Chapter 1

Linear Systems Theory

1.1 Classification of systems

1.1.1 Continuous time versus discrete time

- Continuous time systems evolve with time indices \( t \in \mathbb{R} \).

- Discrete time systems evolve with time indices \( t \in \mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, \ldots\} \). Usually, the symbol \( k \) instead of \( t \) is used to denote discrete time indices.

1.1.2 Static versus dynamic systems

- A system with input \( u \) and output \( y \) is called a static system if \( \exists \) (there exists) a function \( f(u, t) \) such that for all \( t \in T \),
  \[
  y(t) = f(u(t), t). \tag{1.1}
  \]

- A static time invariant system is one with \( y(t) = f(u(t)) \) for all \( t \).

- To determine the output of a static system, only the present input is needed. In contrast, a (causal) dynamic system requires past input to determine the system output. i.e. to determine \( y(t) \) one needs to know \( u(\tau), \tau \in (-\infty, t] \).

1.2 A state determined dynamic system

The state of a dynamic system at time \( t_0 \), \( x(t_0) \), is the extra piece of information needed, so that given the input trajectory \( u(\tau), \tau \geq t_0 \), one is able to determine the behavior of the system for all times \( t \geq t_0 \). The behaviors are usually captured by defining appropriate outputs \( y(t) \).

- State is not unique. Two different pieces of information can both be valid states of the system.

- What constitutes a state depends on what behaviors are of interest.

- Some authors require a state to be a minimal piece of information. In these notes, we do not require this to be so.

An Example: Consider a car with input \( u(t) \) being its acceleration. Let \( y(t) \) be the position of the car.
Remark 1.2.1

1. If the behavior of interest is just the speed of the car, then \( x(t) = \dot{y}(t) \) can be used as the state. It is qualified to be a state because given \( u(\tau), \tau \in [t_0, t] \), the speed at \( t \) is obtained by:

\[
v(t) = \dot{y}(t) = x(t_0) + \int_{t_0}^{t} u(\tau) d\tau.
\]

2. If the behavior of interest is the position of the car, then \( x_a(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix} \in \mathbb{R}^2 \) can be used as the state.

3. An alternate state might be \( x_b(t) = \begin{pmatrix} y(t) + 2\dot{y}(t) \\ \dot{y}(t) \end{pmatrix} \). Obviously, since we can determine the old state vector \( x_a(t) \) from this alternate one \( x_b(t) \), and vice versa, both are valid state vectors. This illustrates the point that state vectors are not unique.

\[
\text{Remark 1.2.1}
\]

1. If \( y(t) \) is defined to be the behavior of interest, then by taking \( t = t_0 \), the definition of a state determined system implies that one can determine \( y(t) \) from the state \( x(t) \) and input \( u(t) \), at time \( t \). i.e. there is a static output function \( h(\cdot, \cdot, \cdot) \) so that the output \( y(t) \) is given by:

\[
y(t) = h(x(t), u(t), t)
\]

\( h : (x, u, t) \mapsto y(t) \) is called the output readout map.

2. The usual representation of continuous time dynamical system is given by the form:

\[
\begin{align*}
\dot{x} &= f(x, u, t) \\
y &= h(x, u, t)
\end{align*}
\]

and for discrete time system,

\[
\begin{align*}
x(k + 1) &= f(x(k), u(k), k) \\
y(k) &= h(x(k), u(k), k)
\end{align*}
\]

3. Notice that a state determined dynamic system defines, for every pair of initial and final times, \( t_0 \) and \( t_1 \), a mapping (or transformation) of the initial state \( x(t_0) = x_0 \) and input trajectory \( u(\tau), \tau \in [t_0, t_1] \) to the state at a time \( t_1 \), \( x(t_1) \). In these notes, we shall use the notation: \( s(t_1, t_0, x_0, u(\cdot)) \) to denote this state transition mapping. i.e.

\[
x(t_1) = s(t_1, t_0, x_0, u(\cdot))
\]

if the initial state at time \( t_0 \) is \( x_0 \), and the input trajectory is given by \( u(\cdot) \).

A state transition map must satisfy two important properties:

- **State transition property** For any \( t_0 \leq t_1 \), if two input signals \( u_1(t) \) and \( u_2(t) \) are such that \( u_1(t) = u_2(t) \ \forall t \in [t_0, t_1] \), then

\[
s(t_1, t_0, x_0, u_1(\cdot)) = s(t_1, t_0, x_0, u_2(\cdot))
\]

i.e. if \( x(t_0) = x_0 \), then the final state \( x(t_1) \) depends only on past inputs (from \( t_1 \) that occur after \( t_0 \), when the initial state is specified. Systems like this is called *causal* because the state does not depend on future inputs.
• **Semi-group property** For all \( t_2 \geq t_1 \geq t_0 \in T \), for all \( x_0 \), and for all \( u(\cdot) \),

\[
s(t_2, t_1, x(t_1), u) = s(t_2, t_1, s(t_1, t_0, x_0, u), u) = s(t_2, t_0, x_0, u)
\]

Thus, when calculating the state at time \( t_2 \), we can first calculate the state at some intermediate time \( t_1 \), and then utilize this result to calculate the state at time \( t_2 \) in terms of \( x(t_1) \) and \( u(t) \) for \( t \in [t_1, t_2] \).

### 1.3 Linear Differential Systems

The most common systems we will study are of the form:

\[
\dot{x} = A(t)x + B(t)u \quad (= f(x, u, t)) \tag{1.2}
\]

\[
y = C(t)x + D(t)u \quad (= h(x, u, t)) \tag{1.3}
\]

where \( x(t) \in \mathbb{R}^n \), \( u : [0, \infty) \to \mathbb{R}^m \) and \( y : [0, \infty) \to \mathbb{R}^p \). \( A(t) \), \( B(t) \), \( C(t) \), \( D(t) \) are matrices with compatible dimensions of real-valued functions.

This system may have been obtained from the Jacobian linearization of a nonlinear system

\[
\dot{x} = f(x, t), \quad y = h(t, x, u)
\]

about a pair of nominal input and state trajectories \( (\bar{u}(t), \bar{x}(t)) \).

**Assumption 1.3.1** We assume that \( A(t) \), \( B(t) \) and the input \( u(t) \) are piecewise continuous functions of \( t \). In other words, a function \( f(t) \) is piecewise continuous in \( t \) if it is continuous in \( t \) except possibly at points of a set which contains at most a finite number of points per unit interval. Also, at points of discontinuity, both the left and right limits exist.

**Remark 1.3.1**

- **Assumption 1.3.1** allows us to claim that given an initial state \( x(t_0) \), and a input function \( u(\tau), \tau \in [t_0, t_1] \), the solution,

\[
x(t) = s(t, t_0, x(t_0), u(\cdot)) = x(t_0) + \int_{t_0}^{t} [A(\tau)x(\tau) + B(\tau)u(\tau)]d\tau
\]

exists and is unique via the Fundamental Theorem of Ordinary Differential Equations (see below) often proved in a course on nonlinear systems.

- **Uniqueness and existence of solutions of nonlinear differential equations are not generally guaranteed. Consider the following examples.**

An example in which existence is a problem:

\[
\dot{x} = 1 + x^2
\]

with \( x(t = 0) = 0 \). This system does not have a solution for \( t \geq \pi/2 \) \( (x(t) = \tan t \) otherwise. This phenomenon is called finite escape.

The other issue is whether if a differential equation can have more than one solution for the same initial condition (non-uniqueness). e.g. for the system:

\[
\dot{x} = 3x^2, \quad x(0) = 0.
\]

both \( x(t) = 0, \forall t \geq 0 \) and \( x(t) = t^3 \) are both valid solutions.
Theorem 1.3.1 (Fundamental Theorem of ODE’s) Consider the following ODE
\[ \dot{x} = f(x, t) \] (1.4)
where \( x(t) \in \mathbb{R}^n, t \geq 0, \) and \( f: \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^n. \)

On a time interval \([t_0, t_1] \), if the function \( f(x, t) \) satisfies

1. For any \( x \in \mathbb{R}^n \), the function \( t \mapsto f(x, t) \) is piecewise continuous,
2. There exists a piecewise continuous function, \( k(t) \), such that for all \( x_1, x_2 \in \mathbb{R}^n \),
\[ \|f(x_1, t) - f(x_2, t)\| \leq k(t)\|x_1 - x_2\|. \]

Then,

1. there exists a solution to the differential equation \( \dot{x} = f(x, t) \) for all \( t \), meaning that: a) For each initial state \( x_0 \in \mathbb{R}^n \), there is a continuous function \( \phi: \mathbb{R}_+ \rightarrow \mathbb{R}^n \) such that
\[ \phi(t_0) = x_0 \]
and
\[ \dot{\phi}(t) = f(\phi, t) \quad \forall t \in \mathbb{R} \]

2. Moreover, the function \( \phi \) is unique. [If \( \phi_1(\cdot) \) and \( \phi_2(\cdot) \) have the same properties above, then they must be the same.]

Remark 1.3.2 This version of the fundamental theorem of ODE is taken from [Callier and Desoer, 91]. A weaker condition (and result) is that on the interval \([t_0, t_1] \), the second condition says that there exists a constant \( k_{[t_0, t_1]} \) such that for any \( t \in [t_0, t_1] \), the function \( x \mapsto f(x, t) \), for all \( x_1, x_2 \in \mathbb{R}^n \), satisfies
\[ \|f(x_1, t) - f(x_2, t)\| \leq k_{[t_0, t_1]}\|x_1 - x_2\|. \]

In this case, the solution exists and is unique on \([t_0, t_1] \).

The proof of this result can be found in many books on ODE’s or dynamic systems and is usually proved in details in a Nonlinear Systems Analysis class.

1.3.1 Linearity Property

The reason why the system (1.2)-(1.3) is called a linear differential system is because of the following linearity property.

Theorem 1.3.2 For any pairs of initial and final times \( t_0, t_1 \in \mathbb{R} \), the state transition map
\[ s: (t_0, t_1, x(t_0), u(\cdot)) \mapsto x(t_1) \]
of the linear differential system (1.2)-(1.3) is a linear map of the pair of the initial state \( x(t_0) \) and the input \( u(\tau), \tau \in [t_0, t_1] \).

\[ ^1 \text{e.g. Vidyasagar, Nonlinear Systems Analysis, 2nd Prentice Hall, 93, or Khalil Nonlinear Systems, McMillan, 92} \]
In order words, for any \( u(\cdot), u'(\cdot) \in \mathbb{R}^{n}_{[t_0,t_1]}, \ x, x' \in \mathbb{R}^n \) and \( \alpha, \beta \in \mathbb{R} \),
\[
s(t, t_0, (\alpha x + \beta x'), (\alpha u(\cdot) + \beta u'(\cdot))) = \alpha \cdot s(t, t_0, x, u(\cdot)) + \beta \cdot s(t, t_0, x', u'(\cdot)).
\] (1.5)

Similarly, for each pair of \( t_0, t_1 \), the mapping
\[
\rho : (t_0, t_1, x(t_0), u(\cdot)) \mapsto y(t_1)
\]
from the initial state \( x(t_0) \) and the input \( u(\tau), \tau \in [t_0, t_1] \), to the output \( y(t_1) \) is also a linear map, i.e.
\[
\rho(t, t_0, (\alpha x + \beta x'), (\alpha u(\cdot) + \beta u'(\cdot))) = \alpha \cdot \rho(t, t_0, x, u(\cdot)) + \beta \cdot \rho(t, t_0, x', u'(\cdot)).
\] (1.6)

Before we prove this theorem, let us point out a very useful principle for proving that two time functions \( x(t) \), and \( x'(t) \), \( t \in [t_0, t_1] \), are the same.

**Lemma 1.3.3** Given two time signals, \( x(t) \) and \( x'(t) \). Suppose that
- \( x(t) \) and \( x'(t) \) satisfies the same differential equation,
  \[
  \dot{p} = f(t, p)
  \] (1.7)
- they have the same initial conditions, i.e. \( x(t_0) = x'(t_0) \)
- The differential equation (1.7) has unique solution on the time interval \([t_0, t_1]\),

then \( x(t) = x'(t) \) for all \( t \in [t_0, t_1] \).

**Proof:** (of Theorem 1.3.2) We shall apply the above lemma (principle) to (1.5).

Let \( t_0 \) be an initial time, \( (x_0, u(\cdot)) \) and \( (x', u'(\cdot)) \) be two pairs of initial state and input, producing state and output trajectories \( x(t) \), \( y(t) \) and \( x'(t) \), \( y'(t) \) respectively.

We need to show that if the initial state is \( x''(t_0) = \alpha x_0 + \beta x_0' \), and input \( u''(t) = \alpha u(t) + \beta u'(t) \) for all \( t \in [t_0, t_1] \), then at any time \( t \), the response \( y''(t) \) is given by the function
\[
y''(t) := \alpha y(t) + \beta y'(t).
\] (1.8)

We will first show that for all \( t \geq t_0 \), the state trajectory \( x''(t) \) is given by:
\[
x''(t) = \alpha x(t) + \beta x'(t).
\] (1.9)

Denote the RHS of (1.9) by \( g(t) \).

To prove (1.9), we use the fact that (1.2) has unique solutions. Clearly, (1.9) is true at \( t = t_0 \),
\[
x''(t_0) = \alpha x_0 + \beta x_0' = \alpha x(t_0) + \beta x'(t_0) = g(t_0).
\]

By definition, if \( x'' \) is a solution to (1.2),
\[
\dot{x}''(t) = A(t)x''(t) + (\alpha u(t) + \beta u'(t)).
\]

Moreover,
\[
\dot{g}(t) = \alpha \dot{x}(t) + \beta \dot{x}'(t)
\]
\[
= \alpha \left[ A(t)x(t) + B(t)u(t) \right] + \beta \left[ A(t)x'(t) + B(t)u'(t) \right]
\]
\[
= A(t)[\alpha x(t) + \beta x'(t)] + \alpha u(t) + \beta u'(t)
\]
\[
= A(t)g(t) + [\alpha u(t) + \beta u'(t)].
\]

Hence, \( g(t) \) and \( x''(t) \) satisfy the same differential equation (1.2). Thus, by the existence and uniqueness property of the linear differential system, the solution is unique for each initial time \( t_0 \) and initial state. Hence \( x''(t) = g(t) \).
1.4 Decomposition of the transition map

Because of the linearity property, the transition map of the linear differential system (1.2) can be decomposed into two parts:

\[ s(t, t_0, x_0, u) = s(t, t_0, x_0, 0_u) + s(t, t_0, 0, u) \]

where \( 0_u \) denotes the identically zero input function \( (u(\tau) = 0 \text{ for all } \tau) \). It is so because we can decompose \( (x_0, u) \in \mathbb{R}^n \times U \) into

\[ (x_0, u) = (x_0, 0_u) + (0, u) \]

and then apply the defining property of a linear dynamical system to this decomposition.

Because of this property, the zero-state response (i.e. the response of the system when the initial state is \( x(t_0) = 0 \)) satisfies the familiar superposition property:

\[ \rho(t, t_0, x = 0, \alpha u + \beta u') = \alpha \rho(t, t_0, x = 0, u) + \beta \rho(t, t_0, x = 0, u'). \]

Similarly, the zero-input response satisfies a superposition property:

\[ \rho(t, t_0, \alpha x + \beta x', 0_u) = \alpha \rho(t, t_0, x, 0_u) + \beta \rho(t, t_0, x', 0_u). \]

1.5 Zero-input transition and the Transition Matrix

From the proof of linearity of the Linear Differential Equation, we actually showed that the state transition function \( s(t, t_0, x_0, u) \) is linear with respect to \( (x_0, u) \). In particular, for zero input \( (u = 0_u) \), it is linear with respect to the initial state \( x_0 \). We call the transition when the input is identically 0, the zero-input transition.

It is easy to show, by choosing \( x_0 \) to be columns in an identity matrix successively (i.e. invoking the so called 1st representation theorem - Ref: Desoer and Callier, or Chen), that there exists a matrix, \( \Phi(t, t_0) \in \mathbb{R}^{n \times n} \) so that

\[ s(t, t_0, x_0, 0_u) = \Phi(t, t_0)x_0. \]

This matrix function is called the transition matrix.

Claim: \( \Phi(t, t_0) \) satisfies (1.2). i.e.

\[ \frac{\partial}{\partial t} \Phi(t, t_0) = A(t)\Phi(t, t_0) \] (1.10)

and \( \Phi(t_0, t_0) = I \).

Proof: Consider an arbitrary initial state \( x_0 \in \mathbb{R}^n \) and zero input. By definition,

\[ x(t) = \Phi(t, t_0)x_0 = s(t, t_0, x_0, 0_u). \]

Differentiating the above,

\[ \dot{x}(t) = \frac{\partial}{\partial t} \Phi(t, t_0)x_0 = A(t)\Phi(t, t_0)x_0. \]

Now pick successively \( n \) different initial conditions \( x_0 = e_1, x_0 = e_2, \ldots, x_0 = e_n \) so that \( \{e_1, \ldots, e_n\} \) form a basis of \( \mathbb{R}^n \). (We can take for example, \( e_i \) to be the \( i-th \) column of the identity matrix).

Thus,

\[ \frac{\partial}{\partial t} \Phi(t, t_0)(e_1 \ e_2 \ \cdots \ e_n) = A(t)\Phi(t, t_0)(e_1 \ e_2 \ \cdots \ e_n) \]

Since \( (e_1 \ e_2 \ \cdots \ e_n) \in \mathbb{R}^{n\times n} \) is invertible, we multiply both sides by its inverse to obtain the required answer.

\[ \diamond \]
Definition 1.5.1 A \( n \times n \) matrix \( X(t) \) that satisfies the system equation,
\[
\dot{X}(t) = A(t)X(t)
\]
and \( X(\tau) \) is non-singular for some \( \tau \), is called a fundamental matrix.

Remark 1.5.1 The transition matrix is a fundamental matrix. It is invertible at at least \( t = t_0 \).

Proposition 1.5.1 Let \( X(t) \in \mathbb{R}^{n \times n} \) be a fundamental matrix. Then, \( X(t) \) is non-singular at all \( t \in \mathbb{R} \).

Proof: Since \( X(t) \) is a fundamental matrix, \( X(t_1) \) is nonsingular for some \( t_1 \). Suppose that \( X(\tau) \) is singular, then \( \exists \) a non-zero vector \( k \in \mathbb{R}^n \) s.t. \( X(\tau)k = 0 \), the zero vector in \( \mathbb{R}^n \). Consider now the differential equation:
\[
\dot{x} = A(t)x, \quad x(\tau) = X(\tau)k = 0 \in \mathbb{R}^n.
\]
Then, \( x(t) = 0 \) for all \( t \) is the unique solution to this system with \( x(\tau) = 0 \). However, the function \( x'(t) = X(t)k \)
satisfies
\[
\dot{x'}(t) = A(t)X(t)k = A(t)x'(t),
\]
and \( x'(\tau) = x(\tau) = 0 \). Thus, by the uniqueness of differential equation, \( x(t) = x'(t) \) for all \( t \). Hence \( x'(t_1) = 0 \) extracting a contradiction because \( X(t_1) \) is nonsingular so \( x'(t_1) = X(t_1)k \) is non-zero.
\( \diamond \)

1.5.1 Properties of Transition matrices

1. (Existence and Uniqueness) \( \Phi(t, t_0) \) exists and is unique for each \( t, t_0 \in \mathbb{R} \) (\( t \geq t_0 \) is not necessary).

2. (Group property) For all \( t, t_1, t_0 \) (not necessarily \( t_0 \leq t \leq t_1 \))
\[
\Phi(t, t_0) = \Phi(t, t_1)\Phi(t_1, t_0).
\]

3. (Non-singularity) \( \Phi(t, t_0) = [\Phi(t_0, t)]^{-1} \).

4. (Spliting property) If \( X(t) \) is any fundamental matrix, then
\[
\Phi(t, t_0) = X(t)X(t_0)^{-1}
\]

5. \( \frac{\partial}{\partial t}\Phi(t, t_0) = A(t)\Phi(t, t_0) \), and \( \Phi(t_0, t_0) = I \) for each \( t_0 \in \mathbb{R} \).

6. \( \frac{\partial}{\partial t_0}\Phi(t, t_0) = -\Phi(t, t_0)A(t_0) \).

Proof:

1. This comes from the fact that the solution \( s(t, t_0, x_0) \) exists for all \( t, t_0 \in \mathbb{R} \) and for all \( x_0 \in \mathbb{R}^n \) (fundamental theorem of differential equation) and that \( s(t, t_0, x_0, 0_n) \) is linear in \( x_0 \).
2. Differentiate both sides with respect to \( t \) to find that RHS and LHS satisfy the same differential equation and have the same value at \( t = t_1 \). So apply existence and uniqueness theorem.

3. From 2), take \( t = t_0 \).

4. Use the existence and uniqueness theorem again: For each \( t_0 \), consider both sides as functions of \( t \). So, \( LHS(t = t_0) = RHS(t = t_0) \). Now, the LHS satisfies, \( \frac{d}{dt} \Phi(t, t_0) = A(t)\Phi(t, t_0) \). The RHS satisfies:

\[
\frac{d}{dt}[X(t)X(t_0)] = \left[ \frac{d}{dt} X(t) \right] X(t_0) = A(t)X(t)X(t_0)^{-1}
\]

Hence \( \frac{d}{dt} RHS(t) = A(t)RHS(t) \). So \( LHS \) and \( RHS \) satisfy same differential equation and agree at \( t = t_0 \). Hence they must be the same at all \( t \).

5. Already shown.

6. From 3) we have \( \Phi(t, t_0)\Phi(t_0, t) = I \), the identity matrix. Differentiate both sides with respect to \( t_0 \),

\[
\left[ \frac{\partial}{\partial t_0} \Phi(t, t_0) \right] \Phi(t_0, t) + \Phi(t, t_0) \left[ \frac{\partial}{\partial t_0} \Phi(t_0, t) \right] = 0
\]

since \( \frac{d}{dt_0}[X(t_0)Y(t_0)] = \frac{d}{dt_0}X(t_0)Y(t_0) + X(t_0)\frac{d}{dt_0}Y(t_0) \) for \( X(\cdot), Y(t_0) \) a matrix of functions of \( t_0 \) (verify this!). Hence,

\[
\left[ \frac{\partial}{\partial t_0} \Phi(t, t_0) \right] \Phi(t_0, t) = -\Phi(t, t_0) \left[ \frac{\partial}{\partial t_0} \Phi(t_0, t) \right]
\]

\[
\Rightarrow \left[ \frac{\partial}{\partial t_0} \Phi(t, t_0) \right] \Phi(t_0, t) = -\Phi(t, t_0)A(t_0)\Phi(t_0, t)
\]

\[
\Rightarrow \frac{\partial}{\partial t_0} \Phi(t, t_0) = -\Phi(t, t_0)A(t_0)\Phi(t_0, t)\Phi(t_0, t)^{-1} = -\Phi(t, t_0)A(t_0)
\]

\( \diamond \)

### 1.5.2 Explicit formula for \( \Phi(t, t_0) \)

The Peano-Baker formula is given by:

\[
\Phi(t, t_0) = I + \int_{t_0}^{t} A(\sigma) \, d\sigma + \int_{t_0}^{t} A(\sigma_1) \left[ \int_{t_0}^{\sigma_1} A(\sigma_2) \, d\sigma_2 \right] \, d\sigma_1 \\
+ \int_{t_0}^{t} A(\sigma) \left[ \int_{t_0}^{\sigma_1} A(\sigma_2) \left[ \int_{t_0}^{\sigma_2} A(\sigma_3) \, d\sigma_3 \right] \, d\sigma_2 \right] \, d\sigma_1 + \ldots \quad (1.11)
\]

This formula can be verified formally (do it) by checking that the \( RHS(t = t_0) \) is indeed \( I \) and that the RHS satisfies the differential equation

\[
\frac{\partial}{\partial t} RHS(t, t_0) = A(t)RHS(t, t_0).
\]

Let us define the exponential of a matrix, \( A \in \mathbb{R}^{n \times n} \) using the power series,

\[
exp(A) = I + \frac{A}{1} + \frac{A^2}{2!} + \cdots \frac{A^k}{k!} + \cdots
\]
by mimicking the power series expansion of the exponential function of a real number.

If \( A(t) = A \) is a constant, then the Peano-Bakar formula reduces to (prove it!)

\[
\Phi(t, t_0) = e^{(t - t_0) \cdot A}
\]

For example, examining the 4th term,

\[
\int_{t_0}^{t} A(\sigma_1) \left[ \int_{t_0}^{\sigma_2} A(\sigma_2) \left[ \int_{t_0}^{\sigma_3} A(\sigma_3) d\sigma_3 \right] d\sigma_2 \right] d\sigma_1
\]

\[
= A^3 \int_{t_0}^{t} \int_{t_0}^{\sigma_1} \int_{t_0}^{\sigma_2} d\sigma_3 d\sigma_2 d\sigma_1
\]

\[
= A^3 \int_{t_0}^{t} (\sigma_2 - t_0) d\sigma_2 d\sigma_1
\]

\[
= A^3 \int_{t_0}^{t} \frac{(\sigma_1 - t_0)^2}{2} d\sigma_1 = A^3 \frac{(t - t_0)^3}{3!}
\]

A more general case is that if \( A(t) \) and \( \int_{t_0}^{t} A(\tau) d\tau \), commute for all \( t \),

\[
\Phi(t, t_0) = e^{\int_{t_0}^{t} A(\tau) d\tau}
\]

(1.12)

An intermediate step is to show that:

\[
\frac{1}{k + 1} \frac{d}{dt} \left[ \int_{t_0}^{t} A(\tau) d\tau \right]^{k+1} = A(t) \left[ \int_{t_0}^{t} A(\tau) d\tau \right]^k = \left[ \int_{t_0}^{t} A(\tau) d\tau \right]^k A(t).
\]

Notice that (1.12) does not in general apply unless the matrices commute. i.e.

\[
A(t) \left[ \int_{t_0}^{t} A(\tau) d\tau \right]^k = \left[ \int_{t_0}^{t} A(\tau) d\tau \right]^k A(t).
\]

Each of the following situations will guarantee that the condition \( A(t) \) and \( \int_{t_0}^{t} A(\tau) d\tau \) commute for all \( t \) is satisfied:

1. \( A(t) \) is constant.
2. \( A(t) = \alpha(t)M \) where \( \alpha(t) \) is a scalar function, and \( M \in \mathbb{R}^{n \times n} \) is a constant matrix;
3. \( A(t) = \sum_i \alpha_i(t)M_i \) where \( \{M_i\} \) are constant matrices that commute: \( M_iM_j = M_jM_i \), and \( \alpha_i(t) \) are scalar functions;
4. \( A(t) \) has a time-invariant basis of eigenvectors spanning \( \mathbb{R}^n \).

1.5.3 Computation of \( \Phi(t, 0) = e^{\text{exp}(tA)} \) for constant \( A \)

The defining computational method for \( \Phi(t, t_0) \) not necessarily for time varying \( A(t) \) is:

\[
\frac{\partial}{\partial t} \Phi(t, t_0) = A(t)\Phi(t, t_0); \quad \Phi(t_0, t_0) = I
\]

When \( A(t) = \) constant, there are algebraic approaches available:

\[
\Phi(t_1, t_0) = e^{\text{exp}(A(t_1 - t_0))}
\]

\[
:= I + \frac{A(t_1 - t_0)}{1!} + \frac{A^2(t_1 - t_0)^2}{2!} + \ldots
\]

Matlab provides \( \text{expm}(A \ast (t_1 - t_0)) \).
Laplace transform approach

\[ \mathcal{L}(\Phi(t,0)) = [sI - A]^{-1} \]

Proof: Since \( \Phi(0,0) = I \), and
\[ \frac{\partial}{\partial t} \Phi(t,0) = A(t)\Phi(t,0) \]
take Laplace transform of the (Matrix) ODE:
\[ s\hat{\Phi}(s,0) - \Phi(0,0) = A\hat{\Phi}(s,0) \]
where \( \hat{\Phi}(s) \) is the Laplace transform of \( \Phi(t,0) \) when treated as function of \( t \). This gives:
\[ (sI - A)\hat{\Phi}(s,0) = I \]

Example

\[ A = \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix} \]

\[ (SI - A)^{-1} = \begin{pmatrix} s + 1 & 0 \\ -1 & s + 2 \end{pmatrix}^{-1} = \frac{1}{(s + 1)(s + 2)} \begin{pmatrix} s + 2 & 0 \\ 1 & s + 1 \end{pmatrix} \]

Taking the inverse Laplace transform (using e.g. a table) term by term,
\[ \Phi(t,0) = \begin{pmatrix} e^{-t} & 0 \\ e^{-t} - e^{-2t} & e^{-2t} \end{pmatrix} \]

Similarity Transform (decoupled system) approach

Let \( A \in \mathbb{R}^{n \times n} \), if \( v \in \mathbb{C}^n \) and \( \lambda \in \mathbb{C} \) satisfy
\[ A \cdot v = \lambda v \]
then, \( v \) is an eigenvector and \( \lambda \) its associated eigenvalue.

If \( A \in \mathbb{R}^{n \times n} \) has \( n \) distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \), then \( A \) is called simple. If \( A \in \mathbb{R}^{n \times n} \) has \( n \) independent eigenvectors then \( A \) is called semi-simple (Notice that \( A \) is necessarily semi-simple if \( A \) is simple).

Suppose \( A \in \mathbb{R}^{n \times n} \) is simple, then let
\[ T = (v_1 \quad v_2 \quad \ldots \quad v_n); \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \]
be the collection of eigenvectors, and the associated eigenvalues. Then, the so-called similarity transform is:
\[ A = T\Lambda T^{-1} \quad (1.13) \]

Remark

- A sufficient condition for \( A \) being semi-simple is if \( A \) has \( n \) distinct eigenvalues (i.e. it is simple);
When $A$ is not semi-simple, a similar decomposition as (1.13) is available except that $A$ will be in Jordan form (has 1’s in the super diagonals) and $T$ consists of eigenvectors and generalized eigenvectors. This topic is covered in most Linear Systems or linear algebra textbook.

Now

$$\exp(At) := I + \frac{At}{1!} + \frac{A^2t^2}{2!} + \ldots$$

Notice that $A^k = T\Lambda^k T^{-1}$, thus,

$$\exp(tA) = T \left[I + \frac{At}{1!} + \frac{A^2t^2}{2!} + \ldots\right] T^{-1} = T \exp(\Lambda t) T^{-1}$$

The above formula is valid even if $A$ is not semi-simple. In the semi-simple case,

$$\Lambda^k = \text{diag} \left[\lambda^k_1, \lambda^k_2, \ldots, \lambda^k_n\right]$$

so that

$$\exp(\Lambda t) = \text{diag} [\exp(\lambda_1 t), \exp(\lambda_2 t), \ldots, \exp(\lambda_n t)]$$

If we write $T^{-1} = \begin{pmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n^T \end{pmatrix}$ where $w_i^T$ is the $i$-th row of $T^{-1}$, then,

$$\exp(tA) = T \exp(\Lambda t) T^{-1} = \sum_{i=1}^{n} \exp(\lambda_i t) v_i w_i^T.$$ 

This is the dyadic expansion of $\exp(tA)$. Can you show that $w_i^T A = \lambda_i w_i^T$? This means that $w_i$ are the left eigenvectors of $A$.

The expansion shows that the system has been decomposed into a set of simple, decoupled 1st order systems.

**Example**

$$A = \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix}$$

The eigenvalues and eigenvectors are:

$$\lambda_1 = -1, \ v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \ \lambda_2 = -2, \ v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$ 

Thus,

$$\exp(At) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 \\ e^{-t} - e^{-2t} & e^{-2t} \end{pmatrix},$$

same as by the Laplace transform method.

**Digression: Modal decomposition** The (generalized) eigenvectors are good basis for a coordinate system.

\(^2\)for the case the $A$ is not semi-simple, additional “nice” vectors are needed to form a basis, they are called generalized eigenvectors.
Suppose that $A = T \Lambda T^{-1}$ is the similarity transform. Let $x \in \mathbb{R}^n$, $z = T^{-1}x$, and $x = Tz$. Thus,

$$x = z_1v_1 + z_2v_2 + \ldots + z_nv_n,$$

where $T = [v_1, v_2, \ldots, v_n] \in \mathbb{R}^{n \times n}$. Hence, $x$ is decomposed into the components in the direction of the eigenvectors with $z_i$ being the scaling for the $i$-th eigenvector.

The original linear differential equation is written as:

$$\dot{x} = Ax + Bu$$
$$T\dot{z} = T\Lambda z + Bu$$
$$\dot{z} = \Lambda z + T^{-1}Bu$$

If we denote $\bar{B} = T^{-1}B$, then, since $\Lambda$ is diagonal,

$$\dot{z}_i = \lambda_i z_i + \bar{B}_i u$$

where $z_i$ is the $i$-th component of $z$ and $\bar{B}_i$ is the $i$-th row of $\bar{B}$.

This is a set of $n$ decoupled first order differential equations that can be analyzed independently. Dynamics in each eigenvector direction is called a mode. If desired, $z$ can be used to re-constitute $x$ via $x = Tz$.

### 1.6 Zero-State transition and response

Recall that for a linear differential system,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \quad x(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^m$$

the state transition map can be decomposed into the zero-input response and the zero-state response:

$$s(t, t_0, x_0, u) = \Phi(t, t_0)x_0 + s(t, t_0, 0, u)$$

Having figured out the zero-input state component, we now derive the zero-state response.

#### 1.6.1 Heuristic guess

We first decompose the inputs into piecewise continuous parts $\{u_i : \mathbb{R} \to \mathbb{R}^m\}$ for $i = \ldots, -2, -1, 1, 0, 1, \ldots$,

$$u_i(t) = \begin{cases} u(t_0 + h \cdot i) & t_0 + h \cdot i \leq t < t_0 + h \cdot (i + 1) \\ 0 & \text{otherwise} \end{cases}$$

where $h > 0$ is a small positive number. Intuitively we can see that as $h \to 0$,

$$u(t) = \sum_{i=-\infty}^{\infty} u_i(t) \text{ as } h \to 0.$$

Let $\bar{u}(t) = \sum_{i=-\infty}^{\infty} u_i(t)$. By linearity of the transition map,

$$s(t, t_0, 0, \bar{u}) = \sum_i s(t, t_0, 0, u_i).$$

Now we figure out $s(t, t_0, 0, u_i)$.
1.6. ZERO-STATE TRANSITION AND RESPONSE

- **Step 1:** \( t_0 \leq t < t_0 + h \cdot i \).
  Since \( u(\tau) = 0, \tau \in [t_0, t_0 + h \cdot i) \) and \( x(t_0) = 0 \),
  \[ x(t) = 0 \quad \text{for } t_0 \leq t < t_0 + h \cdot i \]

- **Step 2:** \( t \in [t_0 + h \cdot i, t_0 + h(i + 1)) \).
  Input is active:
  \[ x(t) \approx x(t_0 + h \cdot i) + [A(t_0 + h \cdot i) + B(t_0 + h \cdot i)u(t_0 + h \cdot i)] \cdot \Delta T \]
  \[ = [B(t_0 + h \cdot i)u(t_0 + h \cdot i)] \cdot \Delta T \]
  where \( \Delta T = t - t_0 + h \cdot i \).

- **Step 3:** \( t \geq t_0 + h \cdot (i + 1) \).
  Input is no longer active, \( u_i(t) = 0 \). So the state is again given by the zero-input transition map:
  \[ \Phi(t, t_0 + h \cdot (i + 1)) B(t_0 + i \cdot h)u(t_0 + h \cdot i) \approx x(t_0 + (i + 1) \cdot h) \]
  Since \( \Phi(t, t_0) \) is continuous, if we make the approximation
  \[ \Phi(t, t_0 + (h + 1)i) \approx \Phi(t, t_0 + h \cdot i) \]
  we only introduce second order error in \( h \). Hence,
  \[ s(t, t_0, x_0, u_i) \approx \Phi(t, t_0 + h \cdot i)B(t_0 + h \cdot i)u(t_0 + h \cdot i). \]

The total zero-state state transition due to the input \( u(\cdot) \) is therefore given by:

\[ s(t, t_0, 0, u) \approx \sum_{i=0}^{(t-t_0)/h} \Phi(t, t_0 + h \cdot i) B(t_0 + i \cdot h)u(t_0 + h \cdot i) \]

As \( h \to 0 \), the sum becomes an integral so that:

\[ s(t, t_0, 0, u) = \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau. \quad (1.14) \]

In this heuristic derivation, we can see that \( \Phi(t, \tau)B(\tau)u(\tau) \) is the contribution to the state \( x(t) \) due to the input \( u(\tau) \) for \( \tau < t \).

### 1.6.2 Formal Proof of zero-state transition map

We will show that for all \( t, t_0 \in \mathbb{R}_+ \),

\[ (x(t) = s(t, t_0, 0, u)) = \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau. \quad (1.15) \]

Clearly (1.15) is correct for \( t = t_0 \). We will now show that the LHS of (1.15), i.e. \( x(t) \) and the RHS of (1.15), which we will denote by \( z(t) \) satisfy the same differential equation.
We know that $x(t)$ satisfies, $\dot{x}(t) = A(t)x(t) + B(t)u(t)$.

Observe first that since $\frac{d}{dt}\Phi(t, t_0) = A(t)\Phi(t, t_0)$,

$$\Phi(t, \tau) = I + \int_{\tau}^{t} A(\sigma)\Phi(\sigma, \tau)d\sigma.$$  

Now for the RHS of (1.15) which we will call $z(t)$,

$$z(t) := \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau$$

$$= \int_{t_0}^{t} B(\tau)u(\tau)d\tau + \int_{t_0}^{t} \left\{ \int_{\tau}^{\sigma} A(\sigma)\Phi(\sigma, \tau)d\sigma \right\} B(\tau)u(\tau)d\tau$$

let $f(\sigma, \tau) := A(\sigma)\Phi(\sigma, \tau)B(\tau)u(\tau)$ and then changing the order of the integral (see fig 1.1)

$$= \int_{t_0}^{t} B(\tau)u(\tau)d\tau + \int_{t_0}^{t} \int_{t_0}^{\sigma} f(\sigma, \tau)d\tau d\sigma$$

$$= \int_{t_0}^{t} B(\tau)u(\tau)d\tau + \int_{t_0}^{t} A(\sigma)\int_{t_0}^{\sigma} \Phi(\sigma, \tau)B(\tau)u(\tau)d\tau d\sigma$$

$$= \int_{t_0}^{t} B(\tau)u(\tau)d\tau + \int_{t_0}^{t} A(\sigma)z(\sigma)d\sigma$$

which on differentiation w.r.t. $t$ gives

$$\dot{z}(t) = B(t)u(t) + A(t)z(t).$$

Hence, both $z(t)$ and $x(t)$ satisfy the same differential equation and have the same values at $t = t_0$. Therefore, $x(t) = z(t)$ for all $t$.

1.6.3 Output Response function

The combined effect of initial state $x_0$ and input function $u(\cdot)$ on the state is given by:

$$s(t, t_0, x_0, u) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau.$$
Deriving the output response is simple since:

\[ y(t) = C(t)x(t) + D(t)u(t). \]

where \( C(t) \in \mathbb{R}^{p \times n}, D(t) \in \mathbb{R}^{p \times m} \). Hence, the output response map \( y(t) = \rho(t, t_0, x_0, u) \) is simply

\[ y(t) = \rho(t, t_0, x_0, u) = C(t)\Phi(t, t_0)x_0 + C(t) \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau + D(t)u(t). \] (1.16)

### 1.6.4 Impulse Response Matrix

When the initial state is 0, let the input be an Dirac impulse occurring at time \( \tau \) in the \( j \)-th input,

\[ u_j(t) = \epsilon_j \delta(t - \tau) \]

where \( \epsilon_j \) is the \( j \)-th unit vector.

The output response is:

\[ y_j(t) = [C(t)\Phi(t, \tau)B(\tau) + D(t)\delta(t - \tau)]\epsilon_j \]

The matrix

\[ H(t, \tau) = \begin{cases} [C(t)\Phi(t, \tau)B(\tau) + D(t)\delta(t - \tau)] & \forall t \geq \tau \\ 0 & t < \tau \end{cases} \] (1.17)

is called the Impulse response matrix. The \( j \)-th column of \( H(t, \tau) \) signifies the output response of the system when an impulse is applied at input channel \( j \) at time \( \tau \). The reason why it must be zero for \( t < \tau \) is because the impulse would have no effect on the system before it is applied.

Because \( u(t) \approx \sum_{\tau} H(t, \tau)u(\tau)d\tau \), we can also see that intuitively,

\[ y(t) = \int_{t_0}^{t} H(t, \tau)u(\tau)d\tau \]

if the state \( x(t_0) = 0 \). This agrees with (1.16).

### 1.7 Linear discrete time system response

Discrete time systems are described by difference equation. For any (possibly nonlinear) difference equation:

\[ x(k + 1) = f(k, x(k), u(k)) \]

with initial condition \( x(k_0) \), the solution \( x(k \geq k_0) \) exists and is unique as long as \( f(k, x(k), u(k)) \) is a properly defined function (i.e. \( f(k, x, u) \) is well defined for given \( k, x, u \)). This can be shown by just recursively applying the difference equation forward in time. The existence of the solution backwards in time is not guaranteed, however.

The linear discrete time system is given by the difference equation:

\[ x(k + 1) = A(k)x(k) + B(k)u(k); \]
with initial condition given by \( x(k_0) = x_0 \).

Many properties of linear discrete time systems are similar to the linear differential (continuous time) systems. We now study some of these, and point out some differences.

Let the transition map be given by \( s(k_1, k_0, x_0, u(\cdot)) \) where
\[
x(k_1) = s(k_1, k_0, x_0, u(\cdot)).
\]

**Linearity:** For any \( \alpha, \beta \in \mathbb{R} \), and for any two initial states \( x_a, x_b \) and two input signals \( u_a(\cdot) \) and \( u_b(\cdot) \),
\[
s(k_1, k_0, \alpha x_a + \beta x_b, \alpha u_a(\cdot) + \beta u_b(\cdot)) = \alpha s(k_1, k_0, x_a, u_a(\cdot)) + \beta s(k_1, k_0, x_b, u_b(\cdot))
\]

As corollaries, we have:

1. **Decomposition into zero-input response and zero-state response:**
\[
s(k_1, k_0, x_0, u(\cdot)) = s(k_1, k_0, x_0, 0) + s(k_1, k_0, 0, u(\cdot))
\]
2. Zero input response can be expressed in terms of a transition matrix,
\[
x(k := s(k_1, k_0, x_0, 0) = \Phi(k, k_0)x_0
\]

### 1.7.1 \( \Phi(k, k_0) \) properties

The discrete time transition function can be explicitly written as:
\[
\Phi(k, k_0) = \prod_{k'=k_0}^{k-1} A(k')
\]

The discrete time transition matrix has many similar properties as the continuous one, particularly in the case of \( A(k) \) is invertible for all \( k \). The main difference results from the possibility that \( A(k) \) may not be invertible.

1. **Existence and uniqueness:** \( \Phi(k_1, k_0) \) exists and is unique for all \( k_1 \geq k_0 \). If \( k_1 < k_0 \), then \( \Phi(k_1, k_0) \) exists and is unique if and only if \( A(k) \) is invertible for all \( k_0 > k \geq k_1 \).
2. **Existence of inverse:** If \( k_1 \geq k_0 \), \( \Phi(k_1, k_0)^{-1} \) exists and is given by
\[
\Phi(k_1, k_0)^{-1} = A(k_0)^{-1}A(k_0+1)^{-1}\cdots A(k_1-1)^{-1}
\]
if and only if \( A(k) \) is invertible for all \( k_1 > k \geq k_0 \).
3. **Semi-group property:**
\[
\Phi(k_2, k_0) = \Phi(k_2, k_1)\Phi(k_1, k_0)
\]
for \( k_2 \geq k_1 \geq k_0 \) only, unless \( A(k) \) is invertible.
4. **Matrix difference equations:**
\[
\Phi(k+1, k_0) = A(k)\Phi(k, k_0) \\
\Phi(k_1, k-1) = \Phi(k_1, k)A(k-1)
\]

Can you formulate a property for discrete time transition matrices similar to the splitting property for continuous time case?
1.7.2 Zero-initial state response

The zero-initial state response can be obtained easily

\[
x(k) = A(k - 1)x(k - 1) + B(k - 1)u(k - 1)
\]
\[
= A(k - 1)A(k - 2)x(k - 2) + A(k - 1)B(k - 2)u(k - 2) + B(k - 1)u(k - 1)
\]
\[
\vdots
\]
\[
= A(k - 1)A(k - 2)\ldots A(k_0)x(k_0) + \sum_{i=k_0}^{k-1} \Pi_{j=i+1}^{k-1} A(j)B(i)u(i)
\]

Thus, since \( x(k_0) = 0 \) for the zero-initial state response:

\[
s(k, k_0, 0_x, u) = \sum_{i=k_0}^{k-1} \Pi_{j=i+1}^{k-1} A(j)B(i)u(i)
\]
\[
\sum_{i=k_0}^{k-1} \Phi(k, i + 1)B(i)u(i)
\]