin{array}{rl} X\left(z\right) & =z^2\left(1-\frac{1}{2}z^{-1}\right)\left(1+z^{-1}\right)\left(1-z^{-1}\right) \\ & =z^2-\frac{1}{2}z-1+\frac{1}{2}z^{-1} \end{array} \tag{2.54}$$

Then

$$x\left[n\right]=\delta\left[n+2\right]-\frac{1}{2}\delta\left[n+1\right]-\delta\left[n\right]+\frac{1}{2}\delta\left[n-1\right]. \tag{2.55}$$

**Example 2.13**
Suppose

$$X\left(z\right)=log\left(1+az^{-1}\right),\quad \left|z\right|>\left|a\right| \tag{2.56}$$

where *log* denotes the complex logarithm. Recalling the Laurent series expansion

$$log\left(1+x\right)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n} \tag{2.57}$$

we can write

$$X\left(z\right)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}a^n}{n}z^{-n}. \tag{2.58}$$

Thus we can infer that

$$x\left[n\right]=\{\begin{array}{cc}\frac{(-1)^{n+1}a^n}{n} & n\geq 1\\ 0 & n\leq 0.\end{array} \tag{2.59}$$

# 2.12 Fourier Representations[12]

## 2.12.1 Fourier Representations

Throughout the course we have been alluding to various Fourier representations. We first recall the appropriate transforms:

**Fourier Series (CTFS):** $x\left(t\right)$: continuous-time, finite/periodic on $[-\pi,\pi]$

$$X\left[k\right]=\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi}x\left(t\right)e^{-jkt}dt \tag{2.60}$$

$$x\left(t\right)=\frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}X\left[k\right]e^{jkt} \tag{2.61}$$

**Discrete-Time Fourier Transform (DTFT):** $x\left[n\right]$: infinite, discrete-time

$$X\left(e^{j\omega}\right)=\frac{1}{\sqrt{2\pi}}\sum_{n=-\infty}^{\infty}x\left[n\right]e^{-j\omega n} \tag{2.62}$$

$$x\left[n\right]=\frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi}X\left(e^{j\omega}\right)e^{j\omega n}d\omega \tag{2.63}$$

**Discrete Fourier Transform (DFT):** $x\left[n\right]$: finite, discrete-time

$$X\left[k\right]=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}x\left[n\right]e^{-j\frac{2\pi}{N}kn} \tag{2.64}$$

$$x\left[n\right]=\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}X\left[k\right]e^{j\frac{2\pi}{N}kn} \tag{2.65}$$

[12] This content is available online at <http://cnx.org/content/m34811/1.2/>.

**Continuous-Time Fourier Transform (CTFT):** $x(t)$: infinite, continuous-time

$$X(\Omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t) e^{-j\Omega t} dt \tag{2.66}$$

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(\Omega) e^{j\Omega t} d\Omega \tag{2.67}$$

We will think of Fourier representations in two complimentary senses:

1. **"Eigenbasis" representations:** Each Fourier transform pair is very naturally related to an appropriate class of LTI systems. In some cases we can think of a Fourier transform as a change of basis.

2. **Unitary operators:** While we often use Fourier transforms to *analyze* certain operators, we can also think of a Fourier transform as itself being an operator.

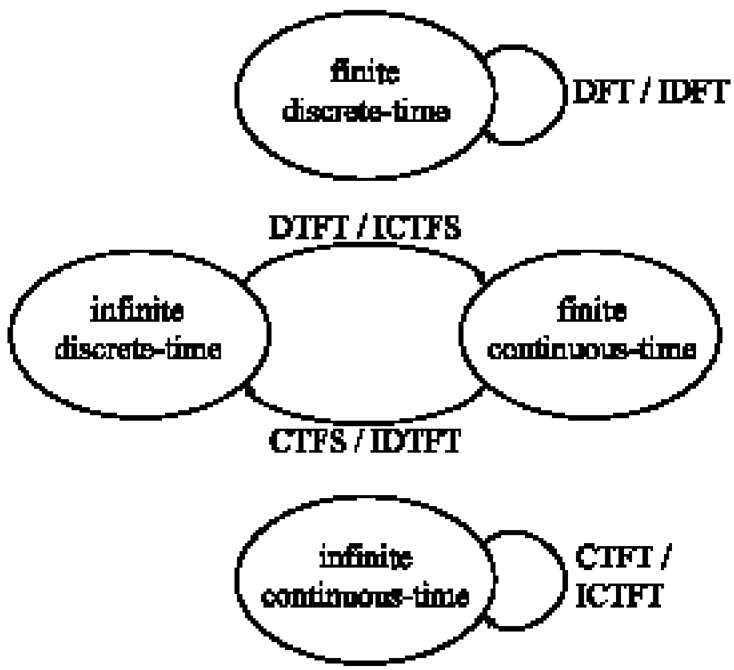



**Figure 2.17**

# 2.13 Normalized DTFT as an Operator[13]

## 2.13.1 Normalized DTFT as an operator

Note that by taking the DTFT of a sequence we get a function defined on $[-\pi, \pi]$. In vector space notation we can view the DTFT as an operator (transformation). In this context it is useful to consider the normalized

[13] This content is available online at <http://cnx.org/content/m34816/1.2/>.

DTFT

$$\mathcal{F}(x) := X\left(e^{j\omega}\right) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} x[n]\, e^{-j\omega n}. \tag{2.68}$$

One can show that the summation converges for any $x \in \ell_2(\pi)$, and yields a function $X\left(e^{j\omega}\right) \in L_2[-\pi, \pi]$. Thus,

$$\mathcal{F} : \ell_2(\mathbb{Z}) \rightarrow L_2[-\pi, \pi] \tag{2.69}$$

can be viewed as a linear operator!

**Note:** It is not at all obvious that $\mathcal{F}$ can be defined for all $x \in \ell_2(\mathbb{Z})$. To show this, one can first argue that if $x \in \ell_1(\mathbb{Z})$, then

$$\begin{aligned} \left|X\left(e^{j\omega}\right)\right| & \leq \left|\tfrac{1}{\sqrt{2\pi}} \textstyle\sum_{n=-\infty}^{\infty} x[n]\, e^{-j\omega n}\right| \\ & \leq \tfrac{1}{\sqrt{2\pi}} \textstyle\sum_{n=-\infty}^{\infty} |x[n]| \left|e^{-jwn}\right| \\ & = \tfrac{1}{\sqrt{2\pi}} \textstyle\sum_{n=-\infty}^{\infty} |x[n]| < \infty \end{aligned} \tag{2.70}$$

For an $x \in \ell_2(\mathbb{Z}) \backslash \ell_1(\mathbb{Z})$, one must show that it is always possible to construct a sequence $x_k \in \ell_2(\mathbb{Z}) \cap \ell_1(\mathbb{Z})$ such that

$$\lim_{k \to \infty} \| x_k - x \|_2 = 0. \tag{2.71}$$

This means $\{x_k\}$ is a Cauchy sequence, so that since $\ell_2(\mathbb{Z})$ is a Hilbert space, the limit exists (and is $x$). In this case

$$X\left(e^{j\omega}\right) = \lim_{k \to \infty} X_k\left(e^{j\omega}\right). \tag{2.72}$$

So for any $x \in \ell_2(\mathbb{Z})$, we can define $\mathcal{F}(x) = X\left(e^{j\omega}\right)$, where $X\left(e^{j\omega}\right) \in L_2[-\pi, \pi]$.

Can we always get the original $x$ back? Yes, the DTFT is invertible

$$\mathcal{F}^{-1}(X) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} X\left(e^{j\omega}\right) \cdot e^{j\omega n} d\omega \tag{2.73}$$

To verify that $\mathcal{F}^{-1}(\mathcal{F}(x)) = x$, observe that

$$\begin{aligned} \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{-\pi}^{\pi} \left(\tfrac{1}{\sqrt{2\pi}} \textstyle\sum_{k=-\infty}^{\infty} x[k]\, e^{-j\omega k}\right) e^{j\omega n} d\omega & = \tfrac{1}{2\pi} \textstyle\sum_{k=-\infty}^{\infty} x[k] \int_{-\pi}^{\pi} e^{-j\omega(k-n)} d\omega \\ & = \tfrac{1}{2\pi} \textstyle\sum_{k=-\infty}^{\infty} x[k] \cdot 2\pi\delta[n-k] \\ & = x[n] \end{aligned} \tag{2.74}$$

One can also show that for any $X \in L_2[-\pi, \pi]$, $\mathcal{F}\left(\mathcal{F}^{-1}(X)\right) = X$.

Operators that satisfy this property are called *unitary operators* or *unitary transformations*. Unitary operators are nice! In fact, if $A = X \rightarrow Y$ is a unitary operator between two Hilbert spaces, then one can show that

$$< x_1, x_2 >=< Ax_1, Ax_2 > \quad \forall\ x_1, x_2 \in X, \tag{2.75}$$

i.e., unitary operators obey Plancherel's and Parseval's theorems!

# 2.14 Fourier Transforms as Unitary Operators[14]

## 2.14.1 Fourier transforms as unitary operators

We have just seen that the DTFT can be viewed as a unitary operator between $\ell_2(\mathbb{Z})$ and $L_2[-\pi,\pi]$. One can repeat this process for each Fourier transform pair. In fact due to the symmetry between the DTFT and the CTFS, we have already established this for CTFS, i.e.,

$$\text{CTFS:} \quad L_2[-\pi,\pi] \to \ell_2(\mathbb{Z}) \tag{2.76}$$

is a unitary operator. Similarly, we have

$$\text{CTFS:} \quad L_2(\mathbb{R}) \to L_2(\mathbb{R}) \tag{2.77}$$

is a unitary operator as well. The proof of this fact closely mirrors the proof for the DTFT. Finally, we also have

$$\text{DFT:} \quad \mathbb{C}^N \to \mathbb{C}^N. \tag{2.78}$$

This operator is also unitary, which can be easily verified by showing that the DFT matrix is actually a unitary matrix: $U^H U = UU^H = I$.

Note that this discussion *only* applies to finite-energy ($\ell_2/L_2$) signals. Whenever we talk about infinite-energy functions (things like the unit step, delta functions, the all-constant signal) having a Fourier transform, we need to be very careful about whether we are talking about a truly convergent Fourier representation or whether we are merely using an engineering "trick" or convention.

# 2.15 The DTFT as an "Eigenbasis"[15]

## 2.15.1 The DTFT as an "Eigenbasis"

We saw Parseval/Plancherel in the context of orthonormal basis expansions. This begs the question, do $\mathcal{F}$ and $\mathcal{F}^{-1}$ just take signals and compute their representation in another basis?

Let's look at $\mathcal{F}^{-1} : L_2[-\pi,\pi] \to \ell_2(\mathbb{Z})$ first:

$$\mathcal{F}^{-1}\left(X\left(e^{jw}\right)\right) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} X\left(e^{j\omega}\right) e^{j\omega n} d\omega. \tag{2.79}$$

Recall that $X\left(e^{j\omega}\right)$ is really just a function of $\omega$, so if we replace $\omega$ with $t$, we get

$$\mathcal{F}^{-1}(X(t)) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} X(t)\, e^{jtn} dt. \tag{2.80}$$

Does this seem familiar? If $X(t)$ is a periodic function defined on $[-\pi,\pi]$, then $\mathcal{F}^{-1}(X(t))$ is just computing (up to a reversal of the indicies) the continuous-time Fourier series of $X(t)$!

We said before that the Fourier series is a representation in an orthobasis, the sequence of coefficients that we get are just the weights of the different basis elements. Thus we have $\to x[n] = \mathcal{F}\mathcal{F}^{-1}(X(t))$ and

$$X(t) = \sum_{n=-\infty}^{\infty} x[n] \left(\frac{e^{-jtn}}{\sqrt{2\pi}}\right). \tag{2.81}$$

[14] This content is available online at <http://cnx.org/content/m34812/1.2/>.
[15] This content is available online at <http://cnx.org/content/m34820/1.2/>.

What about $\mathcal{F}$? In this case we are taking an $x \in \ell_2(\mathbb{Z})$ and mapping it to an $X \in L_2[-\pi,\pi]$. $X$ represents an infinite set of numbers, and when we weight the functions $e^{j\omega n}$ by $X(\omega)$ and sum them all up, we get back the original signal

$$x[n] = \int_{-\pi}^{\pi} X(\omega)\left(\frac{e^{j\omega n}}{\sqrt{2\pi}}\right) d\omega. \tag{2.82}$$

Unfortunately, $\left\|\frac{e^{j\omega n}}{\sqrt{2\pi}}\right\| = \infty\ (\neq 1)$ so technically, we can't really think of this as a change of basis.

However, as a unitary transformation, $\mathcal{F}$ has everything we would ever want in a basis and more: We can represent any $x \in \ell_2(\mathbb{Z})$ using $\{e^{j\omega n}\}_{\omega\in[-\pi,\pi]}$, and since it is unitary, we have Parseval and Plancherel Theorems as well. On top of that, we already showed that the set of vectors $\{e^{j\omega n}\}_{\omega\in[-\pi,\pi]}$ are eigenvectors of LSI systems – if this really were a basis, it would be called an *eigenbasis*.

Eigenbases are useful because once we represent a signal using an eigenbasis, to compute the output of a system we just need to know what it does to its eigenvectors (i.e., its eigenvalues). For an LSI system, $H(e^{j\omega})$ represents a set of eigenvalues that provide a complete characterization of the system.

## 2.16 Eigenbases and LSI Systems[16]

Why is an eigenbasis so useful? It allows us to greatly simplify the computation of the output for a given input. For example, suppose that $X$ is a vector space and that $L : X \to X$ is a linear operator with eigenvectors $\{v_k\}_{k\in\Gamma}$. If $\{v_k\}_{k\in\Gamma}$ form a basis for $X$, then for any $x \in X$ we can write $x = \sum_{k\in\Gamma} c_k v_k$. In this case we have that

$$\begin{aligned} y &= Lx \\ &= L\left(\sum_{k\in\Gamma} c_k v_k\right) \\ &= \sum_{k\in\Gamma} c_k L(v_k) \\ &= \sum_{k\in\Gamma} c_k \lambda_k v_k \end{aligned} \tag{2.83}$$

In the case of a DT, LSI system $H$, we have that $\frac{1}{\sqrt{2\pi}} e^{-j\omega n}$ is an eigenvector of $H$ and for any $x[n]$ we can write

$$x[n] = \int_{-\pi}^{\pi} X(e^{j\omega})\left(\frac{e^{-j\omega n}}{\sqrt{2\pi}}\right) d\omega. \tag{2.84}$$

From the same line of reasoning as above, we have that

$$\begin{aligned} y[n] &= H(x[n]) \\ &= \int_{-\pi}^{\pi} X(e^{j\omega})\, H\left(\frac{e^{-j\omega n}}{\sqrt{2\pi}}\right) d\omega \\ &= \int_{-\pi}^{\pi} X(e^{j\omega})\, H(e^{j\omega}) \cdot \left(\frac{e^{-j\omega n}}{\sqrt{2\pi}}\right) d\omega \\ &= \int_{-\pi}^{\pi} Y(e^{j\omega}) \cdot \left(\frac{e^{-j\omega n}}{\sqrt{2\pi}}\right) d\omega \end{aligned} \tag{2.85}$$

Whenever we have an eigenbasis, we can represent our operator as simply a diagonal operator when the input and output vectors are represented in the eigenbasis. The fact that convolution in time is equivalent to multiplication in the Fourier domain is just one instance of this phenomenon. Moreover, while we have been focusing primarily on the DTFT, it should now be clear that each Fourier representation forms an eigenbasis for a specific class of operators, each of which defines a particular kind of convolution.

**DTFT:** discrete-time convolution (infinite)

[16] This content is available online at <http://cnx.org/content/m34810/1.2/>.

**CTFT:** continuous-time convolution (infinite)

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau \tag{2.86}$$

**DFT:** discrete-time circular convolution

$$(x \circledast y)[n] = \sum_{k=0}^{N-1} x[k]\, y_N[n-k] \tag{2.87}$$

**CTFS:** continuous-time circular convolution

$$(f \circledast g)(t) = \int_{0}^{\tau} f(t)\, g_T(b-\tau)\, d\tau \tag{2.88}$$

*This* is the main reason why we have to care about circular convolution. It is something that one would almost never want to do – but if you multiply two DFTs together you are doing it implicitly, so *be careful* and remember what it is doing.

# Glossary

**Z** **zero**

A *zero* of $X(z)$ is a value of $z$ for which $X(z) = 0$ (or $P(z) = 0$). A *pole* of $X(z)$ is a value of $z$ for which $X(z) = \infty$ (or $Q(z) = 0$).

# Index of Keywords and Terms

**Keywords** are listed by the section with that keyword (page numbers are in parentheses). Keywords do not necessarily appear in the text of the page. They are merely associated with that section. *Ex.* apples, § 1.1 (1) **Terms** are referenced by the page they appear on. *Ex.* apples, 1

# Attributions

Collection: *Digital Signal Processing*
Edited by: Mark A. Davenport
URL: http://cnx.org/content/col11172/1.4/
License: http://creativecommons.org/licenses/by/3.0/

Module: "Introduction to Digital Signal Processing"
By: Mark A. Davenport
URL: http://cnx.org/content/m33588/1.2/
Pages: 1-3
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Metric Spaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m33568/1.3/
Pages: 5-6
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Completeness"
By: Mark A. Davenport
URL: http://cnx.org/content/m33586/1.2/
Pages: 6-8
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Vector Spaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m33583/1.2/
Page: 8
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Normed Vector Spaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m33561/1.2/
Pages: 8-12
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Inner Product Spaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m33556/1.2/
Pages: 12-13
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Properties of Inner Products"
By: Mark A. Davenport
URL: http://cnx.org/content/m33573/1.2/
Pages: 13-14
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Complete Vector Spaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m33548/1.2/
Page: 14
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Hilbert Spaces in Signal Processing"
By: Mark A. Davenport
URL: http://cnx.org/content/m34036/1.2/
Pages: 14-15
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Linear Combinations of Vectors"
By: Mark A. Davenport
URL: http://cnx.org/content/m34030/1.2/
Pages: 15-16
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Vector Subspaces"
By: Mark A. Davenport
URL: http://cnx.org/content/m34073/1.2/
Pages: 16-17
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Signal Approximation in a Hilbert Space"
By: Mark A. Davenport
URL: http://cnx.org/content/m34068/1.2/
Pages: 17-18
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Linear Operators"
By: Mark A. Davenport
URL: http://cnx.org/content/m34058/1.2/
Pages: 18-19
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Projections"
By: Mark A. Davenport
URL: http://cnx.org/content/m34064/1.2/
Pages: 19-20
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Linear Independence"
By: Mark A. Davenport
URL: http://cnx.org/content/m34042/1.2/
Page: 20
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Bases"
By: Mark A. Davenport
URL: http://cnx.org/content/m34017/1.2/
Page: 21
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Orthogonal Bases"
By: Mark A. Davenport
URL: http://cnx.org/content/m34067/1.2/
Pages: 21-24
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Computing the Best Approximation"
By: Mark A. Davenport
URL: http://cnx.org/content/m34021/1.2/
Pages: 24-26
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Matrix Representation of the Approximation Problem"
By: Mark A. Davenport
URL: http://cnx.org/content/m34050/1.2/
Pages: 26-27
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Orthobasis Expansions"
By: Mark A. Davenport
URL: http://cnx.org/content/m34055/1.2/
Pages: 27-28
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Parseval's and Plancherel's Theorems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34062/1.2/
Page: 28
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Error of the Best Approximation in an Orthobasis"
By: Mark A. Davenport
URL: http://cnx.org/content/m34038/1.3/
Pages: 28-29
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Approximation in $\ell$_p Norms"
By: Mark A. Davenport
URL: http://cnx.org/content/m34019/1.3/
Pages: 29-32
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Linear Systems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34053/1.2/
Pages: 33-34
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Discrete-time Systems"
Used here as: "Discrete-Time Systems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34041/1.2/
Pages: 34-35
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Eigenvectors of LSI Systems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34026/1.2/
Page: 36
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "The z-Transform"
By: Mark A. Davenport
URL: http://cnx.org/content/m34821/1.2/
Pages: 36-37
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "The Discrete-Time Fourier Transform"
By: Mark A. Davenport
URL: http://cnx.org/content/m34819/1.2/
Pages: 37-38
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "z-Transform Examples"
By: Mark A. Davenport
URL: http://cnx.org/content/m34823/1.2/
Pages: 38-42
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "z-Transform Analysis of Discrete-Time Filters"
By: Mark A. Davenport
URL: http://cnx.org/content/m34822/1.2/
Pages: 42-44
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Poles and Zeros"
By: Mark A. Davenport
URL: http://cnx.org/content/m34817/1.2/
Pages: 44-50
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Stability, Causality, and the z-Transform"
By: Mark A. Davenport
URL: http://cnx.org/content/m34818/1.2/
Pages: 50-52
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Inverse Systems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34814/1.2/
Page: 52
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Inverse z-Transform"
By: Mark A. Davenport
URL: http://cnx.org/content/m34815/1.2/
Pages: 52-54
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Fourier Representations"
By: Mark A. Davenport
URL: http://cnx.org/content/m34811/1.2/
Pages: 54-55
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Normalized DTFT as an Operator"
By: Mark A. Davenport
URL: http://cnx.org/content/m34816/1.2/
Pages: 55-56
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Fourier Transforms as Unitary Operators"
By: Mark A. Davenport
URL: http://cnx.org/content/m34812/1.2/
Page: 57
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "The DTFT as an "Eigenbasis""
By: Mark A. Davenport
URL: http://cnx.org/content/m34820/1.2/
Pages: 57-58
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

Module: "Eigenbases and LSI Systems"
By: Mark A. Davenport
URL: http://cnx.org/content/m34810/1.2/
Pages: 58-59
Copyright: Mark A. Davenport
License: http://creativecommons.org/licenses/by/3.0/

**Digital Signal Processing**

This course provides an overview of discrete-time signal processing from a vector space perspective. Topics will include sampling, filter design, multirate signal processing and filterbanks, Fourier and wavelet analysis, subspace methods, and a variety of topics relating to inverse problems and "least-squares signal processing".

**About Connexions**

Since 1999, Connexions has been pioneering a global system where anyone can create course materials and make them fully accessible and easily reusable free of charge. We are a Web-based authoring, teaching and learning environment open to anyone interested in education, including students, teachers, professors and lifelong learners. We connect ideas and facilitate educational communities.

Connexions's modular, interactive courses are in use worldwide by universities, community colleges, K-12 schools, distance learners, and lifelong learners. Connexions materials are in many languages, including English, Spanish, Chinese, Japanese, Italian, Vietnamese, French, Portuguese, and Thai. Connexions is part of an exciting new information distribution system that allows for **Print on Demand Books**. Connexions has partnered with innovative on-demand publisher QOOP to accelerate the delivery of printed course materials and textbooks into classrooms worldwide at lower prices than traditional academic publishers.