Lecture 1, Jan 6, 2026

Discrete-Time Systems

- Some physical phenomena is best captured with discrete-time dynamics rather than continuous time, e.g. the human saccadic system (where the eye moves around rapidly to sample different objects)
- (LTI) Discrete processes can be modelled in 3 ways:
  - State models: $x(k+1) = Ax(k) + Bu(k)$ where:
    $$y(k) = Cx(k) + Du(k)$$
    * k is a discrete sample time, which is not necessarily evenly spaced
    * $x(k)$ is the state
    * $u(k)$ is the input
    * $y(k)$ is the output
  - Difference equations: $y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) = b_0 u(k) + b_1 u(k-1) + \cdots + b_m u(k-m)$
    * Written in terms of the input, output, and delayed values (as opposed to always a single shift in k as in the state model)
    * The difference equation is to the state-space model the same way that an ODE is to a continuous time state-space model
  - Transfer functions: $\dfrac{Y(z)}{U(z)}$ where $Y(z)$, $U(z)$ are the Z-transformed versions of the output and input
    * Analogous to continuous time transfer functions, except using Z-transforms instead of Laplace transforms
- Example: modelling the saccade process: let the state $x(k) \in \mathbb{R}$ be the change in eye angle at the end of the saccade, the error $e(k) = r(k) - x(k)$ representing the difference between the desired change in eye angle and the actual eye angle
  - How does the brain ensure that the saccade is always correct, despite various disturbances to the system (e.g. glasses)?
  - Augment the error with a disturbance $d(k)$, such that $e(k) = r(k) + d(k) - x(k)$
  - Consider a simple plant model $x(k+1) = u(k)$, which is in the form of a state model
    * The TF is $zX(z) = U(z) \implies \dfrac{X(z)}{U(z)} = \dfrac{1}{z}$
      - For Z-transforms, a shift $x(k) \to x(k+1)$ becomes a multiply, $X(z) \to zX(z)$, analogous to how differentiation becomes multiplication for Laplace transforms
    * The difference equation is obtained by simply shifting back, $x(k) = u(k-1)$
      - In this case the difference between state model and difference equation is trivial, but in higher dimension systems the difference equation becomes higher order
  - Derive the error model, $e(k+1) = r(k+1) + d(k+1) - x(k+1)$
    $$= \bar{r} + \bar{d} - x(k+1)$$
    $$= \bar{r} + \bar{d} - u(k)$$
    * We assume $r(k) = r(k+1) = \bar{r}$ is a constant target, and $d(k) = d(k+1) = \bar{d}$ is a constant disturbance
    * To get the error model, we keep forward shifting the dynamics until the input $u(k)$ appears
    * We only needed to forward shift once, so this is a first-order model
  - Consider a controller $u(k) = u_s(k) + u_m(k)$ where u_m is the internal model, which handles the steady state behaviour, and u_s gives us good transient behaviour
  - Choose $u_s(k) = Ke(k)$ for some gain K, so $e(k+1) = -Ke(k) - u_m(k) + \bar{d} + \bar{r}$
    * Suppose $u_m(k) = 0$; is this stable?
      - For a system in canonical form $x(k+1) = Ax(k)$, the system is asymptotically stable if and only if the eigenvalues of A are inside the unit circle, so in this case the undriven system is stable if $|K| < 1$
      - To find the steady state (we know it exists in this case), set $e_{ss} = e(k) = e(k+1) \implies e_{ss} = \dfrac{1}{1+K}(\bar{d} + \bar{r})$
      - Notice that the higher the gain, the smaller our steady state error, but we can never

eliminate the steady state error if $u_m(k) = 0$ since we can only set K to be so high before the system becomes unstable

  - We add an internal model $\hat{w}(k+1) = \hat{w}(k) + Ge(k)$ and $u_i(k) = \phi\hat{w}(k)$, and we can show that the resulting closed-loop system is asymptotically stable

Note

The failure of the controller with $u_m(k) = 0$ is an example of the *Internal Model Principle*, which intuitively states that "any good regulator must create an internal model of the dynamic structure of the environment in the closed loop system," i.e. the controller must incorporate a model of the *exosystem*, which is capable of producing any disturbance frequencies that could enter the system.

Lecture 2, Jan 8, 2026

Modelling of Discrete-Time Systems

- LTI systems can be modelled in 3 ways:
  1. *Difference equations* (DEs) of the form $y(k)+a_1y(k-1)+\cdots+a_ny(k-n) = b_0u(k)+\cdots+b_mu(k-m)$
     - Note we assume that $n \geq m$, which is required for causality (otherwise inputs in the future can affect outputs in the past)
     - This is analogous to ODEs in continuous domain
  2. *Transfer functions* (TFs): $G(z) = \dfrac{Y(z)}{U(z)}$, assuming zero initial conditions
     - They are analogous to transfer functions in continuous time but using a Z-transform instead of the Laplace transform; we also have the ratio of the Z-transform of the output over the input
     - $\dfrac{Y(z)}{U(z)} = \dfrac{b_0 + b_1z^{-1} + \cdots + b_mz^{-m}}{1 + a_1z^{-1} + \cdots + a_nz^{-n}} = \dfrac{b_0z^n + b_1z^{n-1} + \cdots + b_mz^{n-m}}{z^n + a_1z^{n-1} + \cdots + a_{n-1}z + a_n} = \dfrac{N(z)}{D(z)}$
  3. State-space models: $x(k+1) = Ax(k) + Bu(k)$ (*canonical form*)
     $$y(k) = Cx(k) + Du(k)$$
     - $x(k) = \begin{bmatrix} x_1(k) \\ \cdots \\ x_n(k) \end{bmatrix} \in \mathbb{R}^n, u(k) = \begin{bmatrix} u_1(k) \\ \cdots \\ u_m(k) \end{bmatrix} \in \mathbb{R}^m, y(k) = \begin{bmatrix} y_1(k) \\ \cdots \\ y_p(k) \end{bmatrix} \in \mathbb{R}^p$

Definition

The *Z-transform* of a discrete time signal $x(k)$ is defined as

$$\mathcal{Z}\{x(k)\} = X(z) = \sum_{k=0}^{\infty} x(k)z^{-k}$$

for some $z \in \mathbb{C}$, provided that the sum is convergent.

- Note the forward and backward shift properties of the Z-transform: $x(k+m)$ transforms to $z^mX(z)$ (for positive or negative m), assuming zero initial conditions
  - Therefore Z-transforming the difference equation we get $Y(z) + a_1z^{-1}Y(z) + \cdots + a_nz^{-n}Y(z) = b_0U(z) + \cdots + b_mz^{-m}U(z)$, which we can rearrange to get the expression earlier
  - Just like how the Laplace transform converts differentiations to multiplication, the Z-transform converts time shifts to multiplication, turning a difference equation into algebra

Conversion Between State Space and Transfer Function Representations

- To convert from state space to transfer function, take the Z-transform of the state and output equations, and solve for $X(z)$ and $Y(z)$
  - $x(k+1) = Ax(k) + Bu(k) \implies zX(z) = AX(z) + BU(z) \implies X(z) = (zI - A)^{-1}BU(z)$

    * Note $(zI - A)^{-1}$ exists since its form means it can be written as a convergent power series, so it can always be inverted
  - $y(k) = Cx(k) + Du(k) \implies Y(z) = CX(z) + DU(z)$
    $$= C(zI - A)^{-1}BU(z) + DU(z)$$
    $$= (C(zI - A)^{-1}B + D)U(z)$$
  - This gives the transfer function $G(z) = \frac{Y(z)}{U(z)} = C(zI - A)^{-1}B + D$
  - Note that we normally only do this for SISO systems (taking the ratio of vectors would not make sense anyway) since transfer functions are generally very hard to work with for MIMO systems
- To convert from transfer function to state space, assume $G(z) = \frac{N(z)}{D(z)}$ is a rational proper transfer function, start by converting to a difference equation by inverting the Z-transform, define as many states as necessary (equal to the order of the difference equation) and organize into matrix form
  - Example: $G(z) = \frac{1}{z^2 + a_1 z + a_0}$
  - Convert to difference equation: $$\begin{aligned}&(z^2 + a_1 z + a_0)Y(z) = U(z)\\ \implies &y(k+2) + a_1 y(k+1) + a_2 y(k) = u(k)\end{aligned}$$
  - Define states: $x_1(k) = y(k), x_2(k) = y(k+1) \implies \begin{cases} x_1(k+1) = y(k+1) = x_2(k) \\ x_2(k+1) = y(k+2) = u(k) - a_1 x_2(k) - a_2 x_1(k) \end{cases}$
  - $x(k+1) = \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k)$
  - This is known as the *controllable canonical form* (analogous to the same in continuous domain)
- More generally when the numerator is not a constant, rewrite $Y(z) = \frac{N(z)}{D(z)}U(z) = N(z)V(z)$ where $V(z) = \frac{1}{D(z)}U(z)$, and $V(z)$ becomes the state, $Y(z) = N(z)V(z)$ becomes the output equation, and $V(z) = \frac{1}{D(z)}U(z)$ becomes the state equation
  - $G(z) = \frac{b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0} = \frac{N(z)}{D(z)}$
  - State equation: $$\begin{aligned}&V(z) = \frac{1}{z^3 + a_2 z^2 + a_1 z + a_0}\\ \implies &v(k+3) + a_2 v(k+2) + a_1 v(k+1) + a_0 v(k) = u(k)\end{aligned}$$
    * Define states $x(k) = \begin{bmatrix} v(k) \\ v(k+1) \\ v(k+2) \end{bmatrix} \implies x(k+1) = \begin{bmatrix} x_2(k) \\ x_3(k) \\ -a_2 x_3(k) - a_1 x_2(k) - a_0 x_1(k) + u \end{bmatrix}$
    * Matrix form: $x(k+1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(k)$
  - Output equation: $$\begin{aligned}&Y(z) = (b_1 z + b_0)V(z)\\ \implies &y(k) = b_1 v(k+1) + b_0 v(k)\\ &= b_1 x_2(k) + b_0 x_1(k)\\ &= \begin{bmatrix} b_0 & b_1 & 0 \end{bmatrix} x(k)\end{aligned}$$
- Note that the states obtained by converting from transfer function to state space are not unique
  - Some states are easier to work with than others, so we might want to do a coordinate transform $z(k) = Px(k)$ for some nonsingular P to obtain a new system $z(k+1) = PAP^{-1}z(k) + P^{-1}BU(k)$

Lecture 3, Jan 9, 2026

Time Response

- Given a state space model $x(k+1) = Ax(k) + Bu(k), y(k) = Cx(k) + Du(k)$, an initial condition $x(0)$ and input at all times $\{ u(k) \}_{k \geq 0}$, we want to obtain an explicit formula for $x(k), y(k)$, for $k \geq 0$
- As with continuous systems, since the system is LTI we can again break the total response into a superposition of the initial state response (nonzero initial conditions with zero input) and input response (zero initial conditions with nonzero input)
  - Initial state response: $x(k+1) = Ax(k) \implies x(k) = A^k x(0)$
    * This is an explicit formula because A^k can always be computed non-iteratively as we will see later
    * A^k is the discrete time analogue of e^{At} in continuous time
  - Input response:
$$\begin{aligned} & x(k+1) = Ax(k) + Bu(k), x(0) = 0 \\ \implies & x(k) = Bu(k-1) + ABu(k-2) + A^2Bu(k-3) + \cdots + A^{k-1}Bu(0) \\ & = \sum_{i=0}^{k-1} A^{k-1-i}Bu(i) \end{aligned}$$
    * This is the discrete time analogue of a convolution
- The total response is $x(k) = A^k x(0) + \sum_{i=0}^{k-1} A^{k-1-i}Bu(i)$

Lecture 4, Jan 13, 2026

Computing the Matrix Power

- Given $A \in \mathbb{R}^{n \times n}$, we want to find a closed-form expression for A^k so we can compute the time response
- Let $\sigma(A) = \{ \lambda_1, \dots, \lambda_n \}, \lambda_i \in \mathbb{C}$ denote the spectrum of A, i.e. the set of eigenvalues of A
- Assume that A is diagonalizable (recall that this is equivalent to A having n linearly independent eigenvectors), then we can compute A^k as $A^k = P\Lambda^k P^{-1}$ where P is the matrix of eigenvectors which diagonalizes A
  - In general for nondiagonalizable A this will be replaced with the Jordan form
- Example: $A = \begin{bmatrix} 0 & 1 \\ 3 & -2 \end{bmatrix}$
  - This has characteristic polynomial $s^2 + 2s - 3 = (s+3)(s-1)$ so $\sigma(A) = \{ 1, -3 \}$
  - Substitute λ in $\begin{bmatrix} 0 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = \lambda \begin{bmatrix} v_{11} \\ v_{22} \end{bmatrix}$ and solve for eigenvectors to get $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$
  - $P = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \implies A^k = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & (-3)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix}^{-1}$
- Note that A^k can also be computed using the Z-transform (similarly to how e^{At} can be computed with the Laplace transform)
  - Starting with $x(k+1) = Ax(k)$, which we know the solution to be $x(k) = A^k x(0)$
  - Apply Z-transforms to get
$$\begin{aligned} & zX(z) - zx(0) = AX(z) \\ \implies & X(z) = (zI - A)^{-1} zx(0) \\ \implies & x(k) = \mathcal{Z}^{-1} \left\{ (zI - A)^{-1} z \right\} x(0) \end{aligned}$$
    * Note we don't assume zero initial conditions here since we're not trying to derive a transfer function
  - By existence and uniqueness of solutions, we conclude that $A^k = \mathcal{Z}^{-1} \left\{ (zI - A)^{-1} z \right\} x(0)$
    * Note that this inverse Z-transform can be computed componentwise using the residue theorem (not covered in this course)

Poles and Eigenvalues

- The qualitative behaviour of solutions can be inferred by looking at either the eigenvalues of A or the poles of the transfer function
- What is the relationship between poles and eigenvalues of A?
- Consider the output equation $y(k) = Cx(k)$ with transfer function (SISO) $\frac{Y(z)}{U(z)} = C(zI - A)^{-1}B$ as we've derived last time
  - Expanding the inverse we get $C\left(\frac{\text{adj}(zI - A)}{\det(zI - A)}\right)B = \frac{N'(z)}{D'(z)}$
    * Note that the middle term is a scalar, so we can conclude that $D'(z)$ is the characteristic polynomial of A
  - To get an actual transfer function we need to cancel roots so that the numerator and denominator are coprime, to get $\frac{N(z)}{D(z)}$
  - This means that all the poles (roots of $D(z)$) are eigenvalues of A, but not the other way around due to pole-zero cancellations – some information is lost when we convert from state space to transfer function
    * A system that has a stable transfer function might not necessarily be stable in all its states

Lecture 5

Qualitative Behaviour of Solutions

- Consider a system $x(k+1) = Ax(k) + Bu(k)$ with response $x(k) = A^k x(0) + \sum_{i=0}^{k-1} A^{k-1-i}Bu(i)$; what can we conclude qualitatively about the solution without explicitly computing it?
- Let the system have output $y(k) = Cx(k)$ and transfer function $G(z) = C(zI - A)^{-1}B = k\frac{(z-q_1)(z-q_2)\cdots}{(z-p_1)(z-p_2)\cdots}$
- To get $Y(k)$ we use partial fractions as with Laplace transforms, yielding $\frac{c_1}{z-p_1} + \frac{c_2}{z-p_2} + \cdots$
- What is the behaviour of a typical pole?
- For real distinct poles, $\mathcal{Z}^{-1}\left\{\frac{z}{z-p}\right\} = p^k$
  - Clearly, if $p = 1$ then we get a constant response, or if $p = -1$ we get an alternating response that has constant magnitude
  - For $p \in (0, 1)$ we get a decaying envelope that does not change sign; with the decay faster the closer p is to zero
  - For $p = 0$ we get a constant zero (but we need to watch out since now we have z/z)
  - For $p \in (-1, 0)$ we get a decaying envelope with alternating signs, again with decay faster the closer p is to zero
  - In summary:
    * $p < 0 \to$ solution alternates signs
    * $|p| < 1 \to$ solution decays
    * $|p| > 1 \to$ solution blows up
    * $p = \pm 1 \to$ steady-state response
- For complex conjugate poles, $\mathcal{Z}^{-1}\left\{\frac{z^2}{(z - re^{j\omega})(z - re^{-j\omega})}\right\} = \frac{1}{\sin\omega} r^k \sin(k\omega + \phi)$
  - This contains a constant factor, an exponential envelope r^k and an oscillation at frequency ω
  - Larger values of r decay slower, until $r = 1$ which is a steady-state oscillation, then for $r > 1$ the solution blows up
  - The frequency of oscillation gets faster with increasing ω, as we increase the angle of the pole
    * However since the angles live on a circle, if we have $\omega > \pi$ (i.e. flipping beyond the negative

real line), the effect is the same as reducing the frequency
    * This is because $\omega = \pi$ represents the Nyquist frequency, and any higher frequency contents will be aliased into lower frequencies in the output, so we have a fundamental limit based on the sampling rate
  - In summary:
    * Poles on the same circle have the same exponential envelope
    * Poles with the same angle have the same oscillation frequency

Lecture 6, Jan 16, 2026

Stability of Discrete-Time Systems

- Consider a general open-loop nonlinear system $x(k+1) = f'(x(k), u(k))$ and its closed-loop system, $x(k+1) = f(k, x(k))$, where u no longer appears since we have designed some control law
  - Note that since we have k appearing explicitly in f, our systems are not time-invariant; this is necessary when we talk about adaptive control

Definition

A constant vector $\bar{x} \in \mathbb{R}^n$ is an *equilibrium* of the closed-loop system $x(k+1) = f(k, x(k))$ if $\bar{x} = f(k, \bar{x})$.

- Notice that whereas in continuous time the equilibria are points where f is zero, in discrete time equilibria are fixed points of f
- To make initial conditions explicit, denote $x(k) \equiv x(k; k_0, x_0)$ which means $x(k_0) = x_0$

Definition

Let $\bar{x} \in \mathbb{R}^n$ be an equilibrium of the system $x(k+1) = f(k, x(k))$, then

1. $\bar{x}$ is *stable* if $\forall k_0 \geq 0, \varepsilon > 0, \exists \delta(\varepsilon, k_0) > 0 \text{ s.t. } \|x_0 - \bar{x}\| < \delta \implies \|x(k; k_0, x_0) - \bar{x}\| < \varepsilon, \forall k \geq k_0$.
2. $\bar{x}$ is *asymptotically stable* if it's stable and $\exists \delta(k_0) > 0 \text{ s.t. } \|x_0 - \bar{x}\| < \delta \implies \lim_{k \to \infty} x(k; k_0, x_0) = \bar{x}$ (attractivity condition).
3. $\bar{x}$ is *uniformly asymptotically stable* if it's asymptotically stable and δ in the previous definitions are independent of k_0.
4. $\bar{x}$ is *globally asymptotically stable* if it is asymptotically stable (or *globally uniformly asymptotically stable* if it is also uniformly asymptotically stable), and $\delta(k_0)$ can be arbitrary large, i.e. all initial conditions converge to $\bar{x}$.

- The definition of stability is analogous to the continuous time definition; it requires that for any positive ε, we can find δ such that starting within δ of the equilibrium guarantees that the solution never goes outside ε of the equilibrium
- Similarly for asymptotic stability, like in the continuous case, we require that solutions near the equilibrium converge to the equilibrium
- Uniform asymptotic stability is important for reasons of robustness
  - Note that this is only an issue for time-dependent systems; for time-invariant systems we never have this k_0 dependence, but adaptive control is time-dependent
- In general GUAS is the best outcome

Lecture 7, Jan 20, 2026

Lyapunov's Method

- Note: assume without loss of generality that $\bar{x} = 0$ in the definitions below

Definition

A continuous function $V : \mathbb{R}^n \mapsto \mathbb{R}$ is *positive definite* at $x = 0$ if $V(0) = 0$ and $V(x) > 0$ for all $x \neq 0$.

V is *positive semidefinite* at $x = 0$ if $V(0) = 0$ and $V(x) \geq 0$ for all x, i.e. the function can be zero at nonzero locations.

- Note that these definitions can be adjusted accordingly if studying a nonzero $\bar{x}$ equilibrium
- We can similarly define negative definite and negative semidefinite using the same conditions but for $-V(x)$

Definition

A function $\kappa : [0, \infty) \mapsto [0, \infty)$ is a *class $\mathcal{K}_\infty$ function* if it has the properties:
1. $\kappa(0) = 0$
2. κ is stricty increasing on $[0, \infty)$
3. $\lim_{x \to \infty} \kappa(s) \to \infty$

Definition

A function $V : \mathbb{R}^n \mapsto \mathbb{R}$ is *radially unbounded* if there exists a function $\kappa : [0, \infty) \mapsto [0, \infty)$ of class $\mathcal{K}_\infty$ such that $V(k, x) \geq \kappa(\|x\|)$ for all $k \geq 0, x \in \mathbb{R}^n$.

- Intuitively, being radially unbounded means that if we go out along any ray from the origin, the function keeps increasing and never flattens out
  - A positive definite function that is not radially unbounded would flatten out along some direction, i.e. it would have a level set that goes out to infinity
  - This is important since Lyapunov's method works similarly to a gradient descent, so if the function becomes flat we can "get stuck"
- Note that a more restrictive definition would be $V(k, x) \geq c_1 \|x\|^2$, i.e. a quadratic bound, which works in the cases we will be examining but not in general
- Let ΔV denote the change in V along a particular solution of $x(k)$, called the *forward difference*
  - For time-invariant systems, we define $\Delta V = \Delta V(x)$ to be a function of only the state x
    * $\Delta V(x) = \Delta V(x(k)) = V(x(k + 1)) - V(x(k)) = V(f(x)) - V(x)$
  - For time-varying systems, we define $\Delta V = \Delta V(k, x)$ with a dependence on k as well
    * $\Delta V(k, x) = \Delta V(k, x(k)) = V(k + 1, x(k + 1)) - V(k, x(k)) = V(k + 1, f(k, x(k))) - V(k, x(k))$
  - This is the analogue of $\dot{V}$ in continuous time

Theorem

Main Theorem of Lyapunov: Suppose there exists a positive definite *Lyapunov function* $V : \mathbb{N}_0 \times \mathbb{R}^n \mapsto \mathbb{R}$ such that $V(k, 0) = 0, \forall k$, and let $\bar{x} = 0$ be an equilibrium, then:
1. If $\Delta V(k, x)$ is negative semi-definite in x, then $\bar{x} = 0$ is stable.
2. If $\Delta V(k, x)$ is negative definite in x, then $\bar{x} = 0$ is asymptotically stable. Moreover if V is radially unbounded, then $\bar{x} = 0$ is globally asymptotically stable.

- Notice that the Lyapunov function is a function of x and k for time-dependent systems, or a function of just x for time-invariant systems
- Example: Consider the system $x(k + 1) = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} x(k)$; prove that the equilibrium $\bar{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is GAS
  - Consider the Lyapunov function $V(x) = x_1^2 + x_2^2$
    * Selecting the appropriate Lyapunov function is the most important part; when in doubt, try a quadratic
  - Clearly $V(x)$ is positive definite at $\bar{x} = (0, 0)$

  - Also, $V(x)$ is radially unbounded (check level sets or use the fact that $V(x) \geq c_1\|x\|^2$
  - Compute the forward difference: $\Delta V(x) = V(x(k+1)) - V(x(k))$

$$
\begin{aligned}
&= x_1^2(k+1) + x_2^2(k+1) - x_1^2(k) - x_2^2(k) \\
&= \frac{1}{4}x_1^2(k) + \frac{1}{4}x_2^2(k) - x_1^2(k) - x_2^2(k) \\
&= -\frac{3}{4}(x_1^2(k) + x_2^2(k))
\end{aligned}
$$

  - $\Delta V(k, x(k))$ is negative definite, so by Lyapunov's theorem and the above conditions $\bar{x} = (0,0)$ is GAS
- Example: Consider the nonlinear system $x(k+1) = -x(k) + x^3(k)$; show that $\bar{x} = 0$ is stable (note that it is not GAS in this case)
  - Consider the Lyapunov function $V(x) = x^2$, which is clearly positive definite and radially unbounded
  - Forward difference: $\Delta V(x) = V(x(k+1)) - V(x(k))$

$$
\begin{aligned}
&= x^2(k+1) - x^2(k) \\
&= (-x(k) + x^3(k))^2 - x^2(k) \\
&= x^2(k) - 2x^4(k) + x^6(k) - x^2(k) \\
&= -(2 - x^2(k))x(k)^4
\end{aligned}
$$

  - Notice that if $|x| < \sqrt{2}$ we have a positive ΔV, but once it gets larger the sign flips, so this is only negative definite within this reduced domain
  - Therefore we conclude that if we restrict the domain to $\{\, x : |x| < \sqrt{2} \,\}$, then $\bar{x} = 0$ is asymptotically stable, but not GAS; outside of this domain, we lose stability

Lecture 8, Jan 22, 2026

Exponential Stability

Definition

An equilibrium $\bar{x} \in \mathbb{R}^n$ for a system $x(k+1) = f(k, x(k))$ is *exponentially stable* if

$$\exists c, \delta > 0, \lambda \in (0,1) \text{ s.t. } \|x(k_0) - \bar{x}\| < \delta \implies \|x(k) - \bar{x}\| \leq c\|x(k_0)\|\lambda^{k-k_0}, \forall k \geq k_0 \geq 0$$

Theorem

Consider an open set $\mathcal{D} \subset \mathbb{R}^n$ with $0 \in \mathcal{D}$; the equilibrium $\bar{x} = 0$ is exponentially stable if and only if there exists $V : \mathbb{N}_0 \times \mathcal{D} \mapsto \mathbb{R}$ and constants $c_1, c_2, c_3 > 0$ such that $c_1\|x\|^2 \leq V(k, x) \leq c_2\|x\|^2$ and $\Delta V(k, x) \leq -c_3\|x\|^2, \forall k \leq 0, x \in \mathcal{D}$.

If $\mathcal{D} = \mathbb{R}^n$, then the equilibrium $\bar{x} = 0$ is globally exponentially stable.

- Note that this is an extension to the Lyapunov theorem we covered last lecture; the conditions here are stricter than the definiteness and radially unboundedness conditions in the previous theorems

Stability of Linear Time-Invariant Systems

Theorem

For a linear time-invariant system $x(k+1) = Ax(k)$, the equilibrium $\bar{x} = 0$ is globally exponentially stable if and only if $|\lambda| < 1, \forall \lambda \in \sigma(A)$, i.e. all eigenvalues of A are inside the open unit disk in $\mathbb{C}$. Such a matrix A is called *Schur stable*.

The equilibrium $\bar{x} = 0$ is stable if and only if $|\lambda| \leq 1, \forall \lambda \in \sigma(A)$, and any eigenvalues $|\lambda| = 1$ have a Jordan block of size 1.

Theorem

For $x(k+1) = Ax(k)$, $\bar{x} = 0$ is globally exponentially stable if and only if for each symmetric positive definite $Q \in \mathbb{R}^{n \times n}$, there exists a unique symmetric positive definite $P \in \mathbb{R}^{n \times n}$ such that $A^T PA - P = -Q$ (the *Lyapunov equation*).

- The idea is to use a Lyapunov function of $V(x) = x^T Px$, which is positive definite if P is positive definite, and
$$\begin{aligned}\Delta V(x) &= V(Ax(k)) - V(x(k))\\ &= (Ax(k))^T P(Ax(k)) - x^T(k)Px(k)\\ &= x^T(k)A^T PAx(k) - x^T(k)Px(k)\\ &= x^T(k)(A^T PA - P)x(k)\\ &= -x^T(k)Qx(k)\end{aligned}$$
  - This gives us global asymptotic stability if Q is positive definite
  - For exponential stability, we make use of the fact that $\lambda_{\min}(P)\|x\|^2 \leq x^T Px \leq \lambda_{\max}(P)\|x\|^2$ where $\lambda_{\min}, \lambda_{\max}$ denote the minimum and maximum eigenvalues
    * Note since P is symmetric positive definite it has all real and positive eigenvalues
  - This also gives $\Delta V(x) = -x^T Qx \leq -\lambda_{\min}(Q)\|x\|^2$, so we have all the conditions for exponential stability

Lecture 9, Jan 23, 2026

Feedback Stabilization

- Consider the LTI system $x(k+1) = Ax(k) + Bu(k), x \in \mathbb{R}^n, u \in \mathbb{R}^m$; we want to find a simple feedback controller $u(k) = Kx(k)$, such that $|\lambda| < 1, \forall \lambda \in \sigma(A + BK)$, i.e. the resulting closed-loop system is exponentially stable
- To study this we look at the *reachability* problem: what are the states reachable in l steps, using any $\{u(0), \ldots, u(l-1)\}$ from $x(0)$?
  - $x(k) = A^k x(0) + \sum_{j=0}^{k-1} A^{k-1-j} Bu(j)$
  - For $k = n$ this can be expanded: $x(n) = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \begin{bmatrix} u(n-1) \\ \vdots \\ u(0) \end{bmatrix} + A^n x(0)$
  - We have the *controllability matrix* $Q_c = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n \times (nm)}$
    * Note this is identical to the controllability matrix in the continuous case
  - If we want $x(n) = 0$, then we just need to find the inputs so that $Q_c \begin{bmatrix} u(n-1) \\ \vdots \\ u(0) \end{bmatrix} = -A^n x(0)$
    * Therefore if $\operatorname{rank}(Q_c) = n$ then we can always find a set of control inputs to drive the system

to zero in n steps regardless of $x(0)$

    * We can prove this to also be a necessary condition, i.e. we can drive the system to zero in n steps if and only if $\text{rank}(Q_c) = n$

  - As with the continuous case, there is no need to look at $l > n$ since by Cayley-Hamilton, $\text{rank}(\begin{bmatrix} B & AB & \cdots & A^{l-1}B \end{bmatrix}) \leq \text{rank}(Q_c)$ regardless of l

Definition

The LTI system (A, B) is *controllable* if $\text{rank}(Q_c) = n$, where $Q_c = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$.

- In discrete time we also have a controllable canonical form which looks the same as the continuous-time controllable canonical form; a system is controllable if and only if it can be transformed into this form

Theorem

If (A, B) is controllable where $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}$, then for any desired symmetric (i.e. conjugate pairs) spectrum $\{ \lambda_{1d}, \dots, \lambda_{nd} \}$ where $\lambda_{id} \in \mathbb{C}$, there exists a state feedback $u(k) = Kx(k)$ with $K \in \mathbb{R}^{m \times n}$, such that $\sigma(A + BK) = \{ \lambda_{1d}, \dots, \lambda_{nd} \}$.

- Example: pole placement for $x(k+1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & a_2 & a_3 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(k)$
  - This is in controllable canonical form so we know it's controllable
  - Let the desired eigenvalues be $\{ 0, 0.5, 0.5 \}$
  - Let $K = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}$ and expand $A + BK = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 + k_1 & a_2 + k_2 & a_3 + k_3 \end{bmatrix}$
  - Characteristic polynomial: $\det(sI - (A + BK)) = s^3 - (a_3 + k_3)s^2 - (a_2 + k_2)s - (a_1 + k_1)$
  - Expand desired characteristic polynomial: $(s - 0)(s - 0.5)(s - 0.5) = s^3 - s^2 + \frac{1}{4}s$
  - Solve for the terms to get $k_1 = -a_1, k_2 = -\frac{1}{4} - a_2, k_3 = 1 - a_3$

Lecture 10, Jan 27, 2026

Deadbeat Control

- If (A, B) is controllable, it is possible to arbitrary assign $\sigma(A + BK)$; what if we assign the spectrum to be all zeros?
  - This means the characteristic polynomial of $A + BK$ is s^n; by Cayley-Hamilton, we then have $(A + BK)^n = 0$
  - Therefore by using a controller $u(k) = Kx(k)$ and setting a zero spectrum, the closed-loop system goes to 0 in exactly n steps, i.e. $x(n) = x(n+1) = \cdots = 0$
  - This is called *deadbeat control*, and is in theory the perfect controller in discrete time
    * This behaviour of finite-time convergence is not possible in continuous time with only linear feedback

Stabilizability

Definition

The LTI system $x(k+1) = Ax(k) + Bu(k)$ is *stabilizable* if there exists a linear feedback $K \in \mathbb{R}^{m \times p}$ such that $\sigma(A + BK)$ is in the open unit disk in $\mathbb{C}$, i.e. the closed-loop system is asymptotically stable.

- This definition is analogous to that of the continuous time case

- It is possible for a system to be stabilizable but not controllable
  - e.g. $x(k+1) = \begin{bmatrix} 0.1 & 0 \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k)$
  - We can tell that by inspection we have no control over the first state since the input is zero in that state and it is not affected by the second state
  - We can also check $Q_c = \begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$
  - However, we can pick $u(k) = k_2 x_2(k)$ such that $|1 + k_2| < 1$, which gives $A + BK = \begin{bmatrix} 0.1 & 0 \\ 0 & 1 + k_2 \end{bmatrix}$ which is Schur stable, therefore the system is stabilizable despite it not being controllable
  - In general for some uncontrollable systems, the uncontrollable states are already stable, so we can do a partial state feedback to control only the unstable ones (Kalman decomposition)

Definition

The eigenvalue λ is *controllable* for a system (A, B) if $\operatorname{rank}\left(\begin{bmatrix} A - \lambda I & B \end{bmatrix}\right) = n$.

Theorem

PBH Test: The system (A, B) is controllable if and only if every $\lambda \in \sigma(A)$ is controllable. If there are uncontrollable eigenvalues, they must have $|\lambda| < 1$ for the system to be stabilizable, i.e. (A, B) is stabilizable if and only if each $\lambda \in \sigma(A)$ where $|\lambda| \geq 1$ is controllable.

- Example: with the same system from before
  - $\sigma(A) = \{0.1, 1\}$
  - For $\lambda = 0.1$: $\operatorname{rank} \begin{bmatrix} A - 0.1I & B \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.9 & 1 \end{bmatrix} = 1 < 2$ so this eigenvalue is uncontrollable
  - For $\lambda = 1$: $\operatorname{rank} \begin{bmatrix} A - I & B \end{bmatrix} = \operatorname{rank} \begin{bmatrix} -0.9 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 2$ so this eigenvalue is controllable
  - By the PBH test we conclude that the system is uncontrollable but stabilizable

Lecture 11, Jan 29, 2026

Adaptive Control – Static Error Model

- Consider the linear regression problem: given *measurements* $y(k) \in \mathbb{R}$, *regressor* $w(k) \in \mathbb{R}^q$, and a linear model $y(k) = \psi^T w(k)$ where the *parameter vector* $\psi \in \mathbb{R}^q$ is unknown, we want to recover ψ
  - Note that our measurements and regressor are functions of time, so we can't do this instantly
- We want to build an estimate of the parameters, $\hat{\psi}(k) \in \mathbb{R}^q$ (note that this is a function of time)
  - Let the estimated output $\hat{y}(k) = \hat{\psi}^T(k) w(k)$ and prediction error $e(k) = \hat{y}(k) - y(k)$
  - Let the parameter estimation error $\tilde{\psi}(k) = \hat{\psi}(k) - \psi$
  - This lets us write $e(k) = \tilde{\psi}^T(k) w(k)$, which is known as the *static error model* of adaptive control
    * This is called "static" as the error depends instantaneously on the current values of the estimation error and regressor, instead of having a dynamics that depends on past values
- To update $\hat{\psi}(k)$, we use the *gradient law*: $\hat{\psi}(k+1) = \hat{\psi}(k) - \gamma(k) e(k) w(k)$ where $\gamma(k)$ is a time-dependent learning rate
  - Let the cost $J(\hat{\psi}) = \frac{1}{2}\|e\|^2$, where $e = \tilde{\psi}^T w = w^T(\hat{\psi} - \psi)$
  - The gradient of the cost is $\nabla J(\hat{\psi}) = \left(\frac{\partial J(\hat{\psi})}{\partial \hat{\psi}}\right)^T = \left(\frac{\partial J(\hat{\psi})}{\partial e} \frac{\partial e}{\partial \hat{\psi}}\right)^T = (e w^T)^T = e w$, giving the $\gamma(k) e(k) w(k)$ term
  - For the learning rate, we take $\gamma(k) = \frac{\bar{\gamma}}{1 + \|w(k)\|^2}$ where $\bar{\gamma}$ is some nominal learning rate

    * The term in the denominator normalizes by the regressor, so our step size stays bounded, which is necessary for stability
    * Note we will show later that convergence requires $\bar{\gamma} \in (0, 2)$
- However, this alone does not guarantee $\hat{\psi}(k) \to \psi$ (e.g. consider a constant $w(k) = 0$); intuitively, the regressor has to do something "interesting" – this is the *persistence of excitation* (PE) condition

Definition

A regressor $w(k) \in \mathbb{R}^q$ is *persistently exciting* (PE) if

$$\exists \beta_0 > 0, N \in \mathbb{N}, N > 0 \text{ s.t. } \beta_0 I \preccurlyeq W(k, N) = \frac{1}{N} \sum_{\tau=k}^{k+N-1} w(\tau) w^T(\tau), \forall k \in \mathbb{N}_0$$

where $A \preccurlyeq B$ denotes $x^T A x \leq x^T B x, \forall x \in \mathbb{R}^q$.

- Note that the key difference between the adaptive control approach versus the least-squares approach is that least squares is a batch approach, i.e. it processes all the data at once, whereas adaptive control processes the data as it comes; consider the least squares approach:
  - We collect $\{ y(k) \}_{k=0}^{N-1}, \{ w(k) \}_{k=0}^{N-1}$
  - Let $Y(0, N) = \begin{bmatrix} y(0) \\ \vdots \\ y(N-1) \end{bmatrix}$ and $X(0, N) = \begin{bmatrix} w^T(0) \\ \vdots \\ w^T(N-1) \end{bmatrix}$, so $Y(0, N) = X(0, N)\psi$
  - We solve for $\min\limits_{\hat{\psi} \in \mathbb{R}^q} \| X(0, N)\hat{\psi} - Y(0, N) \|^2$
    * This is only solvable if $X(0, N)$ has full column rank
    * It turns out that the full rank condition here is related to the persistence of excitation condition in adaptive control
  - If $X(0, N)$ has full column rank, then $\hat{\psi} = (X^T(0, N)X(0, N))^{-1} X^T(0, N) Y(0, N)$

Lecture 12, Jan 30, 2026

Model Reference Adaptive Control

- In the model reference adaptive control (MRAC) problem, we have a system $x(k+1) = Ax(k) + Bu(k)$ and a reference model $x_r(k+1) = A_r x_r(k) + B_r r(k)$, and our goal is to have $x(k)$ track $x_r(k)$
  - Assume A, A_r and B are known, but B_r is not, and we have state measurements for $x(k), x_r(k)$
  - $r(k) \in \mathbb{R}$ is an exogenous input, which can represent a tracking signal or a disturbance
    * Assume $r(k) = \psi^T w(k)$ where $w(k)$ is a known regressor and ψ is an unknown parameter vector
      - e.g. maybe we know the potential frequencies that can make up the disturbance but not their magnitudes
  - We also require that A_r is Schur stable
- We work with a controller $u(k) = Kx(k) + \hat{\psi}^T(k) w(k)$ for some $K \in \mathbb{R}^{1 \times n}$, where the first term stabilizes the transient behaviour, while the second term gives the desired steady-state behaviour
  - This results in the closed loop system $x(k+1) = (A + BK)x(k) + B\hat{\psi}^T(k) w(k)$
- Let the tracking error $\tilde{x}(k) = x(k) - x_r(k)$; what are its dynamics?
  - $$\begin{aligned} \tilde{x}(k+1) &= x(k+1) - x_r(k+1) \\ &= (A + BK)x(k) + B\hat{\psi}^T(k) w(k) - A_r x_r(k) - B_r \psi^T w(k) \\ &= (A + BK)(\tilde{x}(k) + x_r(k)) - A_r x_r(k) + (B\hat{\psi}^T(k) - B_r \psi^T) w(k) \\ &= (A + BK)\tilde{x}(k) + (A + BK - A_r) x_r(k) + (B\hat{\psi}^T(k) - B_r \psi^T) w(k) \end{aligned}$$
  - To have $\tilde{x}(k) \to 0$, we need to impose the *matching conditions*: $A_r = A + BK$ and $B_r = bB$ for some $b \in \mathbb{R}$

Adaptive Control – Dynamic Error Model

- The general *dynamic error model* is $x_e(k+1) = Ax_e(k) + B\tilde{\psi}^T(k)w(k)$, $e(k) = B^T P x_e(k)$, where $x_e(k)$ is a measurable error state, (A, B) is known, A is Schur stable, $w(k)$ is a known regressor, $\tilde{\psi}(k) = \hat{\psi}(k) - \psi$, and P is a symmetric positive definite matrix which solves the Lyapunov equation $A^T PA - P = -I$
  - Note that A being Schur stable guarantees that a solution exists for $A^T PA - P = -I$, as we have derived with Lyapunov analysis of LTI systems
  - The MRAC problem reduces to this
  - The use of $e(k)$ comes from Lyapunov theory, which we will see later
- Consider the transfer function for the system, using $e(k)$ as the output and $\tilde{\psi}^T w(k)$ as the input, we get $H(z) = B^T P(zI - A)^{-1}B$
  - Using the transfer function, $e(k) = H(z)\left[\tilde{\psi}^T(k)w(k)\right]$
    * Note that the notation here mixes time and Z-domain, with [ ] denoting a domain change
  - Notice the similarity between this and our static error model; can we redefine our regressor to incorporate $H(z)$?
- Let the *augmented error* $e_a(k) = e(k) - \hat{y}(k) - \hat{\psi}^T(k)w_a(k)$ where $w_a(k) = H(z)I\,[w(k)]$ is the *augmented regressor*, and $\hat{y}(k) = H(z)\left[\hat{\psi}^T(k)w(k)\right]$
  - Note $H(z)I\,[w(k)] = \begin{bmatrix} H(z) & & \\ & \ddots & \\ & & H(z) \end{bmatrix} \begin{bmatrix} w_1(k) \\ \vdots \\ w_q(k) \end{bmatrix} = \begin{bmatrix} H(z)\,[w_1(k)] \\ \vdots \\ H(z)\,[w_q(k)] \end{bmatrix}$
    * This can be interpreted as "filtering" the regressor signal by our plant