

# Signals and Systems Quick Reference

## Revision 0.3

Emmy Chow  
"Make Math Make Sense Again"

### Contents

<b>1</b>	<b>Notation and Definitions</b>	<b>4</b>
1.1	Notation . . . . .	4
1.1.1	Logic . . . . .	4
1.1.2	Set Definitions . . . . .	4
1.1.3	Linear Algebra . . . . .	5
1.1.4	Calculus . . . . .	7
1.1.5	Complex Numbers . . . . .	7
1.1.6	Linear Systems Theory . . . . .	8
1.1.7	Digital Signal Processing . . . . .	9
1.2	Function Definitions . . . . .	10
<b>2</b>	<b>Identities</b>	<b>11</b>
2.1	Trig Identities . . . . .	11
2.2	Other Identities . . . . .	12
<b>3</b>	<b>Linear Algebra</b>	<b>13</b>
3.1	LU Decomposition . . . . .	13
3.2	Cholesky Decomposition . . . . .	13
3.3	QR Decomposition . . . . .	14
3.4	Eigen Decomposition / Diagonalization . . . . .	15
3.5	Jordan Decomposition and the Jordan Normal/Cannical Form . . . . .	16
3.6	Singular Value Decomposition . . . . .	18
3.7	Quadratic Forms . . . . .	20
<b>4</b>	<b>Linear Systems Theory</b>	<b>21</b>
4.1	Continuous Time Fourier Transform . . . . .	21
4.1.1	CTFT Properties . . . . .	21
4.1.2	CTFT Table . . . . .	22
4.2	Laplace Transform . . . . .	23
4.2.1	Laplace Transform Properties . . . . .	23
4.2.2	Laplace Transform Table . . . . .	24

4.3	Matrix Calculus . . . . .	25
4.3.1	Matrix Derivatives . . . . .	25
4.4	State Space Representation . . . . .	26
4.4.1	State Space Coordinate Systems . . . . .	26
4.5	The Transfer Function . . . . .	27
4.5.1	Analytic Description of the Transfer Function . . . . .	28
4.5.2	The Impulse Response and Time Domain Solution . . . . .	29
4.6	State Space Realizations of Transfer Functions . . . . .	30
4.6.1	SISO Systems . . . . .	30
4.6.2	MIMO Systems . . . . .	31
4.7	Relationship Between CCF and OCF . . . . .	31
4.8	Equilibria and Stability . . . . .	32
4.8.1	Definition of Equilibrium State . . . . .	32
4.8.2	Equilibria of LTI Systems . . . . .	32
4.8.3	Stability of LTI Systems . . . . .	32
4.8.4	Classification of Equilibria . . . . .	32
4.8.5	Lyapanov Stability . . . . .	33
4.8.6	Asymptotic Stability and Exponential Stability . . . . .	33
4.8.7	Diagram of Stability Types . . . . .	33
4.9	Example: Equilibria Classification of a Damped Pendulum . . . . .	34
4.10	Example: Rotation + Scaling . . . . .	35
4.11	The Lyapanov Function . . . . .	36
4.11.1	Derivation . . . . .	36
4.11.2	Application to Establishing GES of Asymptotically Stable LTI Systems . . .	36
4.11.3	Solving for P Given A and Q . . . . .	37
4.11.4	Lyapanov Function When A is Hurwitz . . . . .	37
4.12	Controllability . . . . .	39
4.12.1	Definitions . . . . .	39
4.12.2	Discrete Time Controllability . . . . .	39
4.12.3	Continuous Time Controllability . . . . .	41
4.12.4	Controllability is Coordinate Independent . . . . .	42
4.12.5	The Controllable and Uncontrollable Subspace . . . . .	43
4.12.6	Impulse Inputs . . . . .	45
4.13	Least Norm Input . . . . .	46
4.13.1	The Controllability Grammian . . . . .	48
4.13.2	PCH Test For Controllability . . . . .	50
<b>5</b>	<b>Digital Signal Processing</b>	<b>51</b>
5.1	Discrete Time Fourier Transform . . . . .	51
5.1.1	DTFT Properties . . . . .	51
5.1.2	DTFT Symmetries . . . . .	52
5.1.3	DTFT Table . . . . .	53
5.2	Z Transform . . . . .	54
5.2.1	Z Transform Properties . . . . .	54

5.2.2	Z Transform Table . . . . .	55
5.3	Sampling, Compression, and Expansion . . . . .	56
5.3.1	C/D Conversion . . . . .	56
5.3.2	D/C Conversion . . . . .	57
5.3.3	Compression and Decimation . . . . .	58
5.3.4	Expansion and Interpolation . . . . .	59
5.3.5	Change Sampling Rate by Noninteger Factor . . . . .	60
5.3.6	Polyphase Decompositions . . . . .	61

# 1 Notation and Definitions

## 1.1 Notation

### 1.1.1 Logic

$:=$  means equal by definition

$\Rightarrow \Leftarrow$  means contradiction

### 1.1.2 Set Definitions

#### Topology and Basic Sets

$\partial U$  is the boundary of  $U$

$\bar{U}$  is the closure of  $U$ . i.e. a set plus its boundary.

$\overset{\circ}{U}$  is the interior of  $U$  i.e. a set plus minus its boundary.

$|U|$  is the size or **cardinality** of set  $U$

$\mathbb{N} := \bigcup_{k=1}^{\infty} \{k\}$  is the set of natural numbers

$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$  is the set of whole numbers

$\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\}$

$|\mathbb{N}| = \aleph_0$  is the size of countably infinite sets.

#### Subsets of $\mathbb{Z}$ and $\mathbb{N}_0$

Given  $k, a \in \mathbb{Z}$ ,  $k\mathbb{Z} + a := \{kn + a : n \in \mathbb{Z}\} = \{a, -k + a, k + a, 2k + a, -2k + a, \dots\}$

Given  $k, a \in \mathbb{N}_0$ ,  $k\mathbb{N}_0 + a := \{kn + a : n \in \mathbb{N}_0\} = \{a, k + a, 2k + a, 3k + a, \dots\}$

I extend interval notation to the integers. Instead of a comma, two dots are used between the upper and lower bound when it's set of integers. For example:  $[a .. b) := [a, b) \cap \mathbb{Z}$ .

#### Other Set definitions

$\mathbb{F}$  is either  $\mathbb{R}$  or  $\mathbb{C}$

$\bar{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$  is the **Riemann Sphere**

$\{a_i\}_{i=k}^n := \{a_i\}_{i \in [k .. n]}$  is an **indexed set**. Informally, for indexed sets  $\{a_1, a_2\} \neq \{a_2, a_1\}$ .

#### Algebraic Structure

$(V, \mathbb{F})$  is a vector space

$(V, \mathbb{F}, \|\cdot\|)$  is a **metric space**

$(V, \mathbb{F}, \langle \cdot, \cdot \rangle)$  is an **inner product space**

$U \leq W$  means  $U$  is a vector subspace of  $W$

$U < W$  means  $U$  is a proper vector subspace of  $W$

### 1.1.3 Linear Algebra

$[0]$  is the matrix with all entries equal to 0

$\mathbf{A} \sim \mathbf{B}$  means  $\mathbf{A}$  is row equivalent to  $\mathbf{B}$

$$\text{diag}\{a_i\}_{i=1}^n := \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \mathbf{I}$$

$\bigoplus_{i=1}^n \mathbf{A}_k = \mathbf{A}_1 \oplus \dots \oplus \mathbf{A}_n := \text{diag}\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$  is the [matrix direct sum](#)

$\mathbf{A} \hat{\otimes} \mathbf{B}$  is the [Kronecker product/matrix direct product](#)

$\mathbf{u} \hat{\otimes} \mathbf{w}$  is the [Kronecker outer product](#)

$\mathbf{A}^+$  is the [Moore Penrose Pseudoinverse](#)

### Sets

#### Basis Representation

$\mathcal{E}_U := \{\hat{\mathbf{u}}_i\}_{i=1}^n$  is the standard basis for  $U$

$\mathcal{B}_U := \{\mathbf{u}_i\}_{i=1}^n$  is a basis for  $U$

$[\mathcal{B}_U]$  is the matrix with columns corresponding to elements of the basis  $\mathcal{B}_U$

$\mathcal{E}_n := \{\mathbf{e}_i\}_{i=1}^n$  is the standard basis for  $\mathbb{R}^n$

$\mathbf{x}_{\mathcal{B}_U}$  is the vector  $\mathbf{x} \in U$  wrt basis  $\mathcal{B}_U$

$\mathbf{P}_{\mathcal{B}_U \rightarrow \mathcal{B}_W} := [\mathcal{B}_U]^{-1}[\mathcal{B}_W]$  is the [change of basis matrix](#) from basis  $\mathcal{B}_U$  to basis  $\mathcal{B}_W$

$\mathfrak{L}(U, W)$  is the set of linear operators that maps  $(U, \mathbb{F})$  to  $(W, \mathbb{F})$

$[\mathcal{T}]_{(\mathcal{B}_U, \mathcal{B}_W)}$  is the matrix representation of linear operator  $\mathcal{T}$  with vectors in the domain wrt  $\mathcal{B}_U$  and vectors in the codomain wrt  $\mathcal{B}_W$

### Eigenvalues and Eigenvectors

$$\chi_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda \mathbf{I})$$

$\Lambda(\mathbf{A})$  is the spectrum (set of all eigenvalues) of  $\mathbf{A}$ .

$\mu_{\mathbf{A}}(\lambda)$  is the minimal polynomial for the matrix  $\mathbf{A}$

$\alpha_{\mathbf{A}}(\lambda)$  is the algebraic multiplicity of eigenvalue  $\lambda$  for the matrix  $\mathbf{A}$

$\gamma_{\mathbf{A}}(\lambda)$  is the geometric multiplicity of eigenvalue  $\lambda$  for the matrix  $\mathbf{A}$

### Definiteness

$\mathbf{A} \succ 0 \iff \mathbf{A}$  is positive definite,

$\mathbf{A} \succeq 0 \iff \mathbf{A}$  is positive semidefinite

$\mathbf{A} \prec 0 \iff \mathbf{A}$  is negative definite,

$\mathbf{A} \preceq 0 \iff \mathbf{A}$  is negative semidefinite

### Sets

$\mathbb{W}_+^n$  is the set of  $n \times n$  Hermitian positive definite matrices.

$\overline{\mathbb{W}}_+^n$  is the set of  $n \times n$  Hermitian positive semidefinite matrices.

Suppose  $\mathbf{A} \succeq 0$  and  $\mathbf{A}^H = \mathbf{A}$

By the [Cholesky Decomposition](#), there is a unique  $\mathbf{L}$  where  $\mathbf{A} = \mathbf{L}\mathbf{L}^H$

$$\sqrt{\mathbf{A}} := \mathbf{L}$$

### Norms and Inner Products

$\|\mathbf{x}\|_p := \left( \sum_{k=0}^n |x_k|^p \right)^{\frac{1}{p}}$  is the  $\ell^p$  norm

$$\|\mathbf{x}\| := \|\mathbf{x}\|_2 := \left( \sum_{k=0}^n x_k^2 \right)^{\frac{1}{2}}$$

$$\|\mathbf{x}\|_\infty := \max \{|x_k|\}_{k=0}^n$$

For the following let  $\mathbf{W} \succ 0$  and  $\mathbf{W}^H = \mathbf{W}$

$$\|\mathbf{x}\|_{\mathbf{W}} := \mathbf{x}^H \mathbf{W} \mathbf{x}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{W}} := \mathbf{x}^H \mathbf{W} \mathbf{y}$$

### 1.1.4 Calculus

#### Nth Derivative/Antiderivative

$$\mathcal{D}^n f(t) := f^{(n)}(t)$$

$$\mathcal{D}^{-n} f(t) := \underbrace{\int \cdots \int}_{n \text{ times}} f(t) dt \cdots dt$$

#### Accumulation

$$\mathcal{A}^k f(t) := \underbrace{\int_{-\infty}^t \cdots \int_{-\infty}^t}_{k \text{ times}} f(\tau) d\tau \cdots d\tau$$

$$\mathcal{A}^k f[n] := \underbrace{\sum_{n=-\infty}^n \cdots \sum_{k=-\infty}^n}_{k \text{ times}} f[k]$$

#### Jacobian and Parital Derivatives:

$$\partial_{x_1}^n f(x_1, \dots, x_n) := \frac{\partial^n f}{\partial x^n}$$

$$\partial_{x_1}^{-n} f := \underbrace{\int \cdots \int}_{n \text{ times}} f(x_1, \dots, x_n) dx_1 \cdots dx_1$$

If mixed partials commute we can write:

$$\partial_{x_1, \dots, x_k}^{(k_1, \dots, k_n)} f(x_1, \dots, x_n) := \frac{\partial^K f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}, \text{ where } K = k_1 + \cdots + k_n$$

$$[\mathcal{D}\mathbf{f}(x_1, \dots, x_n)] = \begin{bmatrix} \partial_{x_1} \mathbf{f} & \cdots & \partial_{x_n} \mathbf{f} \end{bmatrix} := \begin{bmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_n} f_1 \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_n & \cdots & \partial_{x_n} f_n \end{bmatrix}$$

### 1.1.5 Complex Numbers

$(a + bj)^* := a - bj$  is the complex conjugate

$$\Re\{a + bj\} := a$$

$$\Im\{a + bj\} := b$$

$$\mathbf{A}^H := (\mathbf{A}^*)^T = (\mathbf{A}^T)^*$$

### 1.1.6 Linear Systems Theory



### 1.1.7 Digital Signal Processing

#### Delta Operator

$$\Delta x[n] := x[n+1] - x[n]$$

$$\Delta^k x[n] := \Delta^{k-1} x[n+1] - \Delta^{k-1} x[n]$$

#### Circular Convolution

$$f(t) \circledast g(t) := \int_{\theta_0}^{\theta_0+2\pi} f(\tau)g(t-\tau) d\tau, \theta_0 \in \mathbb{R}$$

#### Useful Functions/Equations:

$$\text{sgn}(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

$$\text{clamp}_{[a,b]}(x) := \min(\max(x, a), b)$$

$$\text{Arg}(\theta) := \theta - 2\pi \left\lfloor \frac{\theta - \theta_0}{2\pi} \right\rfloor \quad (\text{i.e. } \theta \in \mathbb{R} \text{ maps to } \hat{\theta} \in [\theta_0, \theta_0 + 2\pi])$$

## 1.2 Function Definitions

Name	Definition
(Heaviside) unit step function	$u_H(t) := \begin{cases} 1 & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$
Discrete rectangular function	$\text{rect}[x] := u_H[x] - u_H[x - 1]$
Discrete Triangular function	$\text{tri}[n] := (1 -  x )\text{rect}[\frac{1}{2}(x + 1)]$
Continuous rectangular function	$\text{rect}(x) := u_H(x + \frac{1}{2}) - u_H(x - \frac{1}{2})$
Triangular function	$\text{tri}(x) := (1 -  x )\text{rect}(\frac{1}{2}x)$
ReLU/Ramp function	$\text{relu}(x) := u_H(x)x$
(Normalized) sinc	$\text{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$
Dirichlet kernel	$\text{diric}_N(x) := \frac{\sin((N + \frac{1}{2})x)}{\sin(\frac{x}{2})}$
(Dirac) comb function	$W_T(x) := \sum_{k=-\infty}^{\infty} \delta(t - kT)$

Note:

Discrete functions can only be scaled by integer factors.

## 2 Identities

### 2.1 Trig Identities

#### Co-Function Identities

$$\cos\left(\frac{\pi}{2} - \theta\right) = \sin(\theta)$$

$$\sin\left(\frac{\pi}{2} - \theta\right) = \cos(\theta)$$

$$\tan\left(\frac{\pi}{2} - \theta\right) = \cot(\theta)$$

#### Supplement Angle Identities

$$\sin(\pi - \theta) = \sin(\theta)$$

$$\sin(\pi + \theta) = -\sin(\theta)$$

$$\cos(\pi - \theta) = -\cos(\theta)$$

$$\cos(\pi + \theta) = -\cos(\theta)$$

$$\tan(\pi - \theta) = -\tan(\theta)$$

$$\tan(\pi + \theta) = \tan(\theta)$$

#### Negative Angle Identities

$$\sin(-\theta) = -\sin(\theta)$$

$$\cos(-\theta) = \cos(\theta)$$

$$\tan(-\theta) = -\tan(\theta)$$

#### Additional and Subtraction Identities

$$\sin(x + y) = \sin(x) \cos(y) + \cos(x) \sin(y)$$

$$\sin(x - y) = \sin(x) \cos(y) - \cos(x) \sin(y)$$

$$\cos(x + y) = \cos(x) \cos(y) - \sin(x) \sin(y)$$

$$\cos(x - y) = \cos(x) \cos(y) + \sin(x) \sin(y)$$

#### Product Identities

$$\sin(x) \cos(y) = \frac{1}{2}(\sin(x + y) + \sin(x - y))$$

$$\cos(x) \sin(y) = \frac{1}{2}(\sin(x + y) - \sin(x - y))$$

$$\cos(x) \cos(y) = \frac{1}{2}(\cos(x + y) + \cos(x - y))$$

$$\sin(x) \sin(y) = \frac{1}{2}(\cos(x - y) - \cos(x + y))$$

#### Superposition

$$A_1 \cos(\theta) + A_2 \sin(\theta) = A \cos(\theta - \phi)$$

$$A = \sqrt{A_1^2 + A_2^2}, \phi = \tan^{-1}\left(\frac{A_2}{A_1}\right)$$

#### Power Reduction Formula

$$\sin^2(\theta) = \frac{1}{2}(1 - \cos(2\theta))$$

$$\cos^2(\theta) = \frac{1}{2}(1 + \cos(2\theta))$$

#### Double Angle Identities

$$\sin(2\theta) = 2 \sin(\theta) \cos(\theta)$$

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$$

$$= 2 \cos^2(\theta) - 1$$

$$= 1 - 2 \sin^2(\theta)$$

$$\tan(2\theta) = \frac{2 \tan(\theta)}{1 - \tan^2(\theta)}$$

#### Half Angle Identities

$$\sin\left(\frac{\theta}{2}\right) = \pm \sqrt{\frac{1 - \cos(\theta)}{2}}$$

$$\cos\left(\frac{\theta}{2}\right) = \pm \sqrt{\frac{1 + \cos(\theta)}{2}}$$

#### Sum Identities

$$\sin(x) + \sin(y) = 2 \sin\left(\frac{x + y}{2}\right) \cos\left(\frac{x - y}{2}\right)$$

$$\sin(x) - \sin(y) = 2 \cos\left(\frac{x + y}{2}\right) \sin\left(\frac{x - y}{2}\right)$$

$$\cos(x) + \cos(y) = 2 \cos\left(\frac{x + y}{2}\right) \cos\left(\frac{x - y}{2}\right)$$

$$\cos(x) - \cos(y) = 2 \cos\left(\frac{x + y}{2}\right) \sin\left(\frac{x - y}{2}\right)$$

#### Complex Exponential Identities

$$\sin(z) = \frac{1}{2j}(e^{jz} - e^{-jz})$$

$$\cos(z) = \frac{1}{2}(e^{jz} + e^{-jz})$$

## 2.2 Other Identities

### Convolutions

$$f(x) \star g(x) = f(x) * g(-x)$$

$$|a|(\text{rect}(ax) * \text{rect}(ax)) = \text{tri}(ax)$$

### Dirichlet Kernel

$$\text{diric}_N(x) := \frac{\sin\left(\left(n + \frac{1}{2}\right)x\right)}{\sin\left(\frac{x}{2}\right)} = \sum_{k=-N}^N e^{jkx} = 1 + 2 \sum_{k=1}^N \cos(kx)$$

$$\lim_{N \rightarrow \infty} \text{diric}_N(Tx) = W_{\frac{1}{T}}(x)$$

### Accumulation

$$\mathcal{A}\{\Delta x[t]\} = \Delta\{\mathcal{A}x[t]\} = x[t]$$

$$\mathcal{A}\{\mathcal{D}x(t)\} = \mathcal{D}\{\mathcal{A}x(t)\} = x(t)$$

## 3 Linear Algebra

Reference Text: [Sheldon and Axler's Linear Algebra Done Right](#)

### 3.1 LU Decomposition

The [LU Decomposition](#) factors a matrix into the product of a lower and upper triangular matrix. Invertible matrices have a unique LU Decomposition. The LU Decomposition represents the process of solving a system of equation by substitution rather than elimination.

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$

$$\mathbf{A} = \mathbf{L}\mathbf{U}$$

#### Solving a System using LU

1. First factor  $\mathbf{A} = \mathbf{L}\mathbf{U}$ . Complexity:  $\mathcal{O}(n^3)$
2. Solve  $\mathbf{L}\mathbf{y} = \mathbf{b}$ . Complexity:  $\mathcal{O}(n^2)$
3. Solve  $\mathbf{U}\mathbf{x} = \mathbf{y}$ . Complexity:  $\mathcal{O}(n^2)$

While this algorithm is overall  $\mathcal{O}(n^3)$ , it is typically used when  $\mathbf{A}$  can be factored ahead of time and  $\mathbf{b}$  is a varying unknown quantity. In the case of precomputing the LU Decomposition the algorithm has time complexity  $\mathcal{O}(n^2)$

#### References:

[MIT OCW 18.06 Lecture 4. Factorization  \$\mathbf{A} = \mathbf{L}\mathbf{U}\$](#)

### 3.2 Cholesky Decomposition

The [Cholesky Decomposition](#) is very similar to the LU decomposition and can be computed using the same algorithm as LU; however, in the case of  $\mathbf{A}$  being symmetric positive definite,  $\mathbf{U} = \mathbf{L}^H$ . In practice, a different algorithm from LU is used since it is possible to make it twice as fast owing to the symmetry of  $\mathbf{A}$  as well as make it more numerically stable.

The Cholesky Decomposition is the fastest way to find the square root of a positive definite matrix.

Let  $\mathbf{A} \in \mathbb{W}_+^n$ .

$$\mathbf{A} = \mathbf{L}\mathbf{L}^H, \sqrt{\mathbf{A}} = \mathbf{L}$$

### 3.3 QR Decomposition

The [QR Decomposition](#) factors a matrix into a unitary matrix and an upper triangular matrix. There are infinitely many choices for  $\mathbf{Q}$ , so the QR decomposition is not unique. This algorithm is most often used to calculate the inverse of a matrix in a numerically stable way.

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$ .

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

#### Calculating the QR Decomposition

Typically the  $\mathbf{Q}$  matrix is computed by turning the columns of  $\mathbf{A}$  into an orthonormal basis. The most basic way to do this is using Gram-Schmidt; however, this is numerically unstable. Householder reflections and Givens Rotations give a numerically stable way to do this.

After this is done, we do  $\mathbf{R} = \mathbf{Q}^H \mathbf{A}$

#### Further References:

[Dr. Peyam's Orthogonality Playlist](#)

[MIT OCW 18.06 Lecture 17. Orthogonal Matrices and Gram-Schmidt](#)

### 3.4 Eigen Decomposition / Diagonalization

The eigendecomposition factors a matrix into a [similarity transform](#) to a diagonal matrix. Intuitively there exists a basis where  $\mathbf{A}$  is diagonal given by change of basis  $\mathbf{P}$ . The decomposition is unique up to the ordering of the columns of  $\mathbf{P}$  and  $\mathbf{D}$ .

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$ .

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

#### Existence and Uniqueness

Recall that a diagonal matrix is just constant scaling along a series of dimensions. If we can find  $n$  linearly independent vectors  $\mathbf{v}$  such that  $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ , we can diagonalize  $\mathbf{A}$  by doing a change of basis.

However, this is not always possible, so not all square matrices have an eigendecomposition.

For example take  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

In the above matrix, there is an eigenvalue of 1 with algebraic multiplicity 2 in the characteristic equation  $\mathbf{A} - \lambda\mathbf{I}_n$ . However, it is only associated with 1 eigenvector. All nondiagonalizable matrices suffer from this problem.

#### $\mathbf{A}^k$ using Diagonalization

One can quickly calculate  $\mathbf{A}^k$  using diagonalization.

$$\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}$$

#### Analytic Functions of a Matrix

This decomposition allows for one to easily calculate [analytic functions of matrices](#). Suppose  $f$  is an analytic function (i.e. it is given by a convergent Taylor series).

$$\begin{aligned} f(\mathbf{A}) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \mathbf{A}^k \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \\ &= \mathbf{P} \left( \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \mathbf{D}^k \right) \mathbf{P}^{-1} \\ &= \mathbf{P}f(\mathbf{D})\mathbf{P}^{-1} \end{aligned}$$

#### Further References:

[Eigenvectors and Eigenvalues — Chapter 14, Essence of linear algebra](#)

[MIT OCW 18.06 Lecture 21. Eigenvalues and Eigenvectors](#)

### 3.5 Jordan Decomposition and the Jordan Normal/Cannonical Form

The [Jordan Decomposition](#) generalizes the Eigen decomposition to matrices not similar to diagonal matrices. All square matrices have a Jordan decomposition. The Eigen decomposition is a special case of the Jordan decomposition.

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$ .

$$\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$$

$\mathbf{J}$  is given by the matrix direct sum of all Jordan blocks associated with the matrix.

Additionally,  $\mathbf{J} = \mathbf{D} + \mathbf{Z}$  where  $\mathbf{Z}$  is a [nilpotent matrix](#).

#### Analytic Functions of a Matrix for Jordan Forms

$$f(\mathbf{A}) = \mathbf{P}f(\mathbf{J})\mathbf{P}^{-1}$$

Let  $\mathbf{J}_i(\lambda_k)$  is the  $i$ th Jordan block associated with eigenvalue  $\lambda_k$ . Let  $\nu_k$  be the number of Jordan blocks associated with  $\lambda_k$

$$f(\mathbf{J}) = \bigoplus_{i=1}^m \bigoplus_{k=1}^{\nu_k} f(\mathbf{J}_k(\lambda_i))$$

Note below we differentiate with respect to  $\lambda_i$ .

$$f(\mathbf{J}_k(\lambda_i)) = \begin{bmatrix} \frac{f(\lambda_i)}{0!} & \frac{f'(\lambda_i)}{1!} & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{f^{(n)}(\lambda_i)}{n!} \\ 0 & \frac{f(\lambda_i)}{0!} & \frac{f'(\lambda_i)}{1!} & \cdots & \frac{f^{(n-1)}(\lambda_i)}{(n-1)!} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} \\ 0 & \cdots & \cdots & 0 & \frac{f(\lambda)}{0!} \end{bmatrix}$$

#### The Real Jordan Form

This is a quick summary of the proof presented by [this](#) Stack Exchange thread.

Suppose we have a complex eigenvector  $\mathbf{v}$  with complex eigenvalue  $\lambda$ . Express these as:

$$\mathbf{v} = \mathbf{x} + \mathbf{y}j, \quad \lambda = \alpha + \beta j$$

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

$$\mathbf{A}(\mathbf{x} + \mathbf{y}j) = (\alpha + \beta j)(\mathbf{x} + \mathbf{y}j)$$

$$\mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}j = (\alpha\mathbf{x} - \beta\mathbf{y}) + (\beta\mathbf{x} + \alpha\mathbf{y})j$$

Equating real with real and imaginary with imaginary parts:

$$\mathbf{A}\mathbf{x} = \alpha\mathbf{x} - \beta\mathbf{y}$$

$$\mathbf{A}\mathbf{y} = \beta\mathbf{x} + \alpha\mathbf{y}$$

Can we replace the complex conjugate eigenvector pair  $\mathbf{v}, \mathbf{v}^*$  with  $\mathbf{x}, \mathbf{y}$ ? In fact we can! It turns out  $\mathbf{x}$  and  $\mathbf{y}$  are linearly independent.



Suppose  $c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}$

Consider:

$$(c_1 + c_2j)\mathbf{v} = (c_1 + c_2j)(\mathbf{x} + \mathbf{y}j)$$

$$(c_1 + c_2j)\mathbf{v} = c_1\mathbf{x} - c_2\mathbf{y} + \cancel{(c_1\mathbf{x} + c_2\mathbf{y})j} \xrightarrow{\mathbf{0}}$$

Any scalar multiple of  $\mathbf{v}$  must also be an eigenvector

$(c_1 + c_2j)\mathbf{v}$  is a real eigenvector, which implies it has a real eigenvalue; however, this is a contradiction. Since it was a scalar multiple of  $\mathbf{v}$  we know it actually has a complex eigenvalue  $\lambda$ . Thus  $\mathbf{x}, \mathbf{y}$  are linearly independent.

Are  $\mathbf{x}, \mathbf{y}$  linearly independent of the other eigenvectors as well?

Suppose they weren't.

$$\mathbf{x} = c\mathbf{v}_k$$

Then we see that  $\mathbf{x}$  must be in the same eigenspace as  $\mathbf{v}_k$  and be associated with eigenvalue  $\lambda_k$ , but we know it's not, so  $\mathbf{x}, \mathbf{y}$  are linearly independent of the other eigenvectors. Suppose  $\mathbf{v}_1, \dots, \mathbf{v}_m$  are real eigenvalues and  $\mathbf{v}_{m+1}, \mathbf{v}_{m+1}^* \dots, \mathbf{v}_{(n-m)/2}$  are complex conjugate eigenvalues.

$$\text{Let } \mathbf{P} = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_m & \mathbf{x}_1 & \mathbf{y}_1 & \dots & \mathbf{x}_{(n-m)/2} & \mathbf{y}_{(n-m)/2} \end{bmatrix}$$

$$\mathbf{J}_r = \begin{bmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_m & & & & & \\ & & & \begin{pmatrix} \alpha_1 & -\beta_1 \\ \beta_1 & \alpha_1 \end{pmatrix} & & & & \\ & & & & \ddots & & & \\ & & & & & \begin{pmatrix} \alpha_{(n-m)/2} & -\beta_{(n-m)/2} \\ \beta_{(n-m)/2} & \alpha_{(n-m)/2} \end{pmatrix} & & \end{bmatrix}$$

$$\mathbf{A} = \mathbf{P}\mathbf{J}_r\mathbf{P}^{-1}$$

## Further References

[Texas AM University Jordan Normal Form Notes](#)

[MIT OCW 18.06 Lecture 28. Similar Matrices and Jordan Form](#)

[Sheldon and Axler Jordan Form](#)

### 3.6 Singular Value Decomposition

The [singular value decomposition](#) generalizes the Eigen Decomposition to nonsquare matrices and also has applications in optimal approximations. It factors a matrix in two the product of a unitary matrix, a diagonal matrix, and another unitary matrix.

$\Sigma$  is a diagonal matrix of singular values.

Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$

$\Sigma \in \mathbb{R}^{m \times n}$ ,  $\Sigma_r \in \mathbb{R}^{r \times r}$ ,

$\mathbf{U} \in \mathbb{R}^{m \times m}$ ,  $\mathbf{U}_r \in \mathbb{R}^{m \times r}$ ,  $\mathbf{U}_0 \in \mathbb{R}^{m \times (m-r)}$

$\mathbf{V} \in \mathbb{R}^{n \times n}$   $\mathbf{V}_r \in \mathbb{R}^{n \times r}$ ,  $\mathbf{V}_0 \in \mathbb{R}^{n \times (n-r)}$

$\mathbf{V}$  is a matrix of eigenvectors of  $\mathbf{A}^H \mathbf{A}$  (right singular vectors).

$\mathbf{U}$  is a matrix of eigenvectors of  $\mathbf{A} \mathbf{A}^H$  (left singular vectors).

$\Sigma$  is a matrix of the squared eigenvalues (singular values) of either  $\mathbf{A}^H \mathbf{A}$ ,  $\mathbf{A} \mathbf{A}^H$  on the diagonal (possibly augmented by 0s).

The SVD is unique up to the ordering of the singular values. In general, people order the singular values in descending order.

#### Relationship Between Full and Reduced SVD

The full SVD is given by:

$$\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^H$$

The reduced (or compact) SVD is given by:

$$\mathbf{A} = \mathbf{U}_r \Sigma_r \mathbf{V}_r^H$$

In this case  $\mathbf{U}_r$  and  $\mathbf{V}_r$  are not unitary; however, they have orthonormal columns.

They are related by:

$$\mathbf{A} = \begin{bmatrix} \mathbf{U}_r \\ \mathbf{U}_0 \end{bmatrix} \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_r \\ \mathbf{V}_0 \end{bmatrix}^H$$

Some properties:

$$\mathbf{x} \in \text{img}(\mathbf{A}) \iff \mathbf{x} \in \text{img}(\mathbf{U}_r)$$

$$\mathbf{x} \in \text{img}(\mathbf{A}^H) \iff \mathbf{x} \in \text{img}(\mathbf{V}_r)$$

$$\mathbf{x} \in \text{null}(\mathbf{A}) \iff \mathbf{x} \in \text{img}(\mathbf{V}_0)$$

$$\mathbf{x} \in \text{null}(\mathbf{A}^H) \iff \mathbf{x} \in \text{img}(\mathbf{U}_0)$$

### Norms and the SVD:

$$\sigma_{\max} = \|\mathbf{A}\|_2$$

$$\frac{1}{\sigma_{\min}} = \|\mathbf{A}^{-1}\|_2$$

### The Condition Number:

The [condition number](#) measures the sensitivity of a function

For square matrices

$$\kappa(\mathbf{A}) = \|\mathbf{A}^{-1}\|_2 \|\mathbf{A}\|_2 = \frac{\sigma_{\max}}{\sigma_{\min}}$$

We extend the definition to nonsquare matrices by defining:

$$\kappa(\mathbf{A}) := \frac{\sigma_{\max}}{\sigma_{\min}}$$

### Optimal Approximation:

[Moore Penrose Pseudoinverse](#) is efficiently computed by:

$\mathbf{A}^+ = \mathbf{U}\mathbf{\Sigma}^+\mathbf{V}^H$  where  $\mathbf{\Sigma}^+$  reciprocates nonzero values on the diagonal.

If  $\mathbf{A}$  is underdetermined  $\mathbf{A}^+\mathbf{y}$  gives the least norm solution.

If  $\mathbf{A}$  is overdetermined  $\mathbf{A}^+\mathbf{y}$  gives the least squares solution.

### Further References:

[MIT OCW 18.06 Lecture 29. Singular Value Decomposition](#)

[Prof. Steve Brunton's Lectures on the SVD](#)

### 3.7 Quadratic Forms

A [quadratic form](#) is simply an expression is a binomial where each term has degree 2.

All quadratic forms may be written as:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \text{ where } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{A} \in \mathbb{R}^{n \times n}$$

Example:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + a_{12}x_2^2 + (a_{12} + a_{21})x_1x_2$$

In general the cross terms are off diagonal terms and the pure terms are diagonal terms.

Positive definite quadratic forms may be written as a sum of positive squares. They are associated with positive definite matrices. A negative definite quadratic form is the sum of negative squares.

The above representation is not unique; however, it is unique over the set of symmetric matrices. For this reason, and for ease of analyzing eigenvalues and eigenvectors of  $\mathbf{A}$ , we prefer using the symmetric representation of a quadratic form.

There's a quick trick for turning a nonsymmetric representation into a symmetric representation.

**Lemma:**

$$\mathbf{A} = \mathbf{A}_s + \mathbf{A}_k \text{ where } \mathbf{A}_s^T = \mathbf{A}_s \text{ and } \mathbf{A}_k^T = -\mathbf{A}_k$$

**Proof**

Notice that:

$$\mathbf{A} + \mathbf{A}^T \text{ is symmetric.}$$

$$\mathbf{A} - \mathbf{A}^T \text{ is skew-symmetric.}$$

$$\mathbf{A} = \underbrace{\frac{1}{2}(\mathbf{A} + \mathbf{A}^T)}_{\mathbf{A}_s} + \underbrace{\frac{1}{2}(\mathbf{A} - \mathbf{A}^T)}_{\mathbf{A}_k}$$

—

$$\mathbf{x}^T (\mathbf{A}_s + \mathbf{A}_k) \mathbf{x} = \mathbf{x}^T \mathbf{A}_s \mathbf{x} + \mathbf{x}^T \mathbf{A}_k \mathbf{x}$$

Notice  $\mathbf{A}_o$  is zero on the diagonal and since it's skewsymmetric, off diagonal terms will come in positive and negative pairs. We can conclude from this that all terms are zero.

Thus the quadratic form only depends on the symmetric part of  $\mathbf{A}$ :

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{A}_s \mathbf{x}$$

## 4 Linear Systems Theory

### 4.1 Continuous Time Fourier Transform

$$X(e^{j\omega}) = \mathcal{F}\{x(t)\} := \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad x(t) = \mathcal{F}^{-1}\{X(e^{j\omega})\} := \frac{1}{2\pi} \int_{-\infty}^{\infty} X(e^{j\omega})e^{j\omega t} d\omega$$

#### 4.1.1 CTFT Properties

Property	Time Domain	Frequency Domain
Linearity	$c_1x_1(t) + c_2x_2(t)$	$c_1X_1(e^{j\omega}) + c_2X_2(e^{j\omega})$
Time Shift	$x(t - t_0)$	$X(e^{j\omega})e^{-j\omega t_0}$
Time Reversal	$x(-t)$	$X(e^{-j\omega})$
Time Scale	$x(at)$	$\frac{1}{ a }X\left(e^{j\frac{\omega}{a}}\right)e^{-\frac{j\omega}{a}}$
Time Scale + Shift	$x(at - t_0)$	$\frac{1}{ a }X\left(e^{j\frac{\omega}{a}}\right)e^{-\frac{j\omega t_0}{a}}$
Complex Conjugation	$x^*(t)$	$X^*(e^{-j\omega})$
Frequency Shift	$x(t)e^{-j\omega_0 t}$	$X(e^{j(\omega - \omega_0)})$
Convolution	$x_1(t) * x_2(t)$	$X_1(e^{j\omega})X_2(e^{j\omega})$
Modulation	$x_1(t)x_2(t)$	$\frac{1}{2\pi}X_1(e^{j\omega}) * X_2(e^{j\omega})$
Time Differentiation	$\mathcal{D}_t^n\{x(t)\}$	$(j\omega)^n X(e^{j\omega})$
Causal Accumulation	$\mathcal{A}_t\{x(t)u_H(t)\}$	$\frac{1}{j\omega}X(e^{j\omega}) + \pi X(0)\delta(\omega)$
Frequency Differentiation	$t^n x(t)$	$j^n \mathcal{D}_\omega^n\{X(e^{j\omega})\}$
Parseval's Theorem	$\int_{-\infty}^{\infty} x_1(t)x_2^*(t) dt$	$\frac{1}{2\pi} \int_{-\infty}^{\infty} X_1(e^{j\omega})X_2^*(e^{j\omega}) d\omega$
Duality	$X(e^{j\omega t})$	$2\pi x(\omega)$

#### 4.1.2 CTFT Table

$f(t)$	$F(e^{j\omega})$
0	constant $\omega_0$
1	$2\pi\delta(\omega)$
$\delta(t)$	1
$u(t)$	$\pi\delta(\omega) + \frac{1}{j\omega}$
$\text{sgn}(t)$	$\frac{2}{j\omega}$
$e^{j\omega_0 t}$	$2\pi\delta(\omega - \omega_0)$
$\cos(\omega_0 t)$	$\pi(\delta(\omega - \omega_0) + \delta(\omega + \omega_0))$
$\sin(\omega_0 t)$	$-j\pi(\delta(\omega - \omega_0) - \delta(\omega + \omega_0))$
$\cos(\omega_0)u(t)$	$\frac{\pi}{2}(\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) + \frac{j\omega}{\omega_0^2 - \omega^2}$
$\sin(\omega_0)u(t)$	$\frac{\pi}{2j}(\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) + \frac{\omega_0}{\omega_0^2 - \omega^2}$
$e^{-at} \cos(\omega_0)u(t)$	$\frac{\omega_0}{(a + j\omega)^2 + \omega_0^2}$
$e^{-at} \sin(\omega_0)u(t)$	$\frac{a + j\omega}{(a + j\omega)^2 + \omega_0^2}$
$\text{rect}\left(\frac{t}{T}\right), T > 0$	$T \text{sinc}\left(\frac{T}{2\pi}\omega\right)$
$\text{tri}\left(\frac{t}{T}\right), T > 0$	$T \text{sinc}^2\left(\frac{T}{2\pi}\omega\right)$
$\text{Asinc}\left(\frac{1}{2\pi A}t\right)$	$\text{rect}(A\omega)$
$\text{Asinc}^2\left(\frac{1}{2\pi A}t\right)$	$\text{tri}(A\omega)$
$x^n$	$2\pi j^n \delta^{(n)}(\omega)$
$e^{-at}u(t), a > 0$	$\frac{1}{a + j\omega}$
$t^n e^{-at}u(t), a > 0$	$\frac{1}{(a + j\omega)^{n+1}}$
$e^{-a t }, a > 0$	$\frac{2a}{a^2 + \omega^2}$
$e^{-\frac{t^2}{2\sigma^2}}$	$\sqrt{2\pi\sigma^2}e^{-\frac{\sigma^2}{2}\omega^2}$

## 4.2 Laplace Transform

$$X(s) = \mathcal{L}\{x(t)\} := \int_{-\infty}^{\infty} x(t)e^{-st} dt$$

$$x(t) = \mathcal{L}^{-1}\{X(s)\} := \int_{-\infty}^{\infty} x(t)e^{-st} dt$$

### 4.2.1 Laplace Transform Properties

Property	$x[n]$	$X(z)$
Linearity	$c_1x_1(t) + c_2x_2(t)$	$c_1X_1(s) + c_2X_2(s)$
Time Shifting	$x(t - t_0)u(t - t_0)$	$X(s)e^{-st_0}$
Time Scaling	$x(at)$	$\frac{1}{ a }X\left(\frac{s}{a}\right)$
Time Transformation	$x(at - t_0)u(t - t_0)$	$\frac{1}{ a }X\left(\frac{s}{a}\right)e^{-\frac{st_0}{a}}$
Frequency Shift	$e^{at}x(t)$	$X(s - a)$
1st Time Derivative	$x'(t)$	$sX(s) - x(0^-)$
2nd Time Derivative	$x''(t)$	$s^2X(s) - sx(0^-) - x'(0^-)$
General Time Derivative (One-Sided)	$\mathcal{D}_t^n\{x(t)u(t)\}$	$s^nX(s) - \sum_{k=1}^n s^{n-k}\mathcal{D}_t^{k-1}\{x\}(0^-)$
General Time Derivative (Two-Sided)	$\mathcal{D}_t^n\{x(t)\}$	$s^nX(s)$
1st Frequency Derivative	$tx(t)$	$-X'(s)$
General Frequency Derivative	$t^n x(t)$	$(-1)^n \mathcal{D}_s^n\{X(s)\}$
Accumulation	$\mathcal{A}_t^n\{x(t)\}$	$s^{-n}X(s)$
Frequency Integration	$\frac{1}{t^n}x(t)$	$\mathcal{I}_s^n\{X(s)\}$
Convolution	$x_1(t) * x_2(t)$	$X_1(s)X_2(s)$
Modulation	$x_1(t)x_2(t)$	$\frac{1}{2\pi j} \lim_{\omega \rightarrow \infty} \int_{\sigma - j\omega}^{\sigma + j\omega} X_1(u)X_2(s - u) du$
Complex Conjugation	$x^*(t)$	$X^*(s^*)$

**Initial Value Theorem:**

$$x(t) \text{ is causal} \implies x(0^+) = \lim_{s \rightarrow \infty} sX(s)$$

**Final Value Theorem:**

$$x(t) \text{ is causal and stable} \implies \lim_{t \rightarrow \infty} x(t) = \lim_{s \rightarrow 0} sX(s)$$

### 4.2.2 Laplace Transform Table

$x(t)$	$X(s)$	ROC
$\delta(t)$	1	$\mathbb{C}$
$\delta(t - t_0)$	$e^{-st_0}$	$\mathbb{C}$
$u_H(t)$	$\frac{1}{s}$	$\Re\{s\} > 0$
$u_H(t - t_0)$	$\frac{1}{s}e^{-st_0}$	$\Re\{s\} > 0$
$\text{rect}\left(\frac{t}{T}\right)$	$\frac{1 - e^{-Ts}}{s}$	$\Re\{s\} > 0$
$\text{relu}(t)$	$\frac{1}{s^2}$	$\Re\{s\} > 0$
$t^n u_H(t), n \in \mathbb{N}_0$	$\frac{n!}{s^{n+1}}$	$\Re\{s\} > 0$
$t^z u_H(t), \Re\{z\} > -1$	$\frac{\Gamma(z+1)}{s^{z+1}}$	$\Re\{s\} > 0$
$t^{\frac{1}{n}} u_H(t), n \in \mathbb{N}_0$	$\frac{\Gamma(\frac{1}{n}+1)}{s^{\frac{1}{n}+1}}$	$\Re\{s\} > 0$
$e^{-at}u(t)$	$\frac{1}{s+a}$	$\Re\{s\} > -a$
$e^{-a t }$	$\frac{2a}{(a^2 - s^2)}$	$-a < \Re\{s\} < a$
$(1 - e^{-at})u(t)$	$\frac{a}{s(s+a)}$	$\Re\{s\} > 0$
$\sin(\omega t)u(t)$	$\frac{\omega}{s^2 + \omega^2}$	$\Re\{s\} > 0$
$\cos(\omega t)u(t)$	$\frac{s}{s^2 + \omega^2}$	$\Re\{s\} > 0$
$e^{-at} \sin(\omega t)u(t)$	$\frac{\omega}{(s+a)^2 + \omega^2}$	$\Re\{s\} > -a$
$e^{-at} \cos(\omega t)u(t)$	$\frac{s+a}{(s+a)^2 + \omega^2}$	$\Re\{s\} > -a$



### 4.3 Matrix Calculus

Linear transformations commute with differential operators (derivative, integral, CT Fourier transform, Laplace transform).

$$\mathcal{D}\{\mathbf{A}\mathbf{f}(\mathbf{x})\} = \mathbf{A}(\mathcal{D}\mathbf{f}(\mathbf{x}))$$

$$\int \mathbf{A}\mathbf{f}(t) dt = \mathbf{A} \int \mathbf{f}(t) dt$$

$$\mathcal{F}\{\mathbf{A}\mathbf{f}(\mathbf{x})\} = \mathbf{A}\mathcal{F}\{\mathbf{f}(\mathbf{x})\}$$

$$\mathcal{L}\{\mathbf{A}\mathbf{f}(\mathbf{x})\} = \mathbf{A}\mathcal{L}\{\mathbf{f}(\mathbf{x})\}$$

#### 4.3.1 Matrix Derivatives

The following derivatives can be derived using the following definition.

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + \mathbf{h}) - \mathbf{f}(\mathbf{x})}{\|\mathbf{h}\|} - \mathbf{L}(\mathbf{h}) = \mathbf{0} \iff \mathbf{L}(\mathbf{h}) \text{ is the derivative of } \mathbf{f}(\mathbf{x})$$

$$\mathcal{D}\{(\mathbf{x}(t))^{\top}\} = (\mathcal{D}\mathbf{x}(t))^{\top}$$

$$\mathcal{D}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}$$

$$\mathcal{D}_{\mathbf{x}}\{\mathbf{A}\mathbf{x}\} = \mathbf{A}$$

$$\mathcal{D}_{\mathbf{x}}\{\mathbf{x}^{\top}\mathbf{x}\} = \mathcal{D}\|\mathbf{x}\|_2 = 2\mathbf{x}^{\top}$$

$$\mathcal{D}_{\mathbf{x}}\{\mathbf{x}^{\top}\mathbf{A}\mathbf{x}\} = \mathbf{x}^{\top}(\mathbf{A} + \mathbf{A}^{\top})$$

## 4.4 State Space Representation

Suppose you have a system with  $n$  states,  $m$  outputs and  $p$  inputs.

Define:

$$\mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^p, \mathbf{y} \in \mathbb{R}^m$$

$$\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}, \mathbf{C} \in \mathbb{R}^{m \times n}, \mathbf{D} \in \mathbb{R}^{m \times p}$$

The state space representation is given by:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}$$

Note:

Some authors use  $\mathbf{q}$  instead of  $\mathbf{x}$

### 4.4.1 State Space Coordinate Systems

The state space representation is not unique. It is dependent on your choice of state variables. The choice of state variables can be seen as a choice made on the coordinates of the state space.

Suppose:

$$\mathbf{q} = \mathbf{P}\mathbf{x} \text{ where } \mathbf{P} \text{ is invertible.}$$

We can view  $\mathbf{P}$  as a change of basis matrix from basis  $\mathcal{B}_1$  to  $\mathcal{B}_2$ .  $\mathbf{q}$  is simply  $\mathbf{x}$  under a different choice of coordinates.

$$\underbrace{\mathbf{x}_{\mathcal{B}_2}}_{\mathbf{q}} = \mathbf{P}_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} \mathbf{x}_{\mathcal{B}_1}$$

How does this affect the state space representation? Substitute with  $\mathbf{x} = \mathbf{P}^{-1}\mathbf{q}$

$$\mathbf{P}^{-1}\dot{\mathbf{q}} = \mathbf{A}\mathbf{P}^{-1}\mathbf{q} + \mathbf{B}\mathbf{u}$$

$$\mathbf{y} = \mathbf{C}\mathbf{P}^{-1}\mathbf{q} + \mathbf{D}\mathbf{u}$$

Right multiplying by  $\mathbf{P}$ , we have:

$$\dot{\mathbf{q}} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1}\mathbf{q} + \mathbf{P}\mathbf{B}\mathbf{u}$$

$$\mathbf{y} = \mathbf{C}\mathbf{P}^{-1}\mathbf{q} + \mathbf{D}\mathbf{u}$$

## 4.5 The Transfer Function

The transfer function,  $\mathbf{H}(s)$ , transforms input to output in the s-domain given 0 input. For LTI Systems, the transfer function is a matrix dependent on  $s$ .

$$\mathbf{H}(s) = \mathbf{C}\Phi(s)\mathbf{B} + \mathbf{D}$$

$$\Phi(s) = (s\mathbf{I}_n - \mathbf{A})^{-1}$$

$\Phi(t) = e^{\mathbf{A}t}$  is called the state transition matrix. It describes the evolution of the system in state space from an arbitrary initial state with zero input. In the language of ordinary differential equations  $\Phi(t)$  describes the family of homogenous solutions.

$$\mathcal{L}\{\Phi(t)\} = \Phi(s) = (s\mathbf{I}_n - \mathbf{A})^{-1}$$

So we can see that the state transition matrix is not defined for eigenvalues of  $\mathbf{A}$ .

### Derivation:

In the following, we assume  $\mathbf{x}(t)$  to be causal.

$$\mathcal{L}\{\dot{\mathbf{x}}\} = \mathcal{L}\{\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}\}$$

$$s\mathbf{X}(s) - \mathbf{x}_0 = \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s)$$

$$s\mathbf{X}(s) - \mathbf{A}\mathbf{X}(s) - \mathbf{x}_0 = \mathbf{B}\mathbf{U}(s)$$

$$\mathbf{X}(s)(s\mathbf{I}_n - \mathbf{A}) - \mathbf{x}_0 = \mathbf{B}\mathbf{U}(s)$$

$$\mathbf{X}(s)(s\mathbf{I}_n - \mathbf{A}) = \mathbf{B}\mathbf{U}(s) + \mathbf{x}_0$$

$$\mathbf{X}(s) = (s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s) + (s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{x}_0$$

$$\mathbf{X}(s) = \Phi(s)\mathbf{B}\mathbf{U}(s) + \Phi(s)\mathbf{x}_0$$

$$\mathcal{L}\{\mathbf{y}\} = \mathcal{L}\{\mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\}$$

$$\mathbf{Y}(s) = \mathbf{C}[\Phi(s)\mathbf{B}\mathbf{U}(s) + \Phi(s)\mathbf{x}_0] + \mathbf{D}\mathbf{U}(s)$$

$$\mathbf{Y}(s) = \mathbf{C}\Phi(s)\mathbf{B}\mathbf{U}(s) + \mathbf{C}\Phi(s)\mathbf{x}_0 + \mathbf{D}\mathbf{U}(s)$$

$$\mathbf{Y}(s) = \mathbf{C}\Phi(s)\mathbf{B}\mathbf{U}(s) + \mathbf{D}\mathbf{U}(s) + \mathbf{C}\Phi(s)\mathbf{x}_0$$

$$\mathbf{Y}(s) = [\mathbf{C}\Phi(s)\mathbf{B} + \mathbf{D}]\mathbf{U}(s) + \mathbf{C}\Phi(s)\mathbf{x}_0$$

$$\mathbf{Y}(s) = \mathbf{H}(s)\mathbf{U}(s) + \mathbf{C}\Phi(s)\mathbf{x}_0$$

#### 4.5.1 Analytic Description of the Transfer Function

Notice that the transfer function involves taking the inverse of the Laplace transform of the state transition matrix  $\Phi(s)$ . We can do this analytically using the [adjugate](#).

$$\begin{aligned}\Phi^{-1}(s) &= \frac{1}{\det(\Phi(s))} \text{adj}(\Phi(s)) \\ &= \frac{1}{\det(s\mathbf{I}_n - \mathbf{A})} \text{adj}(\Phi(s)) \\ &= \frac{1}{\chi_{\mathbf{A}}(s)} \text{adj}(\Phi(s))\end{aligned}$$

Recall that the entries of the adjugate matrix are cofactors of  $\Phi(s)$ . This means that they are polynomials of  $s$ , but with degree of 1 less than the characteristic polynomial.

This means that so long as  $\mathbf{D} = [0]$ , all entries of the transfer function matrix are strictly proper rational.

Notice also the poles of the transfer function must be eigenvalues of  $\mathbf{A}$ . But eigenvalues are independent of coordinates (i.e.  $\mathbf{PAP}^{-1}$  has the same eigenvalues as  $\mathbf{A}$ ). This leads us to conclude that the poles of the transfer function are also independent of coordinates. Is the transfer function itself totally coordinate independent then? Proposition 3.5.1 says yes.

##### Important Caveat:

While the poles of the transfer function must be eigenvalues of  $\mathbf{A}$ , the converse is not true. Not all eigenvalues of  $\mathbf{A}$  are poles of the transfer function. Why is this the case? If there's a zero with exactly the same value as an eigenvalue of  $\mathbf{A}$ , we'll get a pole-zero cancellation.

#### Proposition 3.5.1: The Transfer Function is Coordinate Independent

Under no transformations:

$$\mathbf{H}(s) = \mathbf{C}\Phi(s)\mathbf{B} + \mathbf{D}$$

But recall under  $\mathbf{q} = \mathbf{Ax}$

$$\dot{\mathbf{q}} = \mathbf{PAP}^{-1}\mathbf{q} + \mathbf{PBu}$$

$$\mathbf{y} = \mathbf{CP}^{-1}\mathbf{q} + \mathbf{Du}$$

Performing the appropriate substitutions:

$$\begin{aligned}\mathbf{H}(s) &= (\mathbf{CP}^{-1})(s\mathbf{I}_n - \mathbf{PAP}^{-1})^{-1}(\mathbf{PB}) + \mathbf{D} \\ &= \mathbf{C}(\mathbf{P}^{-1}(s\mathbf{I}_n - \mathbf{PAP}^{-1})\mathbf{P})^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}(s\mathbf{P}^{-1}\mathbf{I}_n\mathbf{P} - \cancel{\mathbf{P}^{-1}\mathbf{P}}\mathbf{A}\cancel{\mathbf{P}^{-1}\mathbf{P}})^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}(s\mathbf{P}^{-1}\mathbf{I}_n\mathbf{P} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}(s\mathbf{I}_n\cancel{\mathbf{P}^{-1}\mathbf{P}} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{C}(s\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}\end{aligned}$$

#### 4.5.2 The Impulse Response and Time Domain Solution

$$\begin{aligned}
h(t) &= \mathcal{L}^{-1}\{G(s)\} \\
&= \mathcal{L}^{-1}\{C\Phi(s)B + D\} \\
&= C\mathcal{L}^{-1}\{\Phi(s)\}B + \mathcal{L}^{-1}\{D\} \\
&= C\Phi(t)B + D\delta(t)
\end{aligned}$$

We can use this to derive the solution to the system of ODEs.

$$\begin{aligned}
y(t) &= \mathcal{L}^{-1}\{H(s)U(s) + C\Phi(s)x_0\} \\
&= \mathcal{L}^{-1}\{H(s)U(s)\} + \mathcal{L}^{-1}\{C\Phi(s)x_0\} \\
&= h(t) * u(t) + C\Phi(t)x_0 \\
&= \int_0^\infty h(\tau)u(t - \tau) d\tau + C\Phi(t)x_0 \\
&= \int_0^\infty (C\Phi(\tau)B + D\delta(\tau))u(t - \tau) d\tau + C\Phi(t)x_0 \\
&= \int_0^\infty C\Phi(\tau)Bu(t - \tau) d\tau + \int_0^\infty D\delta(\tau)u(t - \tau) d\tau + C\Phi(t)x_0 \\
&= \int_0^\infty C\Phi(\tau)Bu(t - \tau) d\tau + D \int_0^\infty \delta(\tau)u(t - \tau) d\tau + C\Phi(t)x_0 \\
&= \int_0^\infty C\Phi(\tau)Bu(t - \tau) d\tau + Du(t) + C\Phi(t)x_0
\end{aligned}$$

## 4.6 State Space Realizations of Transfer Functions

Since the state space representation of a system is not coordinate independent, we will see that analyzing the transfer function becomes a key part of analyzing systems. However, when we wish to do practical applications of the system, we will need to go back to a state space representation.

How should we choose the state space realization given a transfer function? There are a couple standard forms: the controllable canonical form and the observable canonical form.

### 4.6.1 SISO Systems

$$H(s) = \frac{\sum_{k=0}^{\nu-1} \beta_k s^k}{s^n + \sum_{k=0}^{\nu-1} \alpha_k s^k} + d$$

#### Controllable Canonical Form (CCF)

$$\mathbf{A}_c = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & & \dots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{\nu-2} & -\alpha_{\nu-1} \end{bmatrix}, \mathbf{B}_c = \mathbf{e}_\nu, \mathbf{C}_c = [\beta_0 \quad \dots \quad \beta_{\nu-1}], \mathbf{D}_c = d$$

#### Observable Canonical Form (OCF)

$$\mathbf{A}_o = \begin{bmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & \vdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \ddots & -\alpha_{\nu-2} \\ 0 & 0 & \dots & 1 & -\alpha_{\nu-1} \end{bmatrix}, \mathbf{B}_o = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_{\nu-1} \end{bmatrix}, \mathbf{C}_o = \mathbf{e}_\nu^\top, \mathbf{D}_o = d$$

Astute readers may notice that  $\mathbf{A}_o$  is the companion matrix. See [Texas A & M Chapter 8 Jordan Normal Form Theorem 8.1.2](#) on page 222.

#### 4.6.2 MIMO Systems

$$H(s) = \frac{1}{\alpha(s)}\beta(s) + D = \underbrace{\alpha^{-1}(s)}_{\frac{1}{\alpha(s)}\mathbf{I}_m}\beta(s) + D$$

We may express each  $\alpha(s)$  and  $\beta(s)$  as polynomials of  $s$  (termwise for the matrix).

$$\alpha(s) = \sum_{k=0}^{\nu} s^k \alpha_k, \quad \beta(s) = \sum_{k=0}^{\nu-1} s^k \beta_k \text{ where } \nu = \deg(\alpha(s))$$

##### Controllable Canonical Form (CCF)

$$\mathbf{A}_c = \begin{bmatrix} 0 & \mathbf{I}_m & 0 & \dots & 0 & 0 \\ 0 & 0 & \mathbf{I}_m & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & & \dots & 0 & \mathbf{I}_m \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{\nu-2} & -\alpha_{\nu-1} \end{bmatrix}, \quad \mathbf{B}_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{I}_m \end{bmatrix}, \quad \mathbf{C}_c = [\beta_0 \quad \dots \quad \beta_{\nu-1}], \quad \mathbf{D}_c = D$$

##### Controllable Canonical Form (OCF)

$$\mathbf{A}_o = \begin{bmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ \mathbf{I}_m & 0 & \dots & 0 & -\alpha_1 \\ 0 & \mathbf{I}_m & \vdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \ddots & -\alpha_{\nu-2} \\ 0 & 0 & \dots & \mathbf{I}_m & -\alpha_{\nu-1} \end{bmatrix}, \quad \mathbf{B}_o = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_{\nu-1} \end{bmatrix}, \quad \mathbf{C}_o = [0 \quad \dots \quad 0 \quad \mathbf{I}_m], \quad \mathbf{D}_o = D$$

#### 4.7 Relationship Between CCF and OCF

$$\mathbf{A}_o = \mathbf{A}_c^\top$$

$$\mathbf{B}_o = \mathbf{C}_c^\top$$

$$\mathbf{C}_o = \mathbf{B}_c^\top$$

$$\mathbf{D}_o = \mathbf{D}_c^\top$$

## 4.8 Equilibria and Stability

### 4.8.1 Definition of Equilibrium State

Suppose a system is given by the state space representation:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

The system stops evolving when  $\mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{0}$ . This corresponds to the nullspace of  $\mathbf{f}$ . We call elements of the nullspace  $(\mathbf{x}_{\text{eq}}, \mathbf{u}_{\text{eq}})$  equilibrium states.

If an equilibrium has  $\mathbf{u}_{\text{eq}} = \mathbf{0}$ , we categorize it as an unforced equilibrium.

If an equilibrium has  $\mathbf{u}_{\text{eq}} \neq \mathbf{0}$ , we categorize it as a forced equilibrium.

### 4.8.2 Equilibria of LTI Systems

Now consider a LTI system given by:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

We can see that unforced equilibria are given by  $\text{null}(\mathbf{A})$ .

### 4.8.3 Stability of LTI Systems

Unlike nonlinear systems, we can infer global behavior of an LTI system from their eigenvalues. For this reason, we can simply talk about the stability of an LTI system rather than stability around equilibria. The classification of LTI systems is summarized below.

Classification	Condition on $\lambda \in \Lambda()$
Globally Exponentially Stable	$\Re\{\lambda\} < 0$
Marginally Stable	$\Re\{\lambda\} \leq 0$ and no $\lambda$ is defective
(Marginally) Unstable	$\Re\{\lambda\} \leq 0$ and a $\lambda = 0$ is defective
Unstable	$\exists \lambda : \Re\{\lambda\} > 0$

### 4.8.4 Classification of Equilibria

For nonlinear systems, inferring global behavior can be complicated. We instead focus on behavior about equilibria. Not all equilibria are created equal. Some equilibria are stable and others are unstable. States around unstable equilibria are repelled. Stable equilibria do not repel states. Depending on your definition, stable equilibria may be strictly attractive or simply neither attract nor repel.

In order to classify equilibria, we linearize the system about a particular equilibrium as a Jacobian matrix and then analyze its eigenvalues. We summarize the classification in the table below.

Classification	Condition on $\lambda \in \Lambda([\mathcal{D}f(\mathbf{x}_{\text{eq}})])$
(Locally) Asymptotically Stable	$\Re\{\lambda\} < 0$
Unstable	$\exists \lambda : \Re\{\lambda\} > 0$
Inconclusive	$\Re\{\lambda\} \leq 0$ and $\exists \lambda : \lambda = 0$



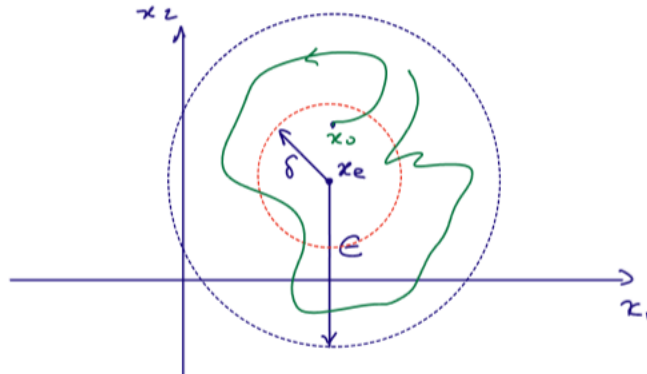
#### 4.8.5 Lyapunov Stability

Let  $\mathbf{x}_{eq} : \mathbf{f}(\mathbf{x}_{eq}) = \mathbf{0}$

The system defined by  $\mathbf{f}$  is said to be Lyapunov stable about  $\mathbf{x}_{eq}$  iff:

$$\forall \epsilon > 0, \exists \delta > 0 : \|\mathbf{x}_0 - \mathbf{x}_{eq}\| < \delta \implies \|\mathbf{x}(t) - \mathbf{x}_{eq}\| < \epsilon$$

Presented below is a diagram of showing parts of the definition



#### 4.8.6 Asymptotic Stability and Exponential Stability

The system defined by  $\mathbf{f}$  is said to be asymptotically stable about  $\mathbf{x}_{eq}$  iff:

$$\forall \epsilon > 0, \exists \delta > 0 : \|\mathbf{x}_0 - \mathbf{x}_{eq}\| < \delta \implies \lim_{t \rightarrow \infty} \|\mathbf{x}(t) - \mathbf{x}_{eq}\| = 0$$

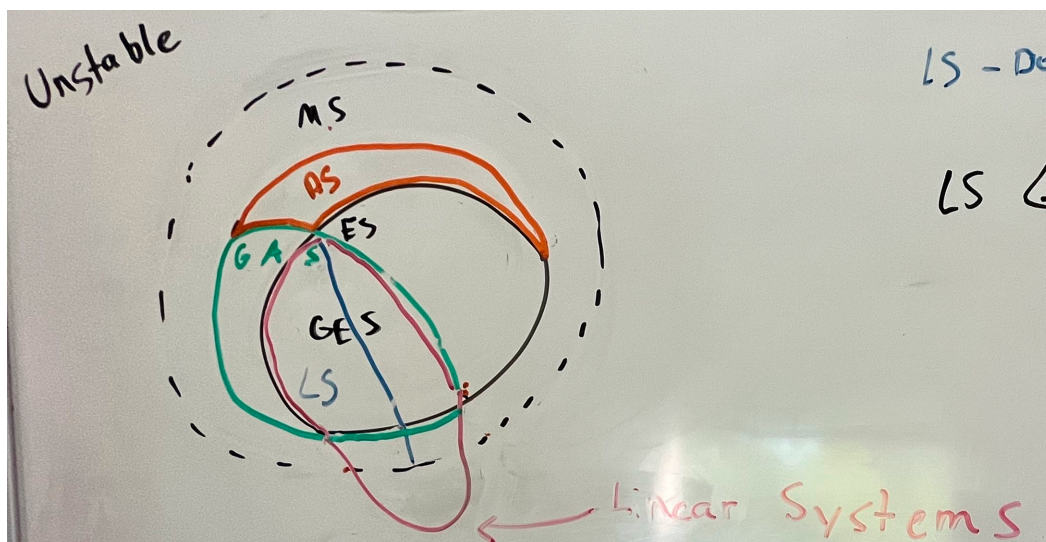
##### Exponentially Stable

A special type of asymptotic stability is exponential stability.

The system defined by  $\mathbf{f}$  is said to be exponentially stable about  $\mathbf{x}_{eq}$  iff:

$$\forall \epsilon > 0, \exists \delta > 0 : \|\mathbf{x}_0 - \mathbf{x}_{eq}\| < \delta \implies \|\mathbf{x}(t) - \mathbf{x}_{eq}\| \leq \alpha \|\mathbf{x}_0 - \mathbf{x}_{eq}\| e^{-\beta t}$$

#### 4.8.7 Diagram of Stability Types

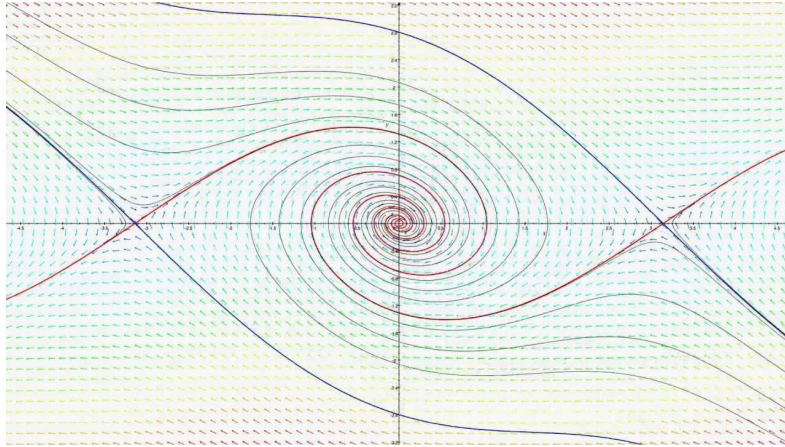


## 4.9 Example: Equilibria Classification of a Damped Pendulum

A damped pendulum with 0 input has the following state space equation:

$$\begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} \dot{\theta} \\ -\frac{g}{L} \sin(\theta) - b\dot{\theta} \end{bmatrix}$$

We can plot this vector field as a phase diagram in phase space as shown below. Notice that the equilibrium at  $\begin{bmatrix} \pi \\ 0 \end{bmatrix}$  is unstable (states are repelled away from it) while the equilibrium at  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$  is stable (states are drawn to it).



Now let's analyze this more systematically:

$$[\mathcal{D}f(\mathbf{x})] = \begin{bmatrix} 0 & 1 \\ -\frac{g}{L} \cos(\theta) & -b \end{bmatrix}$$

But the choice of units is arbitrary. Suppose we chose units such that  $\frac{g}{L} = 1$ ,  $b = 1$ .

$$[\mathcal{D}f(\mathbf{x})] = \begin{bmatrix} 0 & 1 \\ -\cos(\theta) & -1 \end{bmatrix}$$

Equilibrium at  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ .

$[\mathcal{D}f(\mathbf{x})] = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$  has eigenvalues  $-\frac{1+j\sqrt{3}}{2}$ ,  $-\frac{1-j\sqrt{3}}{2}$ . Thus we can see that the system is asymptotically stable about this equilibrium since both real parts are less than 1.

Equilibrium at  $\begin{bmatrix} \pi \\ 0 \end{bmatrix}$ .

$[\mathcal{D}f(\mathbf{x})] = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$  which has eigenvalues  $\frac{-1+\sqrt{5}}{2}$ ,  $\frac{-1-\sqrt{5}}{2}$ . We can see one eigenvalue has real part larger than 1 and thus the system is unstable about this equilibrium.

#### 4.10 Example: Rotation + Scaling

"I'll try spinning, that's a good trick!"

A matrix of the form below is called a rotation scaling matrix:

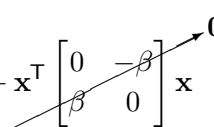
$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$$

Why? Let  $\theta = \tan^{-1}\left(\frac{\beta}{\alpha}\right)$ . We now rewrite the matrix as.

$$\frac{1}{\sqrt{\alpha^2 + \beta^2}} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Now consider:

$$\dot{\mathbf{x}} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \mathbf{x}$$

$$\begin{aligned} \mathcal{D}\{\mathbf{x}^\top \mathbf{x}\} &= 2\mathbf{x}^\top \dot{\mathbf{x}} \\ &= 2\mathbf{x}^\top \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \mathbf{x} \\ &= 2\mathbf{x}^\top \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \mathbf{x} + \cancel{\mathbf{x}^\top \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix} \mathbf{x}} \\ &= 2\alpha \|\mathbf{x}\|^2 \end{aligned}$$


This means  $\|\mathbf{x}(t)\|_2 = e^{\alpha t} \|\mathbf{x}_0\|_2$  which shows this system is globally exponentially stable.

Intuitively, the antidiagonal terms go to 0 because they represent instantaneous rotation by 90 degrees. This corresponds to continuous turning or circular motion (which won't affect the length). You may find this property useful when using the Real Jordan Form (see 3.5).

We could've applied this if we analyzed the damped pendulum using the real Jordan Form around its equilibrium to show the rotation and scaling around the origin (revealed in the phase plot).

## 4.11 The Lyapunov Function

### 4.11.1 Derivation

We saw in the rotation + scaling example that we could use a norm to establish stability properties. We now generalize this approach using Lyapunov functions. These are norms on state space.

$$V(\mathbf{x}) := \|\mathbf{x}\|_P^2 = \mathbf{x}^\top \mathbf{P} \mathbf{x}, \text{ where } \mathbf{P} \in \mathbb{W}_+^n$$

$$\begin{aligned} \dot{V} &= \mathcal{D}\{\mathbf{x}^\top\} \mathbf{P} \mathbf{x} + \mathbf{x}^\top \mathcal{D}\{\mathbf{P} \mathbf{x}\} \\ &= \dot{\mathbf{x}}^\top \mathbf{P} \mathbf{x} + \mathbf{x}^\top \mathbf{P} \dot{\mathbf{x}} \\ &= (\mathbf{A} \mathbf{x})^\top \mathbf{P} \mathbf{x} + \mathbf{x}^\top \mathbf{P} \mathbf{A} \mathbf{x} \\ &= \mathbf{x}^\top \mathbf{A}^\top \mathbf{P} \mathbf{x} + \mathbf{x}^\top \mathbf{P} \mathbf{A} \mathbf{x} \\ &= \mathbf{x}^\top \underbrace{(\mathbf{A}^\top \mathbf{P} + \mathbf{P} \mathbf{A})}_{-\mathbf{Q}} \mathbf{x} \\ &= -\mathbf{x}^\top \mathbf{Q} \mathbf{x} \end{aligned}$$

Thus  $\mathbf{Q} \succ 0 \implies \dot{V} < 0$ .

Some authors define the Lyapunov operator as  $\mathbf{L}_A(\mathbf{P}) = \mathbf{A}^\top \mathbf{P} + \mathbf{P} \mathbf{A}$

### 4.11.2 Application to Establishing GES of Asymptotically Stable LTI Systems

Recall a norm with respect to a matrix is bounded by its eigenvalues times the 2 norm.

$$\begin{aligned} -\lambda_{\max}(\mathbf{W}_1) \|\mathbf{x}\|^2 &\leq \dot{V}(\mathbf{x}) \leq -\lambda_{\min}(\mathbf{W}_1) \|\mathbf{x}\|^2 \\ \implies \dot{V} &\leq -\lambda_{\min}(\mathbf{W}_1) \|\mathbf{x}\|^2 \end{aligned}$$

Let's replace  $\|\mathbf{x}\|^2$  as well. We also know:

$$\begin{aligned} \lambda_{\min}(\mathbf{P}) \|\mathbf{x}\|^2 &\leq V(\mathbf{x}) \leq \lambda_{\max}(\mathbf{P}) \|\mathbf{x}\|^2 \\ \lambda_{\max}(\mathbf{P}) \|\mathbf{x}\|^2 &\leq -V(\mathbf{x}) \leq \lambda_{\min}(\mathbf{P}) \|\mathbf{x}\|^2 \\ \lambda_{\max}(\mathbf{P}) \|\mathbf{x}\|^2 &\leq -V(\mathbf{x}) \\ \|\mathbf{x}\|^2 &\leq -\frac{1}{\lambda_{\max}(\mathbf{P})} V(\mathbf{x}) \end{aligned}$$

Substituting this back in we have:

$$\dot{V} \leq -\underbrace{\frac{\lambda_{\min}(\mathbf{Q})}{\lambda_{\max}(\mathbf{P})}}_{2\alpha} V(\mathbf{x})$$

$$\dot{V} \leq \alpha V(\mathbf{x})$$

$$\implies V(\mathbf{x}) \leq e^{-2\alpha t} V_0$$

So if  $\frac{\lambda_{\min}(\mathbf{Q})}{\lambda_{\max}(\mathbf{P})} > 0$ , the Lyapunov function decays exponentially.

Plugging this back into the inequality we have:

$$\lambda_{\min}(\mathbf{P}) \|\mathbf{x}\|^2 \leq V(\mathbf{x}) \leq e^{-2\alpha t} V_0$$

$$\lambda_{\min}(\mathbf{P}) \|\mathbf{x}\|^2 \leq e^{-2\alpha t} V_0$$

$$\|\mathbf{x}\|^2 \leq e^{-2\alpha t} \frac{V(\mathbf{x}_0)}{\lambda_{\min}(\mathbf{P})}$$

$$\|\mathbf{x}\| \leq \sqrt{e^{-2\alpha t} \frac{V(\mathbf{x}_0)}{\lambda_{\min}(\mathbf{P})}}$$

$$\|\mathbf{x}\| \leq e^{-\alpha t} \sqrt{\frac{V(\mathbf{x}_0)}{\lambda_{\min}(\mathbf{P})}}$$

The 2-norm is decreasing, bounded by  $e^{-\alpha t}$ . So we can now see that all asymptotically stable LTI systems are globally exponentially stable.

#### 4.11.3 Solving for $\mathbf{P}$ Given $\mathbf{A}$ and $\mathbf{Q}$

We'll see how we different choices of  $\mathbf{Q}$  can lead to interesting types of Lyapunov functions which have important physical interpretations. But that's getting ahead of things. For now, we want to know when given  $\mathbf{A}$  and  $\mathbf{Q}$  if we can find a  $\mathbf{P}$  such that:

$$\mathbf{A}^\top \mathbf{P} + \mathbf{P} \mathbf{A} = -\mathbf{Q}$$

In fact the above equation is known as a linear matrix equation (LME). In particular, these types of LMEs are common in control theory and are called [Sylvester equations](#). In general, there are efficient numerical methods which can solve these types of equations in  $\mathcal{O}(n^3)$  operations.

An analytic solution is given through use of [vectorization](#) and the [Kronecker Product](#).

$$(\mathbf{A} \hat{\oplus} \mathbf{I}_n) \text{vec}(\mathbf{P}) = -\text{vec}(\mathbf{Q})$$

So the solution is unique when  $\mathbf{A} \hat{\oplus} \mathbf{I}_n$  is invertible.

$\mathbf{A} \hat{\oplus} \mathbf{I}_n$  has eigenvalues  $\lambda_i + \lambda_j$ ,  $\lambda_i, \lambda_j \in \Lambda(\mathbf{A})$ , so we have unique solutions when  $\lambda_i + \lambda_j \neq 0$

#### 4.11.4 Lyapunov Function When $\mathbf{A}$ is Hurwitz

Notice when solving the Sylvester equation for  $\mathbf{P}$ , it is sufficient for  $\mathbf{A}$  to be Hurwitz. Start by noticing that the Lyapunov operator is a product rule. The idea behind the substitution is very similar to the use of an integrating factor when solving first order linear ODEs.

$$\mathbf{L}_\mathbf{A}(\mathbf{P}) = \mathbf{A}^\top \mathbf{P} + \mathbf{P} \mathbf{A}$$

$$\text{Choose } \mathbf{P} = \int_0^\infty e^{\mathbf{A}^\top t} \mathbf{Q} e^{\mathbf{A} t} dt$$

Substituting back in to the Lyapunov operator, we have:

$$\begin{aligned}
L_A \left( \int_0^\infty e^{A^\top t} Q e^{At} dt \right) &= A^\top \left( \int_0^\infty e^{A^\top t} Q e^{At} dt \right) + \left( \int_0^\infty e^{A^\top t} Q e^{At} dt \right) A \\
&= \int_0^\infty A^\top e^{A^\top t} Q e^{At} + e^{A^\top t} Q e^{At} A dt \\
&= \int_0^\infty \frac{d}{dt} \left[ e^{A^\top t} Q e^{At} \right] dt \\
&= e^{A^\top t} Q e^{At} \Big|_0^\infty \\
&= -Q
\end{aligned}$$

### Interpretation of the Lyapunov Function for Hurwitz Matrices

Let's plug in a constant vector to the Lyapunov equation and see what happens.

$$\begin{aligned}
V(\mathbf{x}_0) &= \mathbf{x}_0^\top \left( \int_0^\infty e^{A^\top t} Q e^{At} dt \right) \mathbf{x}_0 \\
&= \int_0^\infty \mathbf{x}_0^\top e^{A^\top t} Q e^{At} \mathbf{x}_0 dt \\
&= \int_0^\infty \|\mathbf{x}(t)\|_Q^2 dt
\end{aligned}$$

Thus we can see this as a form of cost to move from state  $\mathbf{x}_0$  to  $\mathbf{x}_{eq}$ . Another view is this characterizes the work done passively by the system in moving from  $\mathbf{x}_0$  to  $\mathbf{x}_{eq}$ . Different choices of  $Q$ , of course, will yield different values for how much "work" or "cost" there is.

## 4.12 Controllability

### 4.12.1 Definitions

**Definition:** Set of States Reachable in in  $t$  time or  $\nu$  epochs

The reachable subspace of a LTI system given by  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$

Continuous time:

$$R_t := \left\{ \int_0^t e^{(t-\tau)\mathbf{A}} \mathbf{B} \mathbf{u}(\tau) d\tau \mid \mathbf{u} : [0, t] \rightarrow \mathbb{R}^m \right\}$$

Discrete time:

$$R_\nu := \left\{ \sum_{k=0}^{\nu-1} \mathbf{A}^{\nu-k-1} \mathbf{B} \mathbf{u}[k] \mid \mathbf{u} : [0 .. t] \rightarrow \mathbb{R}^m \right\}$$

Notice that  $t_2 > t_1 \implies \mathcal{R}_{t_2} \subseteq \mathcal{R}_{t_1}$

Intuitively this makes sense. We can reach more states given more time.

**Definition:** Reachable Subspace

$$R := \bigcup_{t \geq 0} R_t$$

This is the set of all states reachable given some input and unlimited time. We'll see why we know this is a subspace in a bit.

**Definition:** Controllable

A system is controllabe iff  $R = \mathbb{R}^m$ .

That is, any state is reachable given enough time.

### 4.12.2 Discrete Time Controllability

In a pratical sense, it would be impossible to search over all possible inputs over infinite time to see if all states are truly reachable. We'll need to develop a better way of analyzing controllability.

Let's start by recalling in discrete time:

$$\mathbf{x}[1] = \mathbf{A}\mathbf{x}[0] + \mathbf{B}\mathbf{u}[0]$$

$$\begin{aligned} \mathbf{x}[2] &= \mathbf{A}\mathbf{x}[1] + \mathbf{B}\mathbf{u}[1] \\ &= \mathbf{A}(\mathbf{A}\mathbf{x}[0] + \mathbf{B}\mathbf{u}[0]) + \mathbf{B}\mathbf{u}[1] \\ &= \mathbf{A}^2\mathbf{x}[0] + \mathbf{A}\mathbf{B}\mathbf{u}[0] + \mathbf{B}\mathbf{u}[1] \end{aligned}$$

Continuing on inductively we have:

$$\mathbf{x}[\nu] = \mathbf{A}^\nu \mathbf{x}[0] + \sum_{k=0}^{\nu-1} \mathbf{A}^{\nu-k-1} \mathbf{B} \mathbf{u}[k]$$

Define:

$$\mathcal{C}_\nu := \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{\nu-1}\mathbf{B} \end{bmatrix}$$

We can now write:

$$\mathbf{x}[\nu] = \mathbf{A}^\nu \mathbf{x}[0] + \mathcal{C}_\nu \begin{bmatrix} \mathbf{u}[\nu-1] \\ \vdots \\ \mathbf{u}[0] \end{bmatrix}$$

We can now see the set of states reachable in  $\nu$  epochs is  $R_\nu = \text{img}(\mathcal{C}_\nu)$ . How big can  $R_\nu$  get? Recall the Cayley-Hamilton theorem.

$\chi_{\mathbf{A}}(\mathbf{A}) = [0]$ . Thus we can see that a linear combination of  $\mathbf{A}, \dots, \mathbf{A}^n$  are not linearly independent.

Thus we can see:

$$\text{rank}(\mathcal{C}_\nu) = \begin{cases} \nu & \text{if } \nu < n \\ n & \text{if } \nu \geq n \end{cases}$$

**Definition:** Controllability Matrix

$$\mathcal{C} := \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}$$

The system is controllable iff  $\text{rank}(\mathcal{C}) = n$



### 4.12.3 Continuous Time Controllability

It turns out that the controllability matrix is the same in continuous time as in discrete time. Let's derive this.

**Lemma:**

$\mathbf{A}^k$  is linearly dependent with  $\mathbf{A}, \dots, \mathbf{A}^{n-1} \quad \forall k \geq n$

$$\begin{aligned} e^{\mathbf{A}t} &:= \sum_{k=0}^{\infty} \frac{(\mathbf{A}t)^k}{k!} \\ &= \sum_{k=0}^{n-1} \frac{(\mathbf{A}t)^k}{k!} + \sum_{k=n}^{\infty} \frac{(\mathbf{A}t)^k}{k!} \\ &= \sum_{k=0}^{n-1} c_k(t) \mathbf{A}^k \end{aligned}$$

$$\begin{aligned} \mathbf{x}(t) &= e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t e^{\mathbf{A}\tau} \mathbf{B} u(t-\tau) d\tau \\ &= e^{\mathbf{A}t} \mathbf{x}_0 + \int_0^t \left( \sum_{k=0}^n c_k(\tau) \mathbf{A}^k \right) \mathbf{B} u(t-\tau) d\tau \\ &= e^{\mathbf{A}t} \mathbf{x}_0 + \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} \int_0^t c_k(\tau) u(t-\tau) d\tau \\ &= e^{\mathbf{A}t} \mathbf{x}_0 + \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}}_{\mathcal{C}} \begin{bmatrix} \int_0^t c_1(\tau) u(t-\tau) d\tau \\ \vdots \\ \int_0^t c_{n-1}(\tau) u(t-\tau) d\tau \end{bmatrix} \\ &= e^{\mathbf{A}t} \mathbf{x}_0 + \mathcal{C} \begin{bmatrix} \int_0^t c_1(\tau) u(t-\tau) d\tau \\ \vdots \\ \int_0^t c_{n-1}(\tau) u(t-\tau) d\tau \end{bmatrix} \end{aligned}$$

Notice that reachability in continuous time only depends on  $\mathcal{C}$ . Given a large enough input vector, we can reach any state that is in  $\text{img}(\mathcal{C})$ , so once again, the system is controllable iff  $\text{rank}(\mathbf{C}) = n$

There is also another context in which the controllability shows up in continuous time.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

$$\ddot{\mathbf{x}} = \mathbf{A}^2\mathbf{x} + \mathbf{A}\mathbf{B}\mathbf{u} + \mathbf{B}\dot{\mathbf{u}}$$

Continuing inductively:

$$\begin{aligned} \mathbf{x}^{(\nu)} &= \mathbf{A}^\nu \mathbf{x} + \sum_{k=0}^{\nu-1} \mathbf{A}^{\nu-k-1} \mathbf{B} \mathbf{u}^{(k)} \\ &= \mathbf{A}^\nu \mathbf{x} + \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{\nu-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{u}^{(n)} \\ \vdots \\ \mathbf{u} \end{bmatrix} \\ &= \mathbf{A}^\nu \mathbf{x} + \mathcal{C} \begin{bmatrix} \mathbf{u}^{(n)} \\ \vdots \\ \mathbf{u} \end{bmatrix} \end{aligned}$$

Thus we have:

$$\mathcal{C} \begin{bmatrix} \mathbf{u}^{(n)} \\ \vdots \\ \mathbf{u} \end{bmatrix} = \mathbf{x}^{(n)} - \mathbf{A}^n \mathbf{x}$$

If  $\mathcal{C}$  is full rank (that is, the system is controllable), we can make the right hand side (the state derivatives of power  $n$ ) anything we want.

#### 4.12.4 Controllability is Coordinate Independent

Intuitively, the choice of units we use for our state space should not affect the controllability. Indeed it doesn't. We now prove this. Suppose we represent our state space with new basis  $\mathcal{B}$ . Recall from the Section 4.4.1 discussion on state space coordinate systems:

$$\mathbf{A} = \mathbf{P}\mathbf{A}_{\mathcal{B}}\mathbf{P}^{-1}, \mathbf{B} = \mathbf{P}\mathbf{B}_{\mathcal{B}}$$

$$\begin{aligned} \mathcal{C} &= \begin{bmatrix} \mathbf{P}\mathbf{B}_{\mathcal{B}} & (\mathbf{P}\mathbf{A}_{\mathcal{B}}\cancel{\mathbf{P}^{-1}})(\cancel{\mathbf{P}}\mathbf{B}_{\mathcal{B}}) & \dots & (\mathbf{P}\mathbf{A}_{\mathcal{B}}^n\cancel{\mathbf{P}^{-1}})(\cancel{\mathbf{P}}\mathbf{B}_{\mathcal{B}}) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{P}\mathbf{B}_{\mathcal{B}} & \mathbf{P}\mathbf{A}_{\mathcal{B}}\mathbf{B}_{\mathcal{B}} & \dots & \mathbf{P}\mathbf{A}_{\mathcal{B}}^n\mathbf{B}_{\mathcal{B}} \end{bmatrix} \\ &= \mathbf{P} \underbrace{\begin{bmatrix} \mathbf{B}_{\mathcal{B}} & \mathbf{A}_{\mathcal{B}}\mathbf{B}_{\mathcal{B}} & \dots & \mathbf{A}_{\mathcal{B}}^n\mathbf{B}_{\mathcal{B}} \end{bmatrix}}_{\mathcal{C}_{\mathcal{B}}} \\ &= \mathbf{P}\mathcal{C}_{\mathcal{B}} \end{aligned}$$

But since  $\mathbf{P}$  is a change of basis matrix, it must be invertible.

$$\implies \text{rank}(\mathcal{C}) = \text{rank}(\mathbf{P}\mathcal{C}_{\mathcal{B}}) = \text{rank}(\mathcal{C}_{\mathcal{B}})$$

#### 4.12.5 The Controllable and Uncontrollable Subspace

An invariant subspace is a subspace  $(V, \mathbb{F}) : \mathbf{A}V = V$

Eigenspaces (span of eigenvectors) are clearly invariant subspaces and so is  $\{\mathbf{0}\}$  and  $\mathbb{R}^n$ . What other subspaces might be invariant? We'll consider an block upper triangular matrix.

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} \text{ where } \mathbf{A}_{11} \in \mathbb{R}^{r \times r}$$

Clearly  $V = \left\{ \begin{bmatrix} \mathbf{v}_r \\ 0 \end{bmatrix} \mid \mathbf{z} \in \mathbb{R}^r \right\}$  is an invariant subspace.

**Theorem:** The Controllable Subspace is  $\mathbf{A}$ -invariant

$$\begin{aligned} \mathbf{A} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \mathbf{v}_k &= \mathbf{A} \sum_{k=0}^{n-1} \mathbf{A}^{n-k-1} \mathbf{B} \mathbf{v}_k \\ &= \sum_{k=0}^{n-1} \mathbf{A}^{n-k} \mathbf{B} \mathbf{v}_k \\ &= \mathbf{A}^n \mathbf{B} \mathbf{v}_0 + \sum_{k=1}^{n-1} \mathbf{A}^{n-k} \mathbf{B} \mathbf{v}_k \\ &= \sum_{k=1}^{n-1} c_k \mathbf{A}^{n-k} \mathbf{B} \mathbf{v}_0 + \sum_{k=1}^{n-1} \mathbf{A}^{n-k} \mathbf{B} \mathbf{v}_k \\ &= \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{v}_1 + c_1 \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{n-1} + c_{n-1} \mathbf{v}_0 \end{bmatrix} \end{aligned}$$

$$\implies \mathbf{A} \text{img}(\mathcal{C}) = \text{img}(\mathcal{C})$$

—

We can characterize an invariant subspace with a matrix.

Let  $(V, \mathbb{F}) \subset \mathbb{R}^{n \times k}$  be an invariant subspace spanned by  $\mathbf{v}_1, \dots, \mathbf{v}_k$

Define:

$$\mathbf{M} = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_k \end{bmatrix}$$

Notice that:

$$\mathbf{A}\mathbf{v}_1 = \sum_{k=1}^k x_{1,k} \mathbf{v}_k$$

$$\mathbf{A}\mathbf{v}_2 = \sum_{k=1}^k x_{2,k} \mathbf{v}_k$$

$x_{i,k}$  forms the entries of a matrix  $\mathbf{X}$

$$\mathbf{A}\mathbf{M} = \mathbf{M} \begin{bmatrix} x_{1,1} & \cdots & x_{1,k} \\ \vdots & & \vdots \\ x_{k,1} & \cdots & x_{k,k} \end{bmatrix}$$

So we have:

$$\mathbf{A}\mathbf{M} = \mathbf{M}\mathbf{X}$$

So  $\text{img}(\mathbf{M})$  is  $\mathbf{A}$ -invariant  $\iff \exists \mathbf{X} : \mathbf{A}\mathbf{M} = \mathbf{M}\mathbf{X}$

Which is a Sylvester equation. We know this has solutions when  $\mathbf{A}$  and  $\mathbf{M}$  don't share eigenvalues.

Now consider  $\tilde{\mathbf{M}} \in \mathbb{R}^{n \times (n-k)}$ .

$$\begin{aligned} \mathbf{A} \begin{bmatrix} \mathbf{M} & \tilde{\mathbf{M}} \end{bmatrix} &= \begin{bmatrix} \mathbf{A}\mathbf{M} & \mathbf{A}\tilde{\mathbf{M}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{M} & \tilde{\mathbf{M}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix} \end{aligned}$$

$$\text{where } \begin{bmatrix} \mathbf{A}_{12} \\ \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \tilde{\mathbf{M}} \end{bmatrix}^{-1} \mathbf{A}\tilde{\mathbf{M}}$$

So if we have  $\mathbf{P} = \begin{bmatrix} \mathbf{M} & \tilde{\mathbf{M}} \end{bmatrix}$

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix}$$

So as long as we select  $\tilde{\mathbf{M}}$  so that  $\mathbf{P}$  is invertible we can reveal the controllable subspace.

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{bmatrix}, \mathbf{P}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{B}_c \\ 0 \end{bmatrix}$$

$$\mathbf{P}^{-1}\mathbf{C} = \begin{bmatrix} \mathbf{B}_c & \mathbf{B}_c\mathbf{A}_c & \cdots & \mathbf{A}_c^{n-1}\mathbf{B}_c \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

#### 4.12.6 Impulse Inputs

Recall in continuous time, we can reach any state given a large enough input. The "largest" input vector we could provide are impulse inputs and inputs that are distributional derivatives of the impulse (in general you don't need to know what this means but that it behaves like a derivative to the Laplace transform).

**Lemma:**

$$\text{Recall that } f(x) = \frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

$$(s\mathbf{I}_n - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} s^{-k} \mathbf{A}^k$$

—

Now for a single input system choose:

$$\text{Consider } u(t) = \mathbf{v}\delta^{(k)}(t)$$

$$U(s) = \mathbf{v}s^\nu$$

$$\begin{aligned} \mathbf{X}(s) &= (s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B}U(s) \\ &= (s\mathbf{I}_n - \mathbf{A})^{-1} \mathbf{B}\mathbf{v}s^\nu \\ &= \left( \sum_{k=0}^{\infty} s^{-k} \mathbf{A}^k \right) \mathbf{B}\mathbf{v}s^\nu \\ &= \left( \sum_{k=0}^{\infty} s^{\nu-k} \mathbf{A}^k \right) \mathbf{B}\mathbf{v} \\ &= \left( \underbrace{\sum_{k=0}^{\nu-1} s^{\nu-k} \mathbf{A}^k}_{\text{Impulse terms}} + \sum_{k=\nu}^{\infty} s^{\nu-k} \mathbf{A}^k \right) \mathbf{B}\mathbf{v} \end{aligned}$$

$$x(t) = \text{Impulse terms} + \sum_{k=\nu}^{\infty} \frac{t^k}{k!} \mathbf{A}^k \mathbf{B}\mathbf{v}$$

So we have:

$$x(0^+) = \mathbf{A}^\nu \mathbf{B}\mathbf{v}$$

$$\text{Now choose } u(t) = \sum_{k=0}^{n-1} \mathbf{v}_k \delta^{(k)}(t)$$

By linearity we'll have:

$$x(0^+) = \sum_{k=0}^{n-1} c_k \mathbf{A}^k \mathbf{B} = \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}}_{\mathcal{C}} \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{n-1} \end{bmatrix} = \mathcal{C} \begin{bmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{n-1} \end{bmatrix}$$

Thus  $\mathbf{x}(0^+)$  can be made to be any element of  $\text{img}(\mathcal{C})$ . Intuitively, given infinite inputs, we can steer the system state to be anything in the reachable subspace instantaneously.

### 4.13 Least Norm Input

Often we want to get to a reachable state while minimizing the work spent to get there. To do this we minimize:

$$\int_0^t \|\mathbf{u}(\tau)\|^2 d\tau$$

#### Discrete Time

Let's start by discretizing  $\mathbf{u}$ . Let it be piecewise constant on  $\mathbf{u}$  with duration  $h = \frac{t_f}{N}$ .

$$u(\tau) = u_d[k], \quad kh \leq \tau < (k+1)h \text{ where } k \in [0..N-1]$$

$$\text{Let } \mathbf{A}_d = e^{\mathbf{A}h}, \mathbf{B}_d = \left( \int_0^h e^{\mathbf{A}\tau} d\tau \right) \mathbf{B}$$

$$\mathbf{x}_{\text{dst}} = \begin{bmatrix} \mathbf{B}_d & \mathbf{A}_d \mathbf{B}_d & \dots & \mathbf{A}_d^{N-1} \mathbf{B}_d \end{bmatrix} \begin{bmatrix} \mathbf{u}_d[N-1] \\ \vdots \\ \mathbf{u}_d[0] \end{bmatrix}$$

$$\begin{aligned} \begin{bmatrix} \mathbf{u}_d[N-1] \\ \vdots \\ \mathbf{u}_d[0] \end{bmatrix}_{\text{ln}} &= \begin{bmatrix} \mathbf{B}_d & \mathbf{A}_d \mathbf{B}_d & \dots & \mathbf{A}_d^{N-1} \mathbf{B}_d \end{bmatrix}^+ \mathbf{x}(t_f) \\ &= \begin{bmatrix} \mathbf{B}_d^\top \\ (\mathbf{A}_d \mathbf{B}_d)^\top \\ \dots \\ (\mathbf{A}_d^{N-1} \mathbf{B}_d)^\top \end{bmatrix} \left( \begin{bmatrix} \mathbf{B}_d & \mathbf{A}_d \mathbf{B}_d & \dots & \mathbf{A}_d^{N-1} \mathbf{B}_d \end{bmatrix} \begin{bmatrix} \mathbf{B}_d^\top \\ (\mathbf{A}_d \mathbf{B}_d)^\top \\ \dots \\ (\mathbf{A}_d^{N-1} \mathbf{B}_d)^\top \end{bmatrix} \right)^{-1} \mathbf{x}(t_f) \\ &= \begin{bmatrix} \mathbf{B}_d^\top \\ (\mathbf{A}_d \mathbf{B}_d)^\top \\ \dots \\ (\mathbf{A}_d^{N-1} \mathbf{B}_d)^\top \end{bmatrix} \sum_{k=0}^{N-1} \left( \mathbf{A}_d^k \mathbf{B}_d \right) \left( \mathbf{A}_d^k \mathbf{B}_d \right)^\top \end{aligned}$$

So we see that:

$$\mathbf{u}[k]_{\text{dlm}} = \left( \mathbf{A}_d^{N-1-k} \mathbf{B}_d \right)^\top \sum_{k=0}^{N-1} \left( \mathbf{A}_d^k \mathbf{B}_d \right) \left( \mathbf{A}_d^k \mathbf{B}_d \right)^\top$$

## Continuous Time

Now approximating  $\mathbf{B}_d$

$$\begin{aligned}
\mathbf{B}_d &= \left( \int_0^h e^{\mathbf{A}\tau} d\tau \right) \mathbf{B} \\
&= \left( \int_0^h \sum_{k=0}^{\infty} \frac{(\mathbf{A}\tau)^k}{k!} d\tau \right) \mathbf{B} \\
&= \left( \int_0^h \sum_{k=0}^{\infty} \frac{(\mathbf{A}\tau)^k}{k!} d\tau \right) \mathbf{B} \\
&= \left( \int_0^h \mathbf{I} + \sum_{k=1}^{\infty} \frac{(\mathbf{A}\tau)^k}{k!} d\tau \right) \mathbf{B} \\
&= \left( \int_0^h \mathbf{I} d\tau + \int_0^h \sum_{k=1}^{\infty} \frac{(\mathbf{A}\tau)^k}{k!} d\tau \right) \mathbf{B} \\
&= h\mathbf{B} + \mathcal{O}(h^2)
\end{aligned}$$

Now we can rewrite:

$$\mathbf{A}^{N-k-1} \mathbf{B} = e^{\mathbf{A}(t_f-t)} \mathbf{B}_d = h(e^{\mathbf{A}(t_f-t)} \mathbf{B}) + \mathcal{O}(h^2)$$

$$\begin{aligned}
\mathbf{u}_{\text{ln}}(t) &= \lim_{N \rightarrow \infty} \left[ h \left( e^{\mathbf{A}(t_f-t)} \mathbf{B} \right)^{\top} \left( \sum_{k=0}^{N-1} \left( e^{\mathbf{A}hk} \mathbf{B} \right) \left( e^{\mathbf{A}hk} \mathbf{B} \right)^{\top} h^2 \right)^{-1} + \mathcal{O}(h) \right] \mathbf{x}_{\text{dst}} \\
&= \lim_{N \rightarrow \infty} \left[ h \left( e^{\mathbf{A}(t_f-t)} \mathbf{B} \right)^{\top} \frac{1}{h} \left( \sum_{k=0}^{N-1} \left( e^{\mathbf{A}hk} \mathbf{B} \right) \left( e^{\mathbf{A}hk} \mathbf{B} \right)^{\top} h \right)^{-1} + \mathcal{O}(h) \right] \mathbf{x}_{\text{dst}} \\
&= \lim_{N \rightarrow \infty} \left[ \left( e^{\mathbf{A}(t_f-t)} \mathbf{B} \right)^{\top} \left( \sum_{k=0}^{N-1} \left( e^{\mathbf{A}hk} \mathbf{B} \right) \left( e^{\mathbf{A}hk} \mathbf{B} \right)^{\top} h \right)^{-1} + \mathcal{O}(h) \right] \mathbf{x}_{\text{dst}} \\
&= \left( e^{\mathbf{A}(t_f-t)} \mathbf{B} \right)^{\top} \left( \int_0^{t_f} (e^{\mathbf{A}\tau} \mathbf{B}) (e^{\mathbf{A}\tau} \mathbf{B})^{\top} d\tau \right)^{-1} \mathbf{x}_{\text{dst}} \\
&= \left( e^{\mathbf{A}(t_f-t)} \mathbf{B} \right)^{\top} \left( \int_0^{t_f} e^{\mathbf{A}\tau} \mathbf{B} \mathbf{B}^{\top} e^{\mathbf{A}^{\top} \tau} d\tau \right)^{-1} \mathbf{x}_{\text{dst}}
\end{aligned}$$

Notice if we want to take less time,  $t_f$ , the inverse operation will turn that into division and make  $u$  larger.

#### 4.13.1 The Controllability Grammian

**Lemma:**

$$\text{Let } \mathbf{W}_c(t_f) = \int_0^{t_f} e^{\mathbf{A}\tau} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top \tau} d\tau$$

$$\mathbf{W}_c^\top(t_f) = \mathbf{W}_c(t_f)$$

—

#### Minimum Energy, Finite Time Horizon

$$\int_0^{t_f} \|\mathbf{u}(\tau)\|^2 d\tau = \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}}$$

**Proof:**

$$\begin{aligned} \int_0^{t_f} \|\mathbf{u}(\tau)\|^2 d\tau &= \int_0^{t_f} \mathbf{u}^\top(\tau) \mathbf{u}(\tau) d\tau \\ &= \int_0^{t_f} \left( \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c^\top(t_f))^{-1} (e^{\mathbf{A}(t_f-\tau)} \mathbf{B}) \right) \left( (e^{\mathbf{A}(t_f-\tau)} \mathbf{B})^\top (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} \right) d\tau \\ &= \int_0^{t_f} \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} e^{\mathbf{A}(t_f-\tau)} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top(t_f-\tau)} (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} d\tau \\ &= \int_0^{t_f} \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} e^{\mathbf{A}(t_f-\tau)} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top(t_f-\tau)} (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} d\tau \\ &= \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \left( \int_0^{t_f} e^{\mathbf{A}(t_f-\tau)} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top(t_f-\tau)} d\tau \right) (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} \end{aligned}$$

Let  $u = t_f - \tau$ ,  $du = -d\tau$

$$\begin{aligned} \int_0^{t_f} \|\mathbf{u}(\tau)\|^2 d\tau &= -\mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \left( \int_{t_f}^0 e^{\mathbf{A}u} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top u} du \right) (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} \\ &= \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \left( \int_0^{t_f} e^{\mathbf{A}u} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top u} du \right) (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} \\ &= \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \cancel{\mathbf{W}_c(t_f)} (\cancel{\mathbf{W}_c(t_f)})^{-1} \mathbf{x}_{\text{dst}} \\ &= \mathbf{x}_{\text{dst}}^\top (\mathbf{W}_c(t_f))^{-1} \mathbf{x}_{\text{dst}} \end{aligned}$$



## Minimum Energy, Infinite Time Horizon

Likewise we can take quantify minimum energy given infinite time by examining the last theorem in the limit.

$$\int_0^\infty \|\mathbf{u}_{\text{in}}(\tau)\|^2 d\tau = \mathbf{x}_{\text{dst}}^\top \underbrace{\left[ \lim_{t_f \rightarrow \infty} \left( \int_0^{t_f} e^{\mathbf{A}\tau} \mathbf{B} \mathbf{B}^\top e^{\mathbf{A}^\top \tau} d\tau \right)^{-1} \right]}_{\mathbf{W}_c} \mathbf{x}_{\text{dst}} = \mathbf{x}_{\text{dst}}^\top \mathbf{W}_c \mathbf{x}_{\text{dst}}$$

$\mathbf{P}$  is uniquely determined this way so long as  $\mathbf{A}$  is Hurwitz. In which case,  $\mathbf{P}$  is positive definite. There is always some nonzero cost to getting to a destination.

If  $\mathbf{A}$  is not Hurwitz,  $\mathbf{W}_c$  may still exist, but  $\mathbf{W}_c$  may have a nontrivial nullspace (in which case it is positive semidefinite instead of positive definite). In that case, there are certain destinations we can get to for arbitrarily small energy. Once you give these systems a little kick, their instability will carry you to the destination.

The quantity  $\mathbf{W}_c$  is called the controllability [Grammian](#). Grammians (or Gram matrices) are quantities given by  $\mathbf{A}^H \mathbf{A}$  where  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . They are hermitian positive semidefinite and capture many intrinsic geometric properties of a space.

The integration in the controllability Grammian may seem familiar if you recall the last section on Lyapunov functions. Indeed we can see that  $\mathbf{W}_c(t_f)$  is given by the solution to the Sylvester equation:

$$\mathbf{P}\mathbf{A} + \mathbf{A}^\top \mathbf{P} = -\mathbf{B}\mathbf{B}^\top \text{ where } \mathbf{A} \text{ is Hurwitz.}$$

**Theorem:**  $\text{rank}(\mathbf{W}_c) = \text{rank}(\mathcal{C})$

Consider  $e^{\mathbf{A}t} \mathbf{B}$ . Recall from that due to Caley-Hamilton we can represent:

$$\begin{aligned} e^{\mathbf{A}t} &= \sum_{k=0}^{n-1} c_k(t) \mathbf{A}^k \\ e^{\mathbf{A}t} \mathbf{B} &= \sum_{k=0}^{n-1} c_k(t) \mathbf{A}^k \mathbf{B} \\ &= \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} c_1(t) \end{bmatrix} \end{aligned}$$

#### 4.13.2 PCH Test For Controllability

**Theorem:**

$\mathcal{C}$  is full row rank  $\iff \begin{bmatrix} \mathbf{A} - \lambda \mathbf{I}_n & \mathbf{B} \end{bmatrix}$  is full row rank where  $\lambda \in \Lambda(\mathbf{A})$

**Proof  $\implies$ :**

Suppose  $\mathcal{C}$  is full row rank but  $\begin{bmatrix} \mathbf{A} - \lambda \mathbf{I}_n & \mathbf{B} \end{bmatrix}$  is not full row rank.

$$\exists \lambda \in \Lambda(\mathbf{A}) : \mathbf{v} \begin{bmatrix} \mathbf{A} - \lambda \mathbf{I}_n & \mathbf{B} \end{bmatrix} = \mathbf{0}$$

$$\implies \mathbf{v}\mathbf{A} = \lambda \mathbf{v} \text{ and } \mathbf{v}\mathbf{B} = \mathbf{0}$$

$$\begin{aligned} \mathbf{v} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} &= \begin{bmatrix} \mathbf{v}\mathbf{B} & \mathbf{v}\mathbf{A}\mathbf{B} & \dots & \mathbf{v}\mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{v}\mathbf{B} & \mathbf{v}\mathbf{A}\mathbf{B} & \dots & \mathbf{v}\mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} & \lambda \mathbf{v}\mathbf{B} & \dots & \lambda^{n-1} \mathbf{v}\mathbf{B} \end{bmatrix} \\ &= \mathbf{0} \end{aligned}$$

So  $\exists \mathbf{v} \neq \mathbf{0} : \mathbf{v}\mathcal{C} = \mathbf{0}$ , which contradicts our assumption that  $\mathcal{C}$  is full row rank.

**Proof  $\Leftarrow$ :**

We'll prove the contrapositive.  $\mathcal{C}$  is not full row rank  $\implies \begin{bmatrix} \mathbf{A} - \lambda \mathbf{I}_n & \mathbf{B} \end{bmatrix}$  is not full row rank

When  $\mathcal{C}$  is not full row rank, under a specific set of coordinates for the state space,  $\mathbf{A}$  and  $\mathbf{B}$  into

$$\tilde{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{A}_c & \tilde{\mathbf{A}}_{12} \\ 0 & \tilde{\mathbf{A}}_{22} \end{bmatrix}, \tilde{\mathbf{B}} = \mathbf{P}\mathbf{B} = \begin{bmatrix} \tilde{\mathbf{B}}_c \\ 0 \end{bmatrix}$$

Recall that controllability is invariant under state space transformations, thus we can check apply the PCH test on the transformed state space.

$$\begin{bmatrix} \tilde{\mathbf{A}} - \lambda \mathbf{I} & \tilde{\mathbf{B}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}}_c - \lambda \mathbf{I} & \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{B}}_c \\ 0 & \tilde{\mathbf{A}}_c - \lambda \mathbf{I} & 0 \end{bmatrix}$$

Notice if we select  $\mathbf{v} : \mathbf{v}\mathbf{A} = \lambda \mathbf{v}$

$$\mathbf{v} \begin{bmatrix} \tilde{\mathbf{A}} - \lambda \mathbf{I} & \tilde{\mathbf{B}} \end{bmatrix} = \mathbf{v} \begin{bmatrix} \tilde{\mathbf{A}}_c - \lambda \mathbf{I} & \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{B}}_c \\ 0 & \tilde{\mathbf{A}}_c - \lambda \mathbf{I} & 0 \end{bmatrix} = [0]$$

So  $\begin{bmatrix} \tilde{\mathbf{A}} - \lambda \mathbf{I} & \tilde{\mathbf{B}} \end{bmatrix}$  is not full row rank.

## 5 Digital Signal Processing

### 5.1 Discrete Time Fourier Transform

$$X(e^{j\omega}) = \mathcal{F}\{x[n]\} := \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \quad x[n] = \mathcal{F}^{-1}\{X(e^{j\omega})\} := \frac{1}{2\pi} \int_{\theta_0}^{\theta_0+2\pi} X(e^{j\omega})e^{j\omega n}d\omega$$

#### 5.1.1 DTFT Properties

Property	Time Domain	Frequency Domain
Linearity	$c_1x_1[n] + c_2x_2[n]$	$c_1X_1 + c_2X_2(e^{j\omega})$
Time Shift	$x[n - k]$	$X(\omega)e^{-j\omega k}$
Time Reversal	$x[-n]$	$X(e^{-j\omega})$
Time Scale	$x[an]$	$\frac{1}{ a }X\left(e^{\frac{j\omega}{a}}\right)$
Time Shift + Scale	$x[an - k]$	$\frac{1}{ a }X\left(e^{\frac{j\omega}{a}}\right)e^{-\frac{j\omega k}{a}}$
Complex Conjugation	$x^*[n]$	$X^*(e^{-j\omega})$
Frequency Shift	$x[n]e^{-j\omega_0 n}$	$X(e^{j(\omega-\omega_0)})$
Convolution	$x_1[n] * x_2[n]$	$X_1(e^{j\omega})X_2(e^{j\omega})$
Modulation	$x_1[n]x_2[n]$	$\frac{1}{2\pi}X_1(e^{j\omega}) \circledast X_2(e^{j\omega})$
First Difference	$\Delta x[n - 1]$	$(1 - e^{-j\omega})X(e^{j\omega})$
Accumulation	$\mathcal{A}_n\{x[n]\}$	$\frac{1}{1 - e^{-j\omega}}X(e^{j\omega})$
Frequency Differentiation	$(-jt)^k x[n]$	$\mathcal{D}_\omega^k\{F(e^{j\omega})\}$
Parseval's Theorem	$\sum_{n=-\infty}^{\infty} x_1[n]x_2^*[n]$	$\frac{1}{2\pi} \int_{-\infty}^{\infty} X_1(e^{j\omega})X_2^*(e^{j\omega}) d\omega$
Duality	$X(e^{2\pi n})$	$2\pi x(-\omega)$

### 5.1.2 DTFT Symmetries

$x[n]$	$X(z)$
$x^*[n]$	$X^*(e^{-j\omega})$
$x^*[-n]$	$X^*(e^{j\omega})$
$\Re\{x[n]\}$	$X_e(e^{j\omega})$
$j\Im\{x[n]\}$	$X_o(e^{j\omega})$
$x_e[n]$	$\Re\{X(e^{j\omega})\}$
$x_o[n]$	$j\Im\{X(e^{j\omega})\}$
Any real $x[n]$	$X(e^{j\omega}) = X^*(e^{-j\omega})$
Any real $x[n]$	$\Re\{X(e^{-j\omega})\} = \Re\{X(e^{j\omega})\}$
Any real $x[n]$	$\Im\{X(e^{-j\omega})\} = -\Im\{X(e^{j\omega})\}$
Any real $x[n]$	$ X(e^{-j\omega})  =  X(e^{j\omega}) $
Any real $x[n]$	$\arg(X(e^{-j\omega})) = -\arg(X(e^{j\omega}))$

### 5.1.3 DTFT Table

$x[n]$	$X(e^{j\omega})$
0	constant $\omega_0$
1	$2\pi W_{2\pi}(\omega)$
$\delta[n]$	1
$u_H[n]$	$\frac{1}{1 - e^{-j\omega}} + \pi W_{2\pi}(\omega)$
$a^n u[n],  a  < 1$	$\frac{1}{1 - ae^{-j\omega}}$
$(n+1)a^n u[n],  a  < 1$	$\frac{1}{(1 - ae^{-j\omega})^2}$
$\frac{(n+r-1)!}{n!(r-1)!} a^n u[n],  a  < 1$	$\frac{1}{(1 - ae^{-j\omega})^r}$
$\text{rect}\left[\frac{t}{N}\right], N \in \mathbb{Z} \setminus \{0\}$	$\text{diric}_{\frac{N}{2}}(\omega)$
$\omega_c \text{sinc}[\omega_c n], \omega_c \leq \pi$	$\text{rect}\left(\frac{1}{\omega_c} \omega\right)$
$\omega_c \text{sinc}^2[\omega_c n], \omega_c \leq \pi$	$\text{tri}\left(\frac{1}{\omega_c} \omega\right)$
$e^{j\omega_0 n}$	$2\pi W_{2\pi}(\omega - \omega_0)$
$\cos[\omega_0 n + \phi]$	$\pi e^{j\phi} W_{2\pi}(\omega - \omega_0) + \pi e^{-j\phi} W_{2\pi}(\omega + \omega_0)$
$\sin[\omega_0 n + \phi]$	$\frac{\pi}{j} e^{j\phi} W_{2\pi}(\omega - \omega_0) - \frac{\pi}{j} e^{-j\phi} W_{2\pi}(\omega + \omega_0)$

## 5.2 Z Transform

$$X(z) = \mathcal{Z}\{x[n]\} := \sum_{n=-\infty}^{\infty} x[n]z^{-n}$$

$$x[n] = \mathcal{Z}^{-1}\{X(z)\} := \frac{1}{2\pi j} \oint_C X(z)z^{n-1} dz$$

### 5.2.1 Z Transform Properties

Property	$x[n]$	$X(z)$
Linearity	$c_1x_1[n] + c_2x_2[n]$	$c_1X_1(z) + c_2X_2(z)$
Time Shifting	$x[n - k]$	$X(z)z^{-k}$
Time Reversal	$x[-n]$	$X(z^{-1})$
Z-Scaling	$z_0^n x[n]$	$X(z_0^{-1}z)$
Backward First Difference	$\Delta x[n - 1]$	$(1 - z^{-1})X(z)$
Forward First Difference	$\Delta x[n]$	$(z - 1)X(z) - zx[0]$
Complex Conjugation	$x^*[n]$	$X^*(z^*)$
Convolution	$x_1[n] * x_2[n]$	$X_1(z)X_2(z)$
Z Differentiation	$n^k x[n]$	$(-z)^k \mathcal{D}_z^k \{X(z)\}$
Parseval's Theorem	$\sum_{n=-\infty}^{\infty} x_1[n]x_2^*[n]$	$\frac{1}{2\pi j} \oint_C X_1(z)X_2^*((z^*)^{-1})z^{-1} dz$

#### Initial Value Theorem:

$$x[n] \text{ is causal} \implies x[0] = \lim_{z \rightarrow \infty} X(z)$$

#### Final Value Theorem:

$$x[n] \text{ is causal} \implies \lim_{n \rightarrow \infty} x[n] = \lim_{z \rightarrow 1} (z - 1)X(z)$$

### 5.2.2 Z Transform Table

$x[n]$	$X(z)$	ROC
$\delta[n - k]$	$z^{-k}$	$\mathbb{C}$
$u_H[n]$	$\frac{1}{1 - z^{-1}}$	$ z  > 1$
$-u_H[-n - 1]$	$\frac{1}{1 - z^{-1}}$	$ z  < 1$
$nu_H[n]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z  > 1$
$n^2u_H[n]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z  > 1$
$-nu_H[-n - 1]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z  < 1$
$-n^2u_H[-n - 1]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z  < 1$
$a^n u_H[n]$	$\frac{1}{1 - az^{-1}}$	$ z  >  a $
$na^n u_H[n]$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z  >  a $
$n^2a^n u_H[n]$	$\frac{az^{-1}(1 + az^{-1})}{(1 - az^{-1})^3}$	$ z  >  a $
$-na^n u_H[-n - 1]$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z  <  a $
$-n^2a^n u_H[-n - 1]$	$\frac{az^{-1}(1 + az^{-1})}{(1 - az^{-1})^3}$	$ z  <  a $
$\binom{n+k-1}{k-1} a^n u[n]$	$\frac{1}{(1 - az^{-1})^k}$	$ z  >  a $
$(-1)^k \binom{-n-1}{k-1} a^n u[-n - k]$	$\frac{1}{(1 - az^{-1})^k}$	$ z  <  a $
$\cos(\omega_0 n) u_H[n]$	$\frac{1 - z^{-1} \cos(\omega_0)}{1 - 2z^{-1} \cos(\omega_0) + z^{-2}}$	$ z  > 1$
$\sin(\omega_0 n) u_H[n]$	$\frac{z^{-1} \sin(\omega_0)}{1 - 2z^{-1} \cos(\omega_0) + z^{-2}}$	$ z  > 1$
$a^n \cos(\omega_0 n) u_H[n]$	$\frac{1 - az^{-1} \cos(\omega_0)}{1 - 2az^{-1} \cos(\omega_0) + a^2 z^{-2}}$	$ z  >  a $
$a^n \sin(\omega_0 n) u_H[n]$	$\frac{az^{-1} \sin(\omega_0)}{1 - 2az^{-1} \cos(\omega_0) + a^2 z^{-2}}$	$ z  >  a $

## 5.3 Sampling, Compression, and Expansion

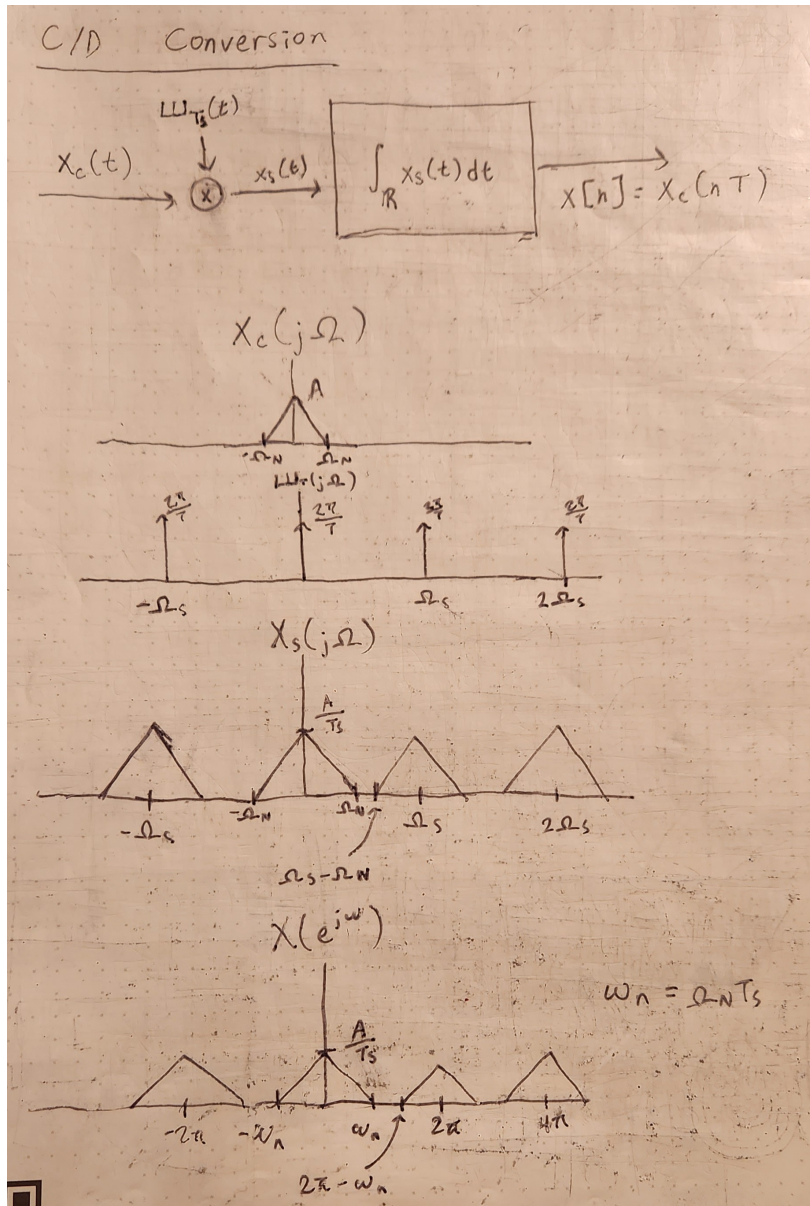
### 5.3.1 C/D Conversion

**Shanon-Nyquist Sampling Theorem:**

Suppose  $X$  is a bandlimited signal with bandwidth  $2\Omega_N$

$\forall \Omega : |\Omega| > \Omega_N, X(j\Omega) = 0$ .

Perfect reconstruction of the signal requires:  $\Omega_s = \frac{2\pi}{T_s} > 2\Omega_N$ .



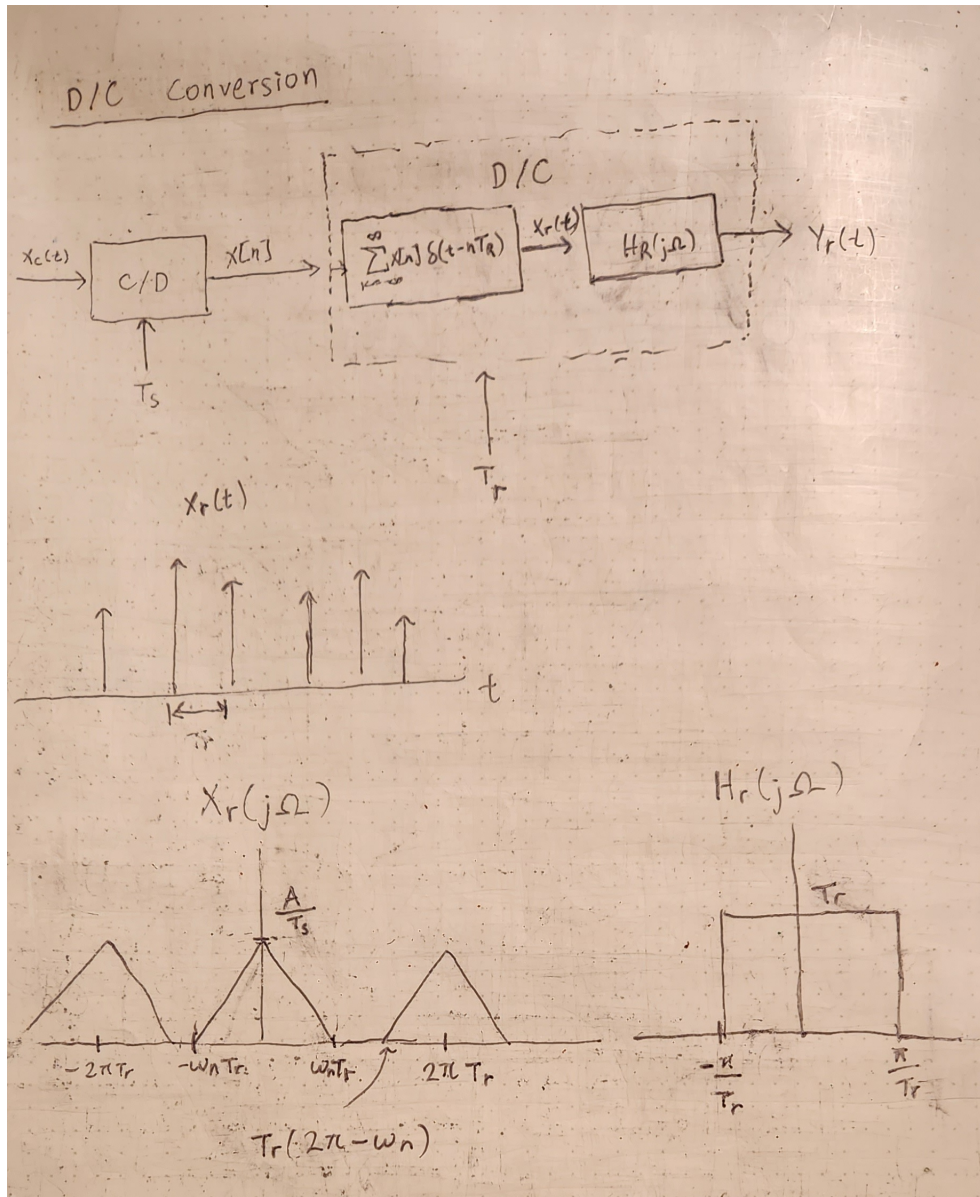


### 5.3.2 D/C Conversion

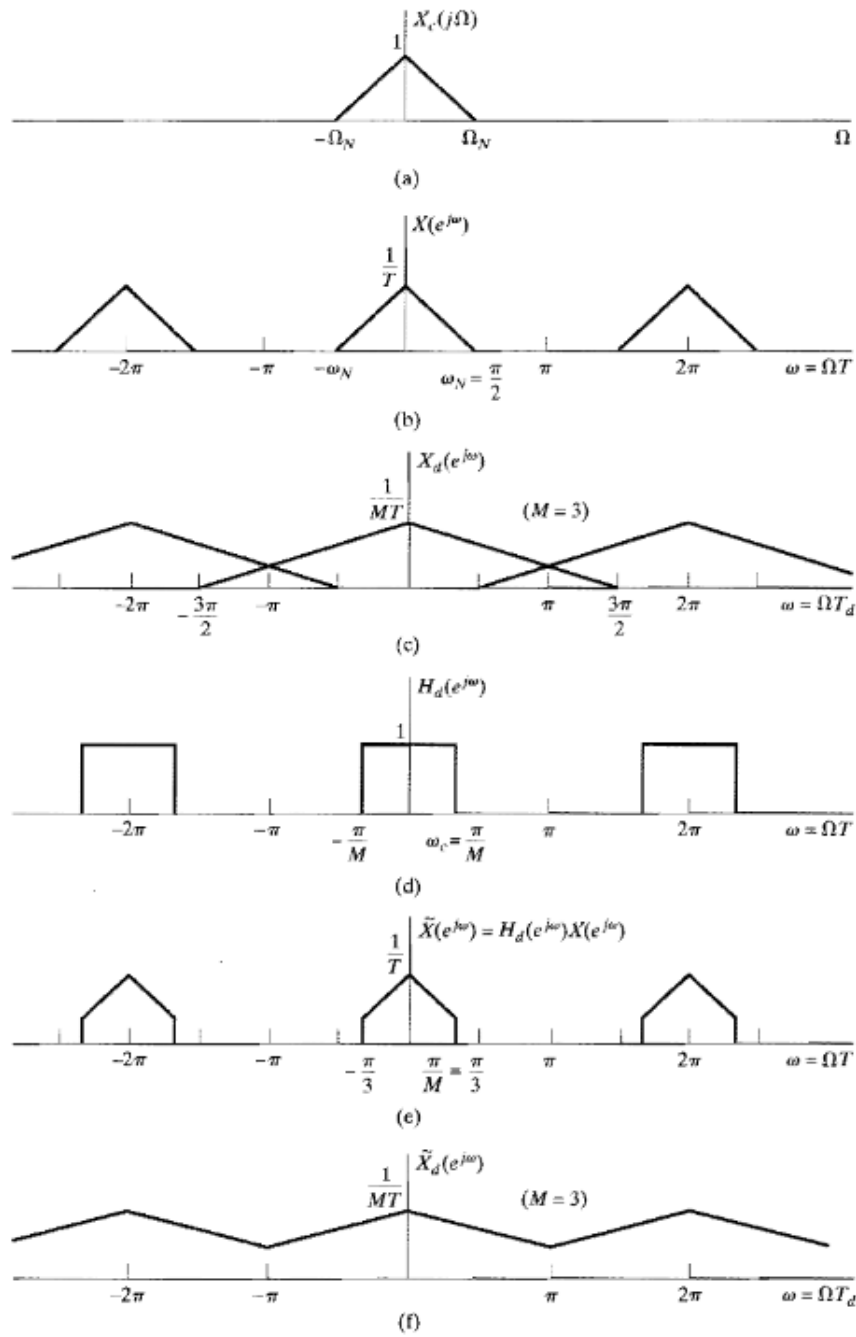
Mismatch in Reconstruction and Sampling Period:

Assuming no aliasing effects, if  $T_r \neq T_s$

$$y_r(t) = \frac{T_r}{T_s} x_c\left(\frac{T_s}{T_r}t\right)$$

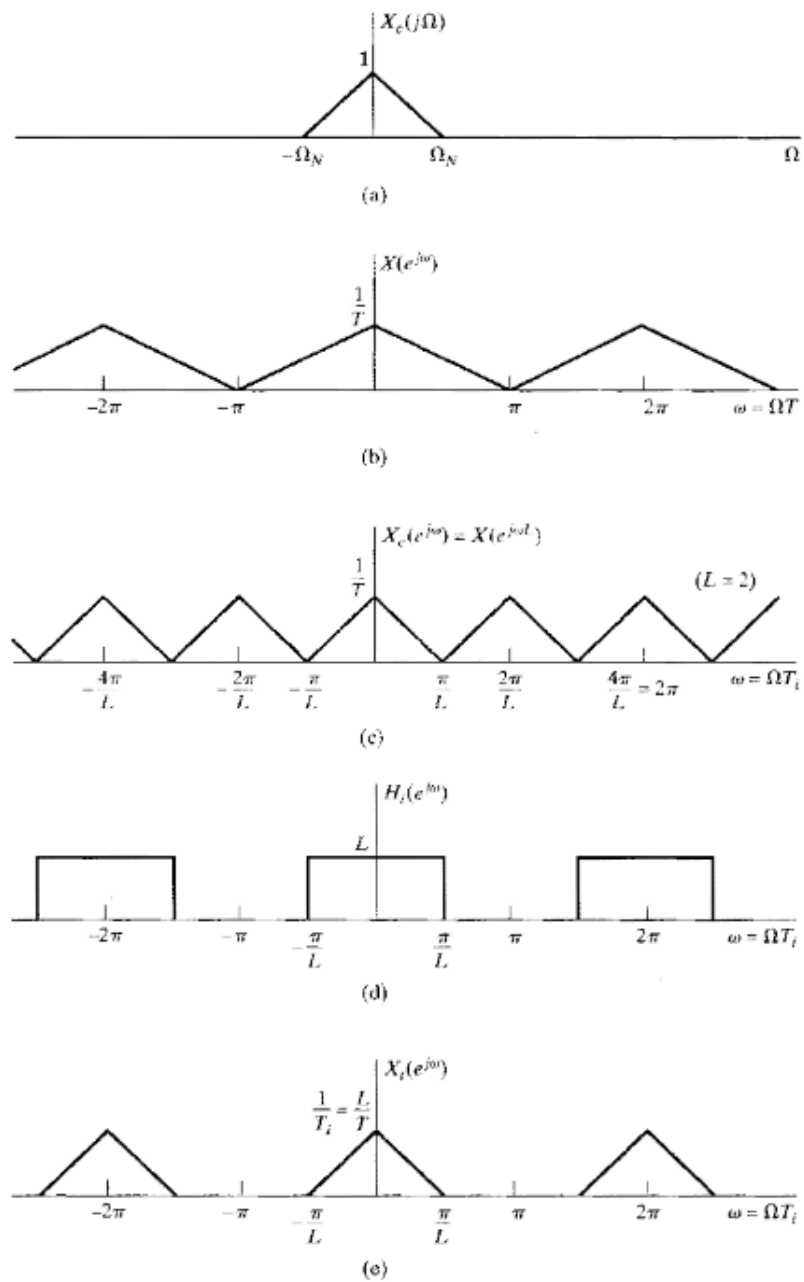


### 5.3.3 Compression and Decimation



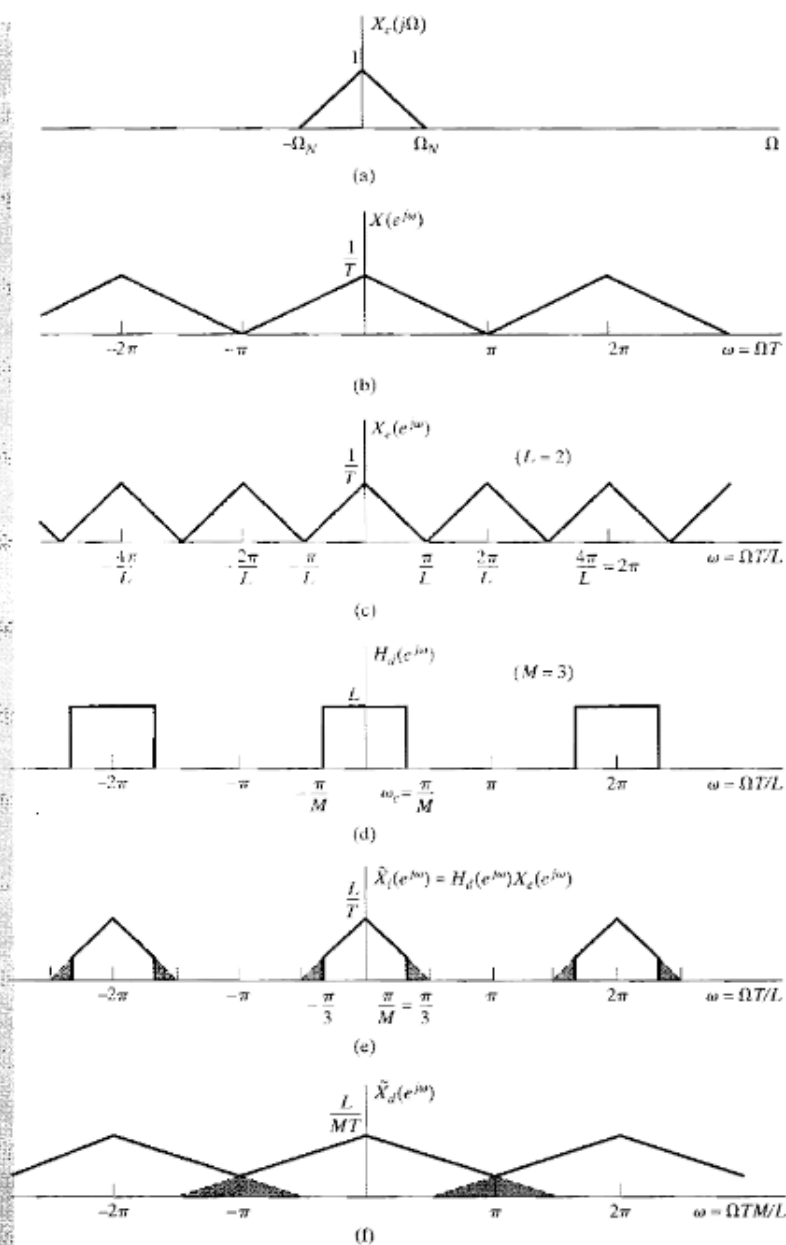
**Figure 4.21** (a)–(c) Downsampling with aliasing. (d)–(f) Downsampling with prefiltering to avoid aliasing.

### 5.3.4 Expansion and Interpolation



**Figure 4.24** Frequency-domain illustration of interpolation.

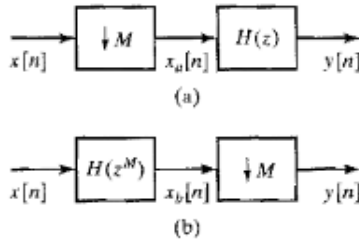
### 5.3.5 Change Sampling Rate by Noninteger Factor



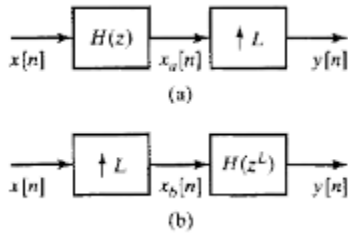
**Figure 4.30** Illustration of changing the sampling rate by a noninteger factor.

### 5.3.6 Polyphase Decompositions

Noble Identities:



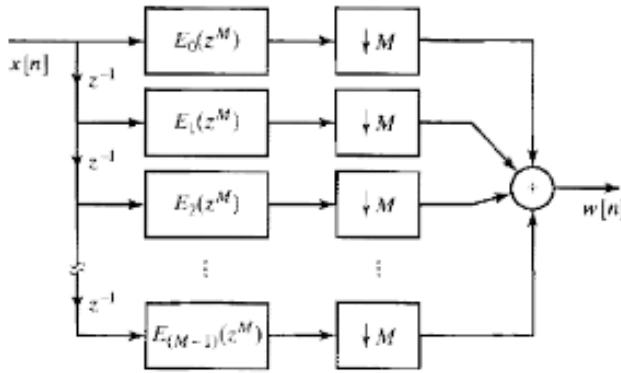
**Figure 4.31** Two equivalent systems based on downsampling identities.



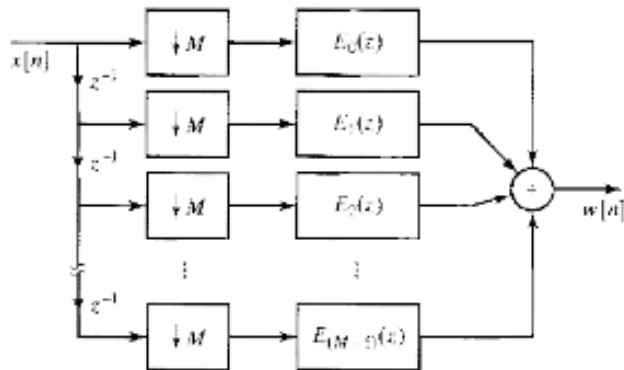
**Figure 4.32** Two equivalent systems based on upsampling identities.

Polyphase Decomposition for Compression:

In the below diagram  $e_k[n] = h[nM + k]$



**Figure 4.39** Implementation of decimation filter using polyphase decomposition.



**Figure 4.40** Implementation of decimation filter after applying the downsampling identity to the polyphase decomposition.