EEET 4071 Advanced Control (2019)
Lecture 10: Lyapunov stability

Dr. Yee Wei Law ⟨yeewei.law@unisa.edu.au⟩

Contents

1 Introduction

After almost two control courses, it should be clear by now that *stability* is the foremost requirement for any useful system. An unstable system will saturate or behave erratically when an input — no matter how small — is applied, and in the worst case it will fail, burn out or even disintegrate.

For LTI systems in continuous time, we know the system matrix must have only eigenvalues with a negative real part. For LTI systems in discrete time, we know that the system matrix must have only eigenvalues with a magnitude less than 1. This lecture deals with more than just LTI systems, and is meant to serve as an introduction to nonlinear systems through *Lyapunov stability theory*. The importance of stability theory cannot be emphasized enough, considering how most controller design techniques depend on definitions of stability and techniques of stabilization for their mathematical foundation.

1.1 Internal vs. external stability

In EEET 3046 Control Systems, we learnt that a SISO LTI system has two components in its output response: natural/free/zero-input response and forced/zero-state response. Because of this, we can qualify the stability of a SISO LTI system in terms of

- its *external stability*, i.e., the stability of its forced response, and
- its *internal stability*, i.e., the stability of its natural response.

In that course, we learnt the definitions of external stability and internal stability for SISO LTI systems (see Figure 1 for visual analogies), but how do they extend to multivariable LTI systems?

External stability is interpreted differently depending on whether system is a transfer function/-matrix or a state-space model.

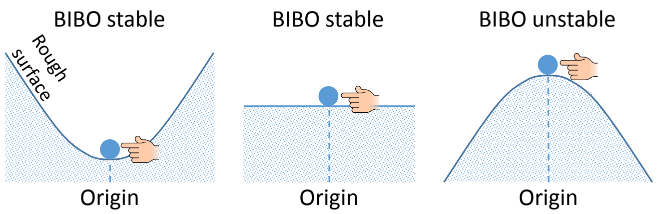


Figure 1: Analogies for external stability. Say the ball is at the origin, and we give the ball a poke of finite strength. Think of the distance between the ball and the origin as the forced response. If the ball eventually stops at a certain spot, the system is externally stable (BIBO stable). If the ball keeps moving away, the system is externally unstable (BIBO unstable).

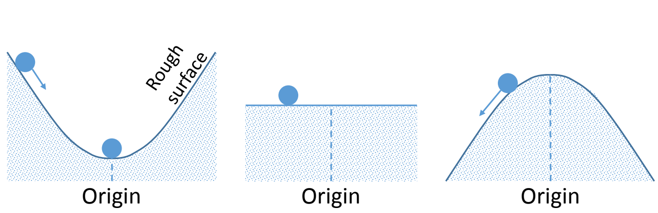


Figure 2: Analogies for internal stability. At $t = 0$, we place a ball at a random spot other than the origin, and see if the ball will settle at the origin (asymptotically stable), or settle at some point other than the origin (stable), or keep moving away from the origin (unstable). Think of the the distance between the ball and the origin as the natural response. Illustration inspired by [WL07, Figure 6.1].

- For a transfer function/matrix, external stability refers to the stability of the system's forced response. Since only the input and output of the system are considered, external stability is synonymous with *input-output stability*, and the more expressive term *bounded-input bounded-output stability* (BIBO stability). A system is BIBO stable if its zero-state response to any bounded input is also bounded [Kai80, Sect. 2.6.1]. Note that if an input is unbounded, the system's BIBO stability cannot be determined. The notion of BIBO stability carries over naturally from SISO systems to multivariable systems, with the only change being where we had "input" or "output" (singular), we now have "inputs" or "outputs" (plural).
- For a linear state-space model, the notion of BIBO stability is applicable.
- For a nonlinear state-space model, the notions of *input-to-state stability* [Son08] and *$\mathcal{L}$ stability* [Kha15, Ch. 6] apply, but the mathematical machinery required to analyze these notions of stability is beyond the undergraduate level.

Figure 3: Alexandr Lyapunov (note "Alexander" without an 'e'): Despite the impact of his work, it remained practically unknown outside Russia until after the World War II. For a great intellectual like him, he had a disproportionately tragic ending, having shot himself in the head on the same day his wife died from tuberculosis. He died three days after. Image from Wikipedia.

Internal stability is interpreted differently depending on whether the system is a transfer function (SISO only) or a state-space model (SISO or multivariable).

- For a transfer function (SISO only), internal stability has the classical meaning: the stability of the transfer function's natural response. A natural response is
  - *stable* if it approaches zero as time approaches infinity;
  - *marginally stable* if it neither decays nor grows but remains constant or oscillates as time approaches infinity;
  - *unstable* if it grows without bound with time [Nis15, Sect. 6.1].

  In this interpretation, internal stability implies BIBO stability but the converse is not true, and internal stability is simply equivalent to the case where all the poles of the system transfer function have a negative real part.

- For a state-space model (SISO or multivariable), the existence of states necessitates a different interpretation of internal stability that is due to Alexandr Lyapunov (see Figure 3). This interpretation is called *Lyapunov stability*, which is defined in Sect. 2. Since Lyapunov stability is defined for the most general state-space representation, it is applicable to LTI and non-LTI systems alike.
  - For LTI systems, Lyapunov stability implies BIBO stability, but the converse is only true for *minimal realizations* (systems that are both controllable and observable) [KA01, Sect. 2.2], [Kai80, p. 177], so Lyapunov stability is the stronger stability criterion.
  - For non-LTI systems and nonlinear systems in particular, Lyapunov stability does not imply external stability [Kha02, p. 175].

  For analyzing Lyapunov stability, *Lyapunov stability theory* is the gold standard. In this lecture, we will study this theory, and apply it to the stability analysis of nonlinear systems (see Sect. 3.2) as well as LTI systems (see Sect. 3.3). We note that for nonlinear systems with nonzero input, Lyapunov stability theory is essential but not enough. Henceforth, we will refer to a system with zero input as an *autonomous system*.

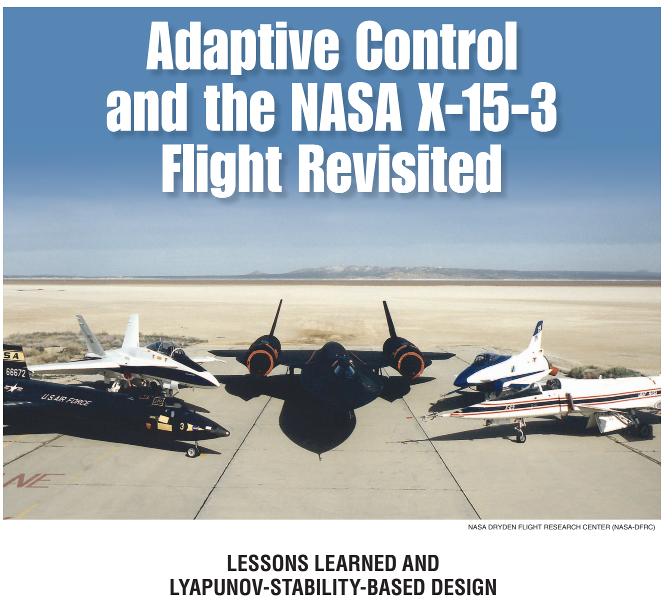


Figure 4: This is the cover page of an article [DAL10] that revisits the fatal accident of the X-15-3 experimental aircraft in 1967, by examining the role of Lyapunov-stability-based design of adaptive flight controllers in several "how and what if" scenarios. The usefulness of Lyapunov stability theory cannot be overstated. When in doubt, think of the legendary X-15 program.

No doubt the central theme of this lecture is Lyapunov stability theory, which is the foundation of nonlinear control, where we lose the ability to determine the stability of a system by means of the eigenvalues of the nonexistent system matrix. Even though this course does not cover nonlinear control, Lyapunov stability theory is so fundamentally important that it is the basis of all but the most basic control approaches. In the opening lecture, we read about NASA's X-15 program, specifically about how the ill-fated X-15-3 was equipped with an adaptive controller without a stability proof, and about how Lyapunov-based model reference adaptive control turned the multi-million dollar program around (see Figure 4). After all, "Lyapunov methods" is a suggested Tier 2 topic in Engineers Australia's "Mechatronic Engineering Foundational Technical Skills" document.

In terms of organization, Sect. 2 introduces the core stability definitions. Sect. 3 introduces Lyapunov's first and second methods, with the emphasis being the second method. Rounding up the lecture, Sect. 4 describes sample applications of Lyapunov theory to optimal control and adaptive control.

2 Lyapunov stability

So far we know stability is related to the "convergence" of some response to its equilibrium state. For LTI systems, we know convergence is exponential, but obviously this is not the case for *all* systems. The following example shows *how* a response converges to its equilibrium state matters.

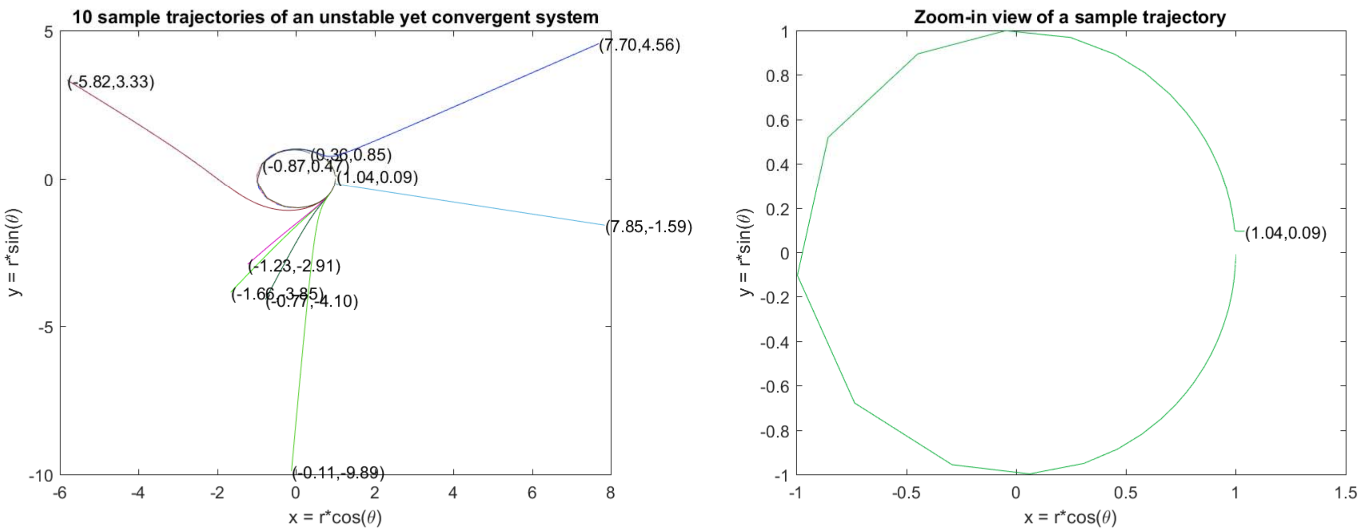


Figure 5: Phase portraits of system (1). The left figure shows 10 trajectories with a random initial state converge at the equilibrium point $(1, 0)$. The right figure shows an example of how a trajectory can start very close to the equilibrium point and yet make a sizeable excursion before reaching the equilibrium point. This is a classic example of how an equilibrium can be *unstable* and yet *convergent* (having all nontrivial trajectories converging to it).

Example 1

This example is adapted from [Ter09, Example 3.11]. Consider the nonlinear system

$$\dot{r} = r(1-r), \quad \dot{\theta} = \sin^2(\theta/2), \tag{1}$$

where (r, θ) are polar coordinates. The equilibrium states are given by

$$\begin{cases} \dot{r} = r(1-r) = 0 \implies r = 0 \text{ or } 1, \\ \dot{\theta} = \sin^2(\theta/2) = 0 \implies \theta = 2N\pi, \text{ where } N \in \mathbb{Z}. \end{cases}$$

So two of the equilibria are $(r, \theta) = (0, 0)$ and $(r, \theta) = (1, 0)$. Using Listing 1, we can generate the *phase portrait* (plot of one state against another state) in Figure 5, which features a conspicuous *limit cycle* (a closed trajectory which at least one other trajectory spirals into). Despite the limit cycle, all trajectories—except the trivial one that starts and stays at the origin—converge at $(r, \theta) = (1, 0)$.

Listing 1: MATLAB code for plotting Figure 5.

```
rng(1);
t0 = 0; tf = 400;
figure; subplot(1,2,1);
for run = 1:10
  if run < 10 % 10 starting points far from equilibrium
    r0 = 10*rand(1); theta0 = 2*pi*rand(1);
  else % 1 close to equilibrium
    r0 = 0.1*rand(1) + 1; theta0 = 0.1*rand(1);
  end
  [t,rtheta] = ode45(@unstable_yet_convergent, [t0,tf], [r0;theta0]);
  r = rtheta(:,1); theta = rtheta(:,2);
  x = r.*cos(theta); y = r.*sin(theta);
  p = plot(x,y); set(p, 'Color', rand(3,1));
  text(x(1), y(1), sprintf('(%.2f,%.2f)',x(1),y(1)));
  hold all
end
xlabel('x = r*cos(\theta)'); ylabel('y = r*sin(\theta)');
title('10 sample trajectories of an unstable yet convergent system');
hold off;

subplot(1,2,2);
p = plot(x,y); set(p, 'Color', rand(3,1));
text(x(1), y(1), sprintf('(%.2f,%.2f)',x(1),y(1)));
xlabel('x = r*cos(\theta)'); ylabel('y = r*sin(\theta)');
title('Zoom-in view of a sample trajectory');

function xdot = unstable_yet_convergent(t,x)
  % x is [r,theta]
  xdot = zeros(2,1);
  r = x(1); theta = x(2);
  xdot(1) = r*(1-r);
  xdot(2) = sin(theta/2)^2;
end
```

A common feature of the system trajectories is that they always make an excursion before converging to the equilibrium state, regardless of how close their initial state is to the equilibrium state. The implication is that even the *slightest* perturbation to the equilibrium state will cause the system to go through an excursion before settling back to its equilibrium state. If the system in question is responsible for cruise control, then since the slightest change in driving conditions or road conditions can set off a large excursion in speed, the system will do a horrible job maintaining its set-point speed. In this sense, which we call *the sense of Lyapunov*, we cannot practically consider the system stable despite the convergence of the system response.

To be exact, we cannot call the system at the specific equilibrium state (because a system can have multiple equilibrium states) stable in the sense of Lyapunov. Later, we will learn that the system at the specific equilibrium state is by definition not stable because no trajectories can be confined to an arbitrarily small neighborhood around the equilibrium state regardless of how close the initial state is to the equilibrium state.

Quiz 1

What is the primary difference between "system response" and "system trajectory"?

Example 1 introduces two key ideas:

- Internal stability is a property of a system with respect to an *equilibrium state*, because a system can have multiple isolated equilibrium states (equilibria), each with different stability properties. As a reminder, the equilibrium states of the autonomous system defined by the state equation

$$\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{2}$$

  are values of $\boldsymbol{x}$ that satisfy $\dot{\boldsymbol{x}} = f(\boldsymbol{x}) = \mathbf{0}$. By convention, instead of saying "system X is stable/unstable at equilibrium state $\boldsymbol{x}_e$", we say "equilibrium state $\boldsymbol{x}_e$ of system X is stable/unstable".

- Stability is different from convergence. *Stability in the sense of Lyapunov*, or *Lyapunov stability* in short, has to do with how a system state behaves around its equilibrium value.

The analysis of the Lyapunov stability of a system necessitates a formal mathematical framework, and this framework in turn requires the formal definitions of a few stability concepts. Below, we introduce these formal definitions:

Definition: Lyapunov stability

An equilibrium $\boldsymbol{x}_e$ of the system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ is

- **Lyapunov stable** or **stable in the sense of Lyapunov** or simply **stable** if for any $t_0 \geq 0$ and $\epsilon > 0$, there exists a $\delta > 0$ s.t.

$$\|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\| < \delta \implies \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| < \epsilon, \quad \forall t \geq t_0; \tag{3}$$

- **unstable** if it is not stable;

- **uniformly stable** if δ in Eq. (3) is independent of t_0;

- **convergent** or **attractive** if there exists a $\delta > 0$ s.t.

$$\|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\| < \delta \implies \lim_{t\to\infty} \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| = 0; \tag{4}$$

- **asymptotically stable** if it is stable and attractive;

- **globally asymptotically stable** if it is stable and $\lim_{t\to\infty} \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| = 0$ for all initial states.

The set of all $\boldsymbol{x}(t_0)$ s.t. $\boldsymbol{x}(t) \to \boldsymbol{x}_e$ as $t \to \infty$ is called the *domain of attraction / region of attraction*.

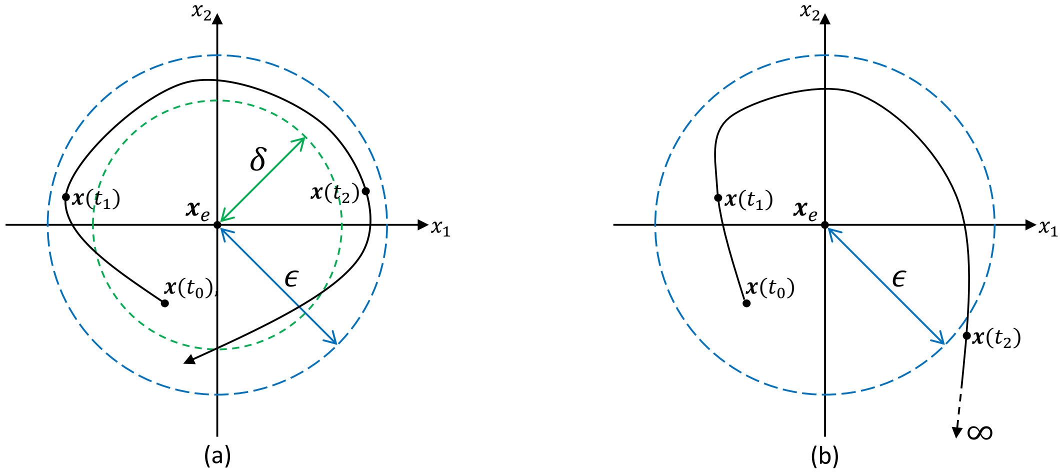


Figure 6: Examples of equilibria in a two-dimensional state space that are (a) Lyapunov stable, (b) unstable.

- It is a common practice to refer to the origin as an equilibrium state because if $\boldsymbol{x}_e$ is not the origin, we can always transfer it to the origin by introducing a new variable $\tilde{\boldsymbol{x}} \overset{\text{def}}{=} \boldsymbol{x} - \boldsymbol{x}_e$. Then, $\dot{\tilde{\boldsymbol{x}}} = \boldsymbol{0}$ at $\tilde{\boldsymbol{x}} = \boldsymbol{0}$. This is why stability is often defined based on the assumption $\boldsymbol{x}_e = \boldsymbol{0}$. However, in the definitions above, we have dispensed with this assumption.

- The concept of "Lyapunov stability" can be understood by imagining we are trying to keep every system trajectory within a neighborhood[1] of some finite radius ϵ centered at the equilibrium state (see Figure 6). After all, for stability's sake, we do not want any system trajectory to go to infinity. Clearly, the initial state $\boldsymbol{x}(t_0)$ must be chosen *at most* ϵ away from the equilibrium; let us define δ as the upper bound of this distance, i.e., $\delta \overset{\text{def}}{=} \inf \|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\|$. The mathematical definition of Lyapunov stability can then be understood through the subsequent challenge-response interpretation. The challenge is: *for any value of ϵ, can we find a value of δ such that as long as $\boldsymbol{x}(t_0)$ is less than δ away from $\boldsymbol{x}_e$, every possible system trajectory emanating from $\boldsymbol{x}(t_0)$ is contained within the ϵ neighborhood?* If the response is "yes", then we conclude $\boldsymbol{x}_e$ is *Lyapunov stable* or just *stable*; otherwise, $\boldsymbol{x}_e$ is not. Through the challenge and response, we can see that δ may depend on ϵ and t_0, thus it is a common practice to express δ as a function of ϵ and t_0, i.e., $\delta(\epsilon, t_0)$.

> **Quiz 2**
>
> In the challenge-response interpretation, ϵ can be of *any* value. However, is the challenger more concerned with *arbitrarily small* or *arbitrarily large* values? You can think of it this way: which case is associated with a smaller "search space" for δ?

- The set of all initial states at time t_0 for which $\boldsymbol{x}_e$ is Lyapunov stable is called the *domain/region of attraction*.

- In the challenge-response interpretation of Lyapunov stability, if δ exists and does not depend on t_0, then $\boldsymbol{x}_e$ is *uniformly stable*. In other words, if the region of attraction changes with t_0, then $\boldsymbol{x}_e$ is not uniformly stable. Uniform stability is a concern for time-varying systems since the property

[1] This neighborhood has the geometry of an "interval" in a one-dimensional state space, a "circle" in a two-dimensional state space, a "ball" in a three-dimensional state space, a "hypersphere" in a higher-dimensional state space.

can change with time, as the following example shows.

Example 2

This example is adapted from [LW13, Example 8.9]. Consider the first-order linear time-varying state equation:

$$\dot{x}(t) = 2t[3\sin(t) - 1]x(t), \tag{5}$$

subject to initial condition $x(t_0) = x_0$. The question is whether the equilibrium point at the origin is stable.

We can get the system response by solving the differential equation manually using the method of the integrating factor, as demonstrated in the EEET 5148 Control Systems M workshop. Alternatively, we can solve it using the MATLAB function `dsolve`. The system response turns out to be

$$x(t) = x_0 c(t_0) \exp\left\{6\sin(t) - 6t\cos(t) - t^2\right\},$$

where $c(t_0) = \exp\left\{-\left[6\sin(t_0) - 6t\cos(t_0) - t_0^2\right]\right\}$. Through constrained optimization (e.g., MATLAB function `fmincon`), we can determine the upper bound of $\exp\left\{6\sin(t) - \cdots\right\}$ to be around 22011, so

$$|x(t)| < 22011 c(t_0)|x_0| \qquad (\text{note } c(t_0) > 0).$$

In order for the origin to be stable, there must be a positive response to the challenge: for any given ϵ, can we find a δ such that

$$|x_0| < \delta \implies |x(t)| < \epsilon, \quad \forall t \geq t_0?$$

Suppose we define

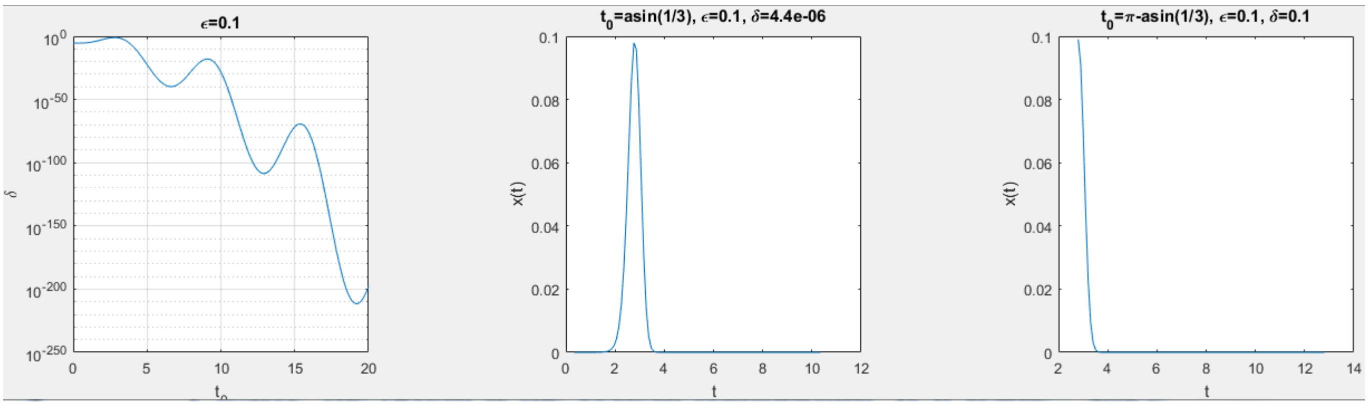
$$\delta \stackrel{\text{def}}{=} \frac{\epsilon}{22011 c(t_0)},$$

then for all $t \geq t_0$,

$$|x_0| < \delta \implies |x(t)| < 22011 c(t_0)|x_0| < \frac{\epsilon}{\delta}\delta = \epsilon.$$

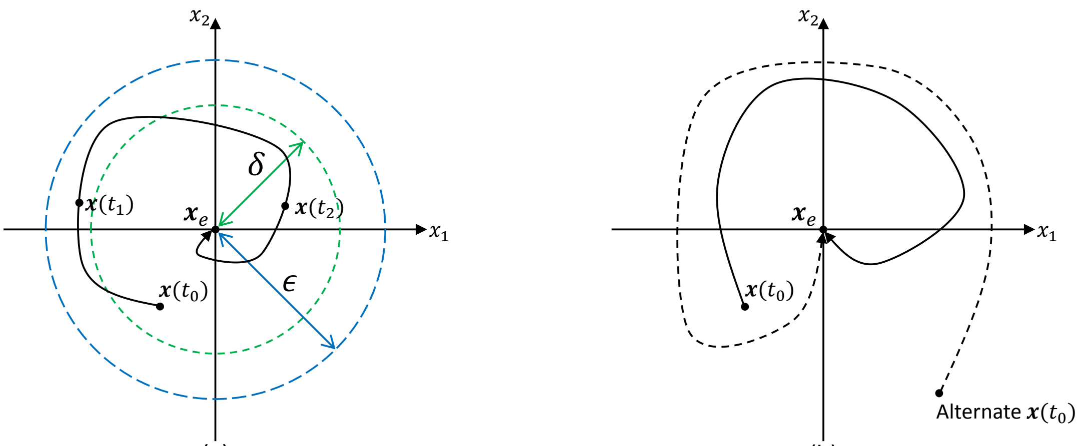
Clearly, δ depends on t_0, so the system is *not* uniformly stable. However, let us see how *not* uniformly stable the system is. By differentiating δ with respect to t_0, we can determine δ reaches its local minima and local maxima at $2N\pi + \arcsin(1/3)$ and $(2N+1)\pi - \arcsin(1/3)$ respectively, where $N \in \mathbb{Z}$. This is to say, for the same ϵ (how close we want the system trajectory to be near the equilibrium), x_0 must be significantly closer to the equilibrium for $t_0 = 2N\pi + \arcsin(1/3)$ than for $t_0 = (2N+1)\pi - \arcsin(1/3)$.

- If $\boldsymbol{x}_e$ is stable, then no matter how small ϵ is, we can always find a δ. However, regardless of how small δ is (i.e., how close $\boldsymbol{x}(t_0)$ is to $\boldsymbol{x}_e$), "stability" alone does not imply $\boldsymbol{x}(t)$ will ever converge to $\boldsymbol{x}_e$. When $\boldsymbol{x}(t)$ does converge to $\boldsymbol{x}_e$, we say $\boldsymbol{x}_e$ is *convergent* or *attractive*. A stable and convergent equilibrium is an *asymptotically stable* equilibrium (see Figure 8). It is easy to imagine a stable but not convergent equilibrium. As for an unstable yet convergent equilibrium, we have actually



(a) (b) (c)

Figure 7: The origin of system (5) is not uniformly stable. (a) shows δ generally decreases with t_0. (b) and (c) show x_0 needs to be significantly closer to the origin for $t_0 = \arcsin(1/3)$ than for $t_0 = \pi - \arcsin(1/3)$, given the same ϵ.



(a) (b)

Figure 8: Examples of equilibria in a two-dimensional state space that are (a) asymptotically stable, and (b) globally asymptotically stable.

seen one such example in Example 1. Asymptotic stability is clearly more desirable than Lyapunov stability.

- If the region of attraction of an asymptotically stable equilibrium is the entire state space, then the equilibrium is *globally asymptotically stable*. Any kind of stability quality that is globally true is clearly more desirable than otherwise.

To qualify how *fast* a system approaches its asymptotically stable equilibrium, we have the following definitions:

Definition: Exponential stability [Kha02, Definition 4.5]

An equilibrium $\boldsymbol{x}_e$ of the system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ is

- **exponentially stable** if for any $t_0 > 0$, there exist positive constants δ, k and λ s.t.

$$\|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\| < \delta \implies \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| \le k e^{-\lambda(t-t_0)} \|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\|, \quad \forall t \ge t_0 \tag{6}$$

(the maximum value of λ for which this holds is the *rate of convergence*);

- **globally exponentially stable** if for any $t_0 > 0$ and $\boldsymbol{x}(t_0)$, there exist positive constants k and λ s.t.
$$\|\boldsymbol{x}(t) - \boldsymbol{x}_e\| < ke^{-\lambda(t-t_0)}\|\boldsymbol{x}(t_0) - \boldsymbol{x}_e\|, \quad \forall t \geq t_0. \tag{7}$$

Exponential stability can be understood through the same challenge-response interpretation introduced earlier. To understand why we need the factor k in the definition, observe $k\exp(-\lambda(t-t_0))$ is another way of writing $\exp(-\lambda(t-t_0) + \text{some constant})$, which has a constant term in the exponent for generality. Global exponential stability is the ultimate goal, but can be difficult to achieve in many applications.

Example 3

Consider the nonlinear system

$$\dot{x}(t) = -(1 + \cos^2 x(t))x(t),$$

with equilibrium state at the origin. It is easily verifiable that the system response takes the form:

$$x(t) = x(0)\exp\left\{-\int_0^t \left[1 + \cos^2 x(\tau)\right] \mathrm{d}\tau\right\} = x(0)\exp(-t - \text{some nonnegative number}).$$

Therefore,
$$|x(t)| \leq |x(0)|\exp(-t),$$
i.e., the nonlinear system is exponentially stable.

It is important to note that the definitions provided so far are applicable to any dynamical system — linear or nonlinear, time-invariant or time-varying. In the next subsection, we will see a specialization of these definitions for LTI systems.

2.1 LTI systems

The linearity of LTI systems allows us to simplify the stability definitions earlier. Before we study these simplifications, let us get to know some useful facts about the equilibria of an LTI system. Consider the equilibria of the LTI system $\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x}$, which can be obtained by solving $\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x} = \mathbf{0}$. Basic linear algebra tells us:

- If $\mathbf{A}$ is nonsingular, the origin is the only equilibrium point. Such an equilibrium point is so-called *isolated*.

- If $\mathbf{A}$ is singular, $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ has infinitely many solutions for $\boldsymbol{x}$, hence the system has infinitely many equilibrium points including the origin.

Since the origin is always an equilibrium point of an LTI system, it is common to refer to the LTI system and the origin of the LTI system interchangeably in terms of stability. As such, "the system is stable/unstable" is really a concise way of saying "the origin of the system is stable/unstable" if the system is LTI.

Quiz 3

In the discussion above, why did we consider $\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x}$ rather than $\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u}$?

Now, the definitions of Lyapunov stability for LTI systems:

Definition: Lyapunov stability of LTI systems [WL07, Definition 6.2], [Hes09, Definition 8.1]

The origin of the system $\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t)$ is

- **Lyapunov stable** or **stable in the sense of Lyapunov** or simply **stable** if for any initial state $\boldsymbol{x}_0$, there exists a positive constant k s.t. $\|\boldsymbol{x}(t)\| \leq k\|\boldsymbol{x}_0\|, \forall t \geq 0$;
- **unstable** if it is not stable;
- **(globally) asymptotically stable** if for any initial state $\boldsymbol{x}_0$, $\lim_{t\to\infty} \|\boldsymbol{x}(t)\| = 0$;
- **(globally) exponentially stable** if for any initial state $\boldsymbol{x}_0$, there exist positive constants k and λ s.t. $\|\boldsymbol{x}(t)\| \leq ke^{-\lambda t}\|\boldsymbol{x}_0\|, \forall t \geq 0$.

The definition of Lyapunov stability for LTI systems is based on the observation:

$$\boldsymbol{x}(t) = e^{\mathbf{A}t}\boldsymbol{x}_0 \implies \|\boldsymbol{x}(t)\| = \|e^{\mathbf{A}t}\boldsymbol{x}_0\| \leq \|e^{\mathbf{A}t}\|\|\boldsymbol{x}_0\|. \tag{8}$$

If the norm of the matrix exponential does not grow unbounded, then $\|e^{\mathbf{A}t}\| \leq k$ for some positive constant k, and we have a Lyapunov stable system satisfying

$$\|\boldsymbol{x}(t)\| \leq k\|\boldsymbol{x}_0\|. \tag{9}$$

Example 4

Suppose $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix}$, then $\|e^{\mathbf{A}t}\|$ starts at 1, peaks at slightly below 1.2 before it decays to zero, so a good choice for k is $k = 1.2$.

In the definitions above, the parentheses surrounding "globally" mean if an LTI system is asymptotically/exponentially stable, it is globally so. Global exponential stability is the ultimate goal. Fortunately for LTI systems, asymptotic stability and exponential stability are equivalent [Ter09, Corollary 3.2], [WL07, p. 216], and achieving asymptotic stability requires just one simple condition: the system matrix of the system must be *Hurwitz* (if the system is continuous-time) or *convergent* (if the system is discrete-time). In either case, the system matrix is called a *stability matrix* [Hes09].

Definition: Hurwitz [Ter09, Defintion 3.2], [Son98, Definition C.5.2]

- A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is said to be **Hurwitz** if all of its eigenvalues have a negative real part.
- A matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is said to be **convergent** or **discrete-time Hurwitz** if all of its eigenvalues have a magnitude of less than 1.

On the relationship between Hurwtiz-ness and asymptotic stability, we have the following essential theorem for continuous-time systems:

Theorem 1: [WL07, Theorem 6.3]

An equilibrium state of $\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t)$ is

- *stable* iff all eigenvalues of $\mathbf{A}$ have a nonpositive real part, and the geometric multiplicity of any eigenvalue with zero real part equals the associated algebraic multiplicity;
- *globally asymptotically stable* iff $\mathbf{A}$ is Hurwitz.

For discrete-time systems, we have the following essential theorem:

Theorem 2: [Che99, p. 131]

An equilibrium state of $\dot{\boldsymbol{x}}[k+1] = \mathbf{A}\boldsymbol{x}[k]$ is

- *stable* iff all eigenvalues of $\mathbf{A}$ have a magnitude of less than or equal to 1, and the geometric multiplicity of any eigenvalue with unit magnitude equals the associated algebraic multiplicity;
- *globally asymptotically stable* iff $\mathbf{A}$ is convergent.

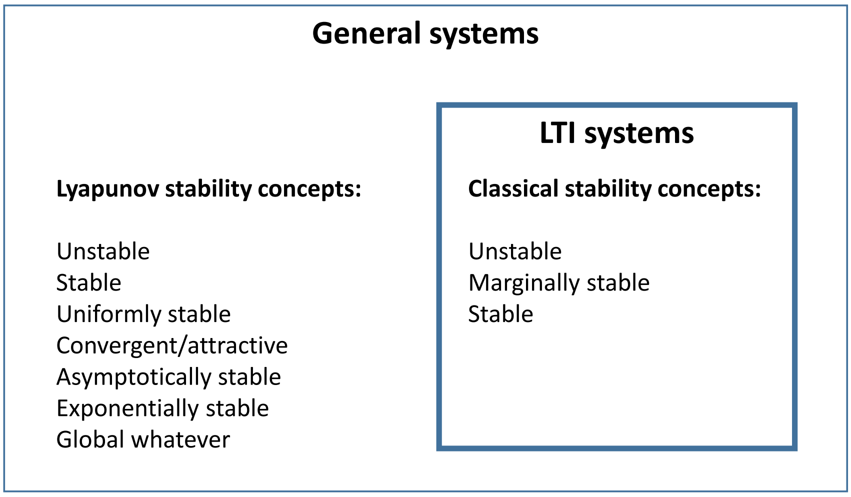


Figure 9: Keywords in the Lyapunov sense and keywords in the classical sense. A keyword in the classical sense can be mapped to multiple keywords in the Lyapunov sense.

Quiz 4

In Figure 9, can you map the keywords in the classical sense to the keywords in the Lyapunov sense? If you have difficulty mapping “marginally stable”, please read on to find out why.

3 Lyapunov stability analysis

Unlike linear systems, nonlinear systems do not have eigenvalues, and moreover closed-form formulas relating the trajectory to the initial state are typically not available for nonlinear systems. Lyapunov analysis is so far the only means by which we can determine the stability of these systems. The underlying theory originated in Lyapunov's doctoral thesis "The general problem of the stability of motion", which in turn was built on Poincaré's work in theoretical mechanics [KA01, p. 10]. Although published in 1892, Lyapunov's work remained practically unknown outside Russia until the Cold War and Space Race era, when there was a huge demand for techniques for analyzing and controlling nonlinear systems commonly found in aerospace applications. Lyapunov proposed two methods for determining stability [BW74]. We shall discuss them in turn.

3.1 Lyapunov's first/indirect method

Since we can determine the stability of a linear system by its eigenvalues, it is natural to ask whether we can determine the stability of a nonlinear system by the eigenvalues of its linear approximation. Lyapunov's first or indirect method [SL91, Sect. 3.3] has an answer to this very question, and essentially, it serves as the fundamental justification for using linear control techniques in practice.

Given a nonlinear system and a specific equilibrium point of the system, the method works by first linearizing the system about the specified equilibrium point, and calculating the eigenvalues of the linearized system. The following theorem is then used to determine the stability of the equilibrium point of the original nonlinear system:

Theorem 3: [SL91, Theorem 3.1]

Denote the equilibrium point of the original nonlinear system by $\boldsymbol{x}_e$.

- If the linearized system is strictly stable, i.e., all eigenvalues have a negative real part, then $\boldsymbol{x}_e$ is asymptotically stable.
- If the linearized system is unstable, i.e., at least one eigenvalue has a positive real part, then $\boldsymbol{x}_e$ is unstable.
- If the linearized system is marginally stable, i.e., all eigenvalues are in the closed left-half complex plane with at least one eigenvalue on the imaginary axis, then nothing can be concluded about the stability of $\boldsymbol{x}_e$.

Therefore, if a linearized system is marginally stable, the original nonlinear system can be stable, asymptotically stable or unstable.

Example 5

Consider the first-order nonlinear system:

$$\dot{x} = -x^5.$$

About the origin, the linearized system has an eigenvalue of 0, so Lyapunov's indirect method fails. In the next subsection, we shall see how the direct method can overcome this problem.

In addition to the limitation associated with marginally stable linearized systems, the indirect method provides no way of finding the region in which asymptotic stability applies, when the linearized system is asymptotically stability. These limitations of the indirect method explain the need for the direct method.

3.2 Lyapunov's second/direct method

To cover the case where the first method failed and also to prove the first method, Lyapunov developed the second or direct method, which provided the first general stability criteria applicable to both linear and nonlinear systems. Lyapunov observed that the stability of an equilibrium state could be based on a general class of energy-like functions. The intuition that real-world systems settle from a high-energy state into a low-energy one forms the basis of the *Lyapunov function*, which for a given system is a *user-defined* scalar function that is supposed to capture the energy of the system — this is why Lyapunov functions are typically quadratic. If a system is stable, then a Lyapunov function exists, and the system state only follows trajectories where the time derivative of the Lyapunov function is nonpositive, i.e., trajectories of nonincreasing energy. The existence of a Lyapunov function implies the stability of the system. Procedure-wise, the stability of a system can be established essentially by just investigating the signs of the Lyapunov function and its time derivative evaluated along the system trajectories. The primary advantage of the direct method is that it does not require the solutions to the differential equations describing the system.

A Lyapunov function is defined as follows:

Definition: Lyapunov function [NPWW02, Definition 2.2]

Let $V : \mathbb{R}^n \mapsto \mathbb{R}$ be a real-valued function, and $\mathcal{S} \subset \mathbb{R}^n$ be a compact (i.e., closed and bounded) region containing the origin in its interior. V is a Lyapunov function for the system $\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t))$ if it satisfies the following properties:

1. V is continuously differentiable, i.e., $\dfrac{\partial V}{\partial x_i}$ $(i = 1, \dots, n)$ exist and are continuous.

2. $V(\boldsymbol{x})$ is *positive definite* in $\mathcal{S}$, i.e.,

$$V(\boldsymbol{x}) \begin{cases} = 0 & \text{if } \boldsymbol{x} = \mathbf{0}; \\ > 0 & \text{if } \boldsymbol{x} \in \mathcal{S} \setminus \{\mathbf{0}\}. \end{cases} \tag{10}$$

3. $\dot{V}(\boldsymbol{x})$ along the system trajectories is *negative semidefinite*, i.e.,

$$\dot{V}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}} V \cdot \dot{\boldsymbol{x}} = \frac{\partial V}{\partial x_1}\dot{x}_1 + \cdots + \frac{\partial V}{\partial x_n}\dot{x}_n \begin{cases} = 0 & \text{if } \boldsymbol{x} = \mathbf{0}; \\ \leq 0 & \text{if } \boldsymbol{x} \in \mathcal{S} \setminus \{\mathbf{0}\}. \end{cases} \tag{11}$$

In Lecture 7, we came across the definition of positive/negative (semi)definiteness for *symmetric matrices*, whereas the definition above is for *functionals*. A functional is a real-valued function on a vector space. Caution: The definitions of positive semidefiniteness and positive definiteness in the textbook [Żak03, Definition 4.8], [Żak03, Definition 4.9] are inaccurate. For any general dynamical system, the following theorem give the conditions for its stability and asymptotic stability:

Theorem 4: [Kha02, Theorem 4.1], [SL91, Theorem 3.3]

- The origin of the system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, $\boldsymbol{f}(\boldsymbol{0}) = \boldsymbol{0}$, is *stable* if there exists a Lyapunov function V for the system.
- If furthermore $\dot{V}$ is *negative definite*, then the origin is *asymptotically stable*.
- If even furthermore V is *radially unbounded*, i.e., $V \to \infty$ as $\|\boldsymbol{x}\| \to \infty$, then the origin is *globally asymptotically stable*.

Essentially, if the origin of a system is stable, then all system trajectories follow the direction of nonincreasing V as illustrated in Figure 10.

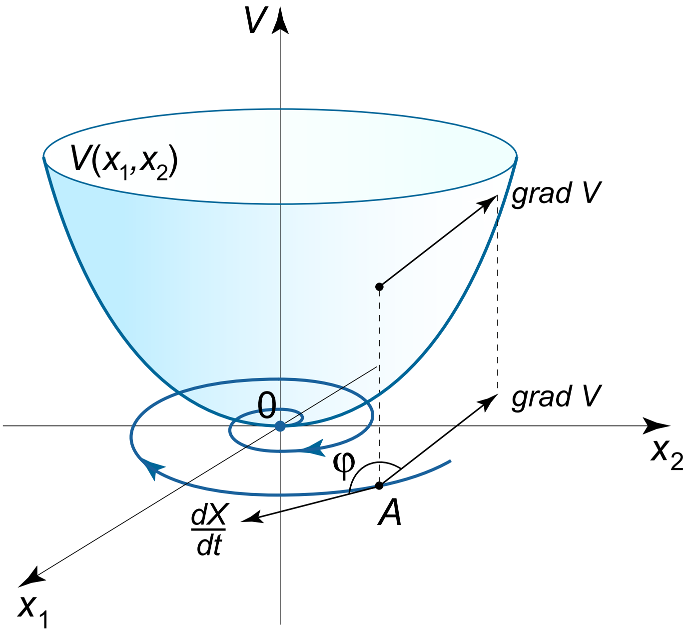


Figure 10: If the origin is stable, all system trajectories follow the direction of nonincreasing V, where V is the Lyapunov function. $\dot{V}$ is exactly the inner product of grad V and $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$ in the figure. Image source: `https://www.math24.net/wp-content/uploads/2016/09/lyapunov-function2.svg`

Lyapunov functions are user-defined. If we can find a Lyapunov function, we can prove a system is stable or asymptotically stable; otherwise, we cannot make any conclusion. Whether or not a Lyapunov function can be found depends on the problem domain, the designer's insight into the problem, his/her skills, and one may even say a bit of luck. However, since Lyapunov functions are energy-like functions, like electrical power and kinetic energy, they are often quadratic, such as the ones in the subsequent examples. This does not mean Lyapunov functions must be quadratic though.

Example 6

Revisiting Example 5. Let us consider a more general system:

$$\dot{x} = -c(x),$$

where $c(x)$ is defined such that $xc(x) > 0$ and $c(0) = 0$, e.g., $c(x) = x^5$. Define the Lyapunov function candidate as $V(x) = x^2$.

- Since $\dfrac{\mathrm{d}V}{\mathrm{d}x} = 2x$ exists and is continuous, V is continuously differentiable.
- V is clearly positive definite.

- $\dot{V} = \dfrac{\mathrm{d}V}{\mathrm{d}x}\,\dot{x} = -2xc(x)$ is clearly negative definite.
- Moreover, $V \to \infty$ as $|x| \to \infty$.

Therefore the origin is globally asymptotically stable. We saw in Example 5 how this result was not attainable using the indirect method.

Quiz 5

For each of the functions below, determine whether it satisfies the definition of $c()$ in Example 6:
- the saturation function,
- the tan function,
- the tanh function.

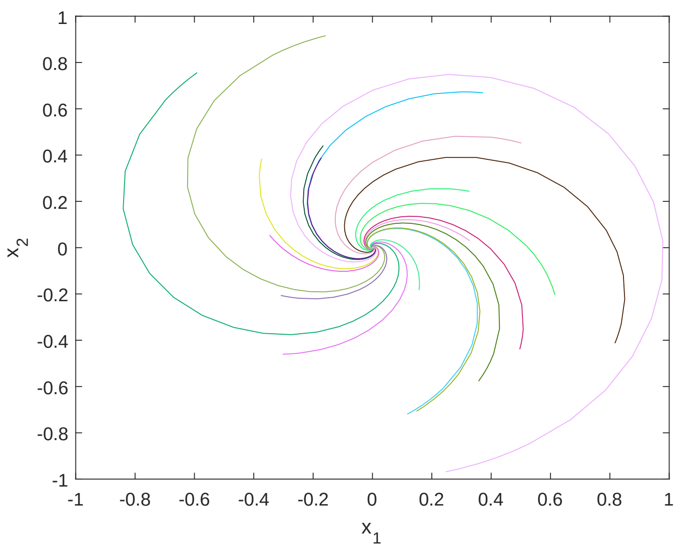


Figure 11: Plot of 20 sample trajectories of the system (12)–(13), with initial states satisfying $x_1^2 + x_2^2 < 1$. The origin is asymptotically stable.

Example 7

This example is adapted from [NPWW02, Example 2.3]. Determine the equilibrium point of the nonlinear system:

$$\dot{x}_1 = x_1(x_1^2 + x_2^2 - 1) - x_2, \tag{12}$$
$$\dot{x}_2 = x_1 + x_2(x_1^2 + x_2^2 - 1), \tag{13}$$

where $x_1, x_2 \in \mathbb{R}$. Show the equilibrium point is the origin, and the origin is asymptotically stable by showing $V(\boldsymbol{x}) = x_1^2 + x_2^2$ is a Lyapunov function and $\dot{V}(\boldsymbol{x})$ is negative definite.

Solution: To obtain the equilibrium point, we set $\dot{x}_1$ and $\dot{x}_2$ to zero, and get

$$(x_1^2 + x_2^2 - 1) = \frac{x_2}{x_1}, \quad (x_1^2 + x_2^2 - 1) = -\frac{x_1}{x_2}.$$

Hence $x_1^2 = -x_2^2$. This equation is only satisfied, since $x_1, x_2 \in \mathbb{R}$, by $x_1 = x_2 = 0$. The equilibrium point is therefore the origin.

Now, let us investigate the Lyapunov function candidate $V(\boldsymbol{x}) = x_1^2 + x_2^2$.

- Since the partial derivatives $\dfrac{\partial V}{\partial x_1} = 2x_1$ and $\dfrac{\partial V}{\partial x_2} = 2x_2$ exist and are continuous, V is continuously differentiable.
- Since
$$V = \begin{cases} 0, & \text{if } (x_1, x_2) = (0,0), \\ \text{sum of squares} > 0, & \text{if } (x_1, x_2) \neq (0,0); \end{cases}$$
V is positive definite.
- For $\dot{V}$ along the system trajectory, we have
$$\begin{aligned} \dot{V}(\boldsymbol{x}) &= \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2 = 2x_1[x_1(x_1^2+x_2^2-1) - x_2] + 2x_2[x_1 + x_2(x_1^2+x_2^2-1)] \\ &= 2(x_1^2+x_2^2)(x_1^2+x_2^2-1) \\ \implies \dot{V}(\boldsymbol{x}) &= \begin{cases} 0, & \text{if } x_1^2+x_2^2 = 0; \\ -\text{ve}, & \text{if } 0 < x_1^2+x_2^2 < 1; \\ 0, & \text{if } x_1^2+x_2^2 = 1; \\ +\text{ve}, & \text{if } 1 < x_1^2+x_2^2. \end{cases} \end{aligned}$$

Therefore within the region $x_1^2+x_2^2 < 1$, $V(\boldsymbol{x})$ is a Lyapunov function with $\dot{V}(\boldsymbol{x})$ being negative definite. Within this region, the origin is asymptotically stable, as confirmed by the plot in Figure 11.

Example 8

This example is adapted from [FPEN15, EXAMPLE 9.16]. Consider the nonlinear system
$$\ddot{x}_1 = -\dot{x}_1 + c(r - x_1),$$
where x_1 is the state variable, r is the *constant* input set-point, and $c()$ is a *continuous* nonlinear function with the property: $ec(e) > 0$ and $c(0) = 0$. The system is second-order, and by defining the state vector as $\boldsymbol{x} = [\dot{e}\ e]^\top$, where $e = r - x_1$, we can rewrite the system as the nonlinear state equation:
$$\dot{\boldsymbol{x}} = \begin{bmatrix} \ddot{e} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} -\dot{e} - c(e) \\ \dot{e} \end{bmatrix}. \tag{14}$$
Furthermore, the system is a servomechanism, because in steady state ($\ddot{e} = 0$, $\dot{e} = 0$),
$$\ddot{e} = -\dot{e} - c(e) \implies c(e) = c(r - x_1) = 0 \implies x_1 = r.$$
Suppose we define the Lyapunov function candidate as $V(\boldsymbol{x}) = \frac{1}{2}\dot{e}^2 + \int_0^e c(\xi)\,\mathrm{d}\xi$. Show that the equilibrium point at the origin is Lyapunov stable.

Solution:

- Since the partial derivatives $\dfrac{\partial V}{\partial \dot{e}} = \dot{e}$ and $\dfrac{\partial V}{\partial e} = c(e)$ (due to Leibniz's rule) exist and are continuous, V is continuously differentiable.

- Since
$$V = \begin{cases} 0, & \text{if } (\dot{e}, e) = (0,0), \\ \text{square } + \text{ positive} > 0, & \text{if } (\dot{e}, e) \neq (0,0); \end{cases}$$
  V is positive definite.

- For $\dot{V}$ along the system trajectory, we have
$$\dot{V} = \frac{\partial V}{\partial \dot{e}}\ddot{e} + \frac{\partial V}{\partial e}\dot{e} = \dot{e}(-\dot{e} - c(e)) + c(e)\dot{e} = -\dot{e}^2.$$
  Note that $\dot{V}$ along the system trajectory only depends on $\dot{e}$ and not e. More specifically, $\dot{V}$ is zero at points $(0, e)$, where e is nonzero, so $\dot{V}$ can be zero outside the origin. In other words, $\dot{V}$ is only negative semidefinite, and not negative definite.

Therefore, $V(\boldsymbol{x})$ is a Lyapunov function with a negative semidefinite $\dot{V}(\boldsymbol{x})$, and thus the origin is Lyapunov stable.

Quiz 6

In Example 8, explain why the state vector is NOT defined as $\boldsymbol{x} = [\dot{x}_1 \ x_1]^\top$ instead.

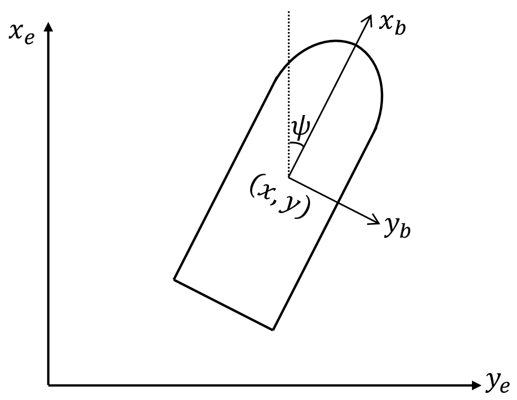


Figure 12: The position parameters of a ship: (x, y, ψ). Notice the earth axes and body axes follow the north-east-down and front-right-down conventions respectively.

Example 9

Examples of Lyapunov functions for physical systems abound in [DBDN13]. In [DBDN13, Sect. 2.3], we can see a case study on the position control of a ship. The control objective is to track the desired position expressed in terms of (x, y, ψ) as shown in Figure 12. Instead of discussing the system model at length, let us just say the following quadratic form has been found to be a Lyapunov function:
$$V = \frac{1}{2}\left(\boldsymbol{r}^\top \mathbf{M}^* \boldsymbol{r} + \boldsymbol{e}^\top \mathbf{K}_p \boldsymbol{e} + \tilde{\boldsymbol{\phi}}^\top \boldsymbol{\Gamma}^{-1} \tilde{\boldsymbol{\phi}}\right),$$
where

- $\boldsymbol{e}$ is the control error;
- $\boldsymbol{r} \stackrel{\text{def}}{=} \dot{\boldsymbol{e}} + \boldsymbol{\alpha}\boldsymbol{e}$ is the filtered tracking error, where $\boldsymbol{\alpha}$ is some diagonal gain matrix;
- $\mathbf{M}^* = \mathbf{R}\mathbf{M}\mathbf{R}^\top$ is the transformed mass-inertia matrix, where $\mathbf{R}$ is the rotation matrix between the earth and body axes, and $\mathbf{M}$ is the mass-inertia matrix;
- $\mathbf{K}_p$ and $\boldsymbol{\Gamma}$ are control gain matrices;
- $\tilde{\boldsymbol{\phi}}$ is the system parameter estimation error.

In fact, a Lyapunov function typically contains all the squared tracking errors and squared estimation errors in the system, since in equilibrium, these tracking errors and estimation errors should all be zero.

3.3 Lyapunov theorem

When applied to LTI systems, Lyapunov's theory gives us the Lyapunov theorem. For LTI systems, the *quadratic form* is especially useful as a Lyapunov function candidate:

$$V(\boldsymbol{x}) \stackrel{\text{def}}{=} \boldsymbol{x}^\top \mathbf{P} \boldsymbol{x}, \tag{15}$$

where $\mathbf{P}$ is a symmetric positive definite matrix. We shall see what other conditions $\mathbf{P}$ must satisfy in order for V to be a Lyapunov function.

- For the continuous-time LTI system $\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t)$, differentiating V gives us

$$\dot{V}(\boldsymbol{x}(t)) = \dot{\boldsymbol{x}}^\top(t)\mathbf{P}\boldsymbol{x}(t) + \boldsymbol{x}^\top(t)\mathbf{P}\dot{\boldsymbol{x}}(t) = \boldsymbol{x}^\top\mathbf{A}^\top\mathbf{P}\boldsymbol{x} + \boldsymbol{x}^\top\mathbf{P}\mathbf{A}\boldsymbol{x} = \boldsymbol{x}^\top(\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A})\boldsymbol{x}. \tag{16}$$

- For the discrete-time LTI systems $\boldsymbol{x}[k+1] = \mathbf{A}\boldsymbol{x}[k]$, differencing V gives us

$$\begin{aligned}\Delta V(\boldsymbol{x}[k]) = V(\boldsymbol{x}[k+1]) - V(\boldsymbol{x}[k]) &= \boldsymbol{x}^\top[k+1]\mathbf{P}\boldsymbol{x}[k+1] - \boldsymbol{x}^\top[k]\mathbf{P}\boldsymbol{x}[k] \\ &= \boldsymbol{x}^\top[k]\mathbf{A}^\top\mathbf{P}\mathbf{A}\boldsymbol{x}[k] - \boldsymbol{x}^\top[k]\mathbf{P}\boldsymbol{x}[k] = \boldsymbol{x}^\top[k](\mathbf{A}^\top\mathbf{P}\mathbf{A} - \mathbf{P})\boldsymbol{x}[k].\end{aligned} \tag{17}$$

Recall for an asymptotically stable equilibrium at the origin, we must have a decreasing V along the system trajectories converging at the origin, i.e., Eqs. (16)–(17) tell us

- if the system is continuous-time, $\dot{V}$ and hence $\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A}$ must be negative definite;
- if the system is discrete-time, ΔV and hence $\mathbf{A}^\top\mathbf{P}\mathbf{A} - \mathbf{P}$ must be negative definite.

The Lyapunov theorem confirms the existence of such a matrix $\mathbf{P}$.

Theorem 5: Lyapunov theorem

For *any* symmetric positive definite matrix $\mathbf{Q}$, the **Lyapunov equation**

$$\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q} \qquad \text{(continuous-time)} \tag{18}$$

$$\mathbf{A}^\top\mathbf{P}\mathbf{A} - \mathbf{P} = -\mathbf{Q} \qquad \text{(discrete-time)} \tag{19}$$

has a *unique* symmetric positive definite solution $\mathbf{P}$ iff $\mathbf{A}$ is a stability matrix.

Figure 13: An easy way to remember the continuous-time Lyapunov equation is to recall the immensely successful Swedish pop sensation called "ABBA" in the 1970s. There is some similarity between "$\mathbf{A}^\top\mathbf{P}+\mathbf{P}\mathbf{A}$" and "ABBA", right? Image from Wikipedia.

Proof: The following proof for the continuous-time case is from [WL07, Theorem 6.4].

Necessity ($\mathbf{A}$ is Hurwitz, hence $\mathbf{P}$ is ...): Suppose all eigenvalues of $\mathbf{A}$ are Hurwitz. Define a symmetric matrix

$$\mathbf{P} \overset{\text{def}}{=} \int_0^\infty e^{\mathbf{A}^\top t}\mathbf{Q}e^{\mathbf{A}t}\,\mathrm{d}t.$$

Since $\mathbf{Q}$ is positive definite, $\mathbf{Q}$ can be expressed as $\mathbf{R}^\top\mathbf{R}$, where $\mathbf{R}$ is nonsingular. Therefore,

$$\begin{aligned}\boldsymbol{x}^\top\mathbf{P}\boldsymbol{x} &= \int_0^\infty \left(e^{\mathbf{A}t}\boldsymbol{x}\right)^\top \mathbf{Q}\left(e^{\mathbf{A}t}\boldsymbol{x}\right)\mathrm{d}t = \int_0^\infty \left(e^{\mathbf{A}t}\boldsymbol{x}\right)^\top \mathbf{R}^\top\mathbf{R}\left(e^{\mathbf{A}t}\boldsymbol{x}\right)\mathrm{d}t \\ &= \int_0^\infty \|\mathbf{R}e^{\mathbf{A}t}\boldsymbol{x}\|^2\,\mathrm{d}t \geq 0.\end{aligned}$$

Since $\mathbf{R}$ and $e^{\mathbf{A}t}$ are nonsingular, $\mathbf{R}e^{\mathbf{A}t}\boldsymbol{x}$ is only zero when $\boldsymbol{x} = \mathbf{0}$. Therefore, $\mathbf{P}$ is positive definite. $\mathbf{P}$ also satisfies the Lyapunov equation because

$$\begin{aligned}\mathbf{A}^\top\mathbf{P}+\mathbf{P}\mathbf{A} &= \int_0^\infty \left\{\mathbf{A}^\top e^{\mathbf{A}^\top t}\mathbf{Q}e^{\mathbf{A}t} + e^{\mathbf{A}^\top t}\mathbf{Q}e^{\mathbf{A}t}\mathbf{A}\right\}\mathrm{d}t = \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\left\{e^{\mathbf{A}^\top t}\mathbf{Q}e^{\mathbf{A}t}\right\}\mathrm{d}t \qquad (\because \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A}) \\ &= e^{\mathbf{A}^\top t}\mathbf{Q}e^{\mathbf{A}t}\Big|_0^\infty = \mathbf{0} - \mathbf{Q} = -\mathbf{Q} \qquad (\because \mathbf{A} \text{ is Hurwitz}).\end{aligned}$$

$\mathbf{P}$ is the unique solution because if $\tilde{\mathbf{P}}$ is another solution,

$$\begin{aligned}\mathbf{A}^\top(\mathbf{P}-\tilde{\mathbf{P}}) + (\mathbf{P}-\tilde{\mathbf{P}})\mathbf{A} = \mathbf{0} &\implies e^{\mathbf{A}^\top t}\left\{\mathbf{A}^\top(\mathbf{P}-\tilde{\mathbf{P}}) + (\mathbf{P}-\tilde{\mathbf{P}})\mathbf{A}\right\}e^{\mathbf{A}t} = \mathbf{0} \\ &\implies \frac{\mathrm{d}}{\mathrm{d}t}\left\{e^{\mathbf{A}^\top t}(\mathbf{P}-\tilde{\mathbf{P}})e^{\mathbf{A}t}\right\} = \mathbf{0} \\ &\implies \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\left\{e^{\mathbf{A}^\top t}(\mathbf{P}-\tilde{\mathbf{P}})e^{\mathbf{A}t}\right\}\mathrm{d}t = -(\mathbf{P}-\tilde{\mathbf{P}}) = \mathbf{0}.\end{aligned}$$

Sufficiency ($\mathbf{P}$ is ..., hence $\mathbf{A}$ is Hurwitz): Suppose for a given positive definite $\mathbf{Q}$, there is a unique positive definite solution $\mathbf{P}$. If we let λ be an eigenvalue of $\mathbf{A}$ and $\boldsymbol{v}$ be the corresponding eigenvector, then

$$\begin{aligned}-\boldsymbol{v}^*\mathbf{Q}\boldsymbol{v} &= \boldsymbol{v}^*(\mathbf{A}^\top\mathbf{P}+\mathbf{P}\mathbf{A})\boldsymbol{v} = \boldsymbol{v}^*\mathbf{A}^*\mathbf{P}\boldsymbol{v} + \boldsymbol{v}^*\mathbf{P}\mathbf{A}\boldsymbol{v} = (\lambda\boldsymbol{v})^*\mathbf{P}\boldsymbol{v} + \boldsymbol{v}^*\mathbf{P}(\lambda\boldsymbol{v}) = (\bar{\lambda}+\lambda)\boldsymbol{v}^*\mathbf{P}\boldsymbol{v} \\ &= 2\Re(\lambda)\boldsymbol{v}^*\mathbf{P}\boldsymbol{v}.\end{aligned}$$

Since $\mathbf{Q}$ and $\mathbf{P}$ are by definition positive definite, $\Re(\lambda)$ must be negative, for every λ of $\mathbf{A}$. Hence $\mathbf{A}$ is Hurwitz. □

Important points about the Lyapunov theorem:

- For LTI systems, the Lyapunov theorem provides a necessary and sufficient condition for stability.

- The Lyapunov theorem is equivalent to saying that given a system matrix $\mathbf{A}$, provided we can find positive definite matrices $\mathbf{P}$ and $\mathbf{Q}$ that satisfy the Lyapunov equation associated with $\mathbf{A}$, then we can show $\mathbf{A}$ is a stability matrix [Ter09, Theorem 3.7].

- The Lyapunov theorem implies that given an asymptotically stable LTI system, we can always find some positive definite $\mathbf{P}$ s.t. $V(\boldsymbol{x}) \stackrel{\text{def}}{=} \boldsymbol{x}^\top \mathbf{P} \boldsymbol{x}$ is a Lyapunov function for the system [Son98, Theorem 18]. This fact will become useful later for solving the LQR problem.

- Continuous-time and discrete-time Lyapunov equations can be solved using the MATLAB functions `lyap` and `dlyap` respectively, but beware: where we have $\mathbf{A}$ or $\mathbf{A}^\top$ in Eqs. (18)–(19) the MATLAB functions use its transpose instead.

Example 10

This example is adapted from [KA01, Problem 3.6.9]. Consider the continuous-time system:

$$\begin{aligned}\dot{\boldsymbol{x}} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u, \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}.\end{aligned}$$

Apply the Lyapunov theorem to determine whether the system is asymptotically stable. Note: Of course we can tell the system is unstable because the system matrix has eigenvalues -1 and 1, but the purpose of this example is to show how the same conclusion can be deduced from the Lyapunov theorem.

Solution: Define the symmetric matrix

$$\mathbf{P} \stackrel{\text{def}}{=} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$$

as the solution to the Lyapunov equation $\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}$. Since $\mathbf{Q}$ can be any positive definite matrix, let us define $\mathbf{Q} = \mathbf{I}$. Then,

$$\begin{aligned}\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{I} &\implies \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &\implies \begin{bmatrix} p_{12} & p_{22} \\ p_{11} & p_{12} \end{bmatrix} + \begin{bmatrix} p_{12} & p_{11} \\ p_{22} & p_{12} \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &\implies \begin{bmatrix} 2p_{12} & p_{11} + p_{22} \\ p_{11} + p_{22} & 2p_{12} \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &\therefore p_{12} = -1/2, \quad p_{11} + p_{22} = 0.\end{aligned}$$

Since $\mathbf{P}$ is not unique, we conclude that $\mathbf{A}$ is not Hurwitz.

4 Sample applications

This section presents some applications of Lyapunov stability theory to controller design.

4.1 Optimal control

In Lecture 8, we have gone through a lot of trouble deriving the continuous-time algebraic Riccati equation starting from the state and costate equations, but it is actually possible to derive the equation using Lyapunov theorem alone. An optimal control input, denoted $\boldsymbol{u}^*(t)$, should be such that the closed-loop system $\dot{\boldsymbol{x}}(t) = (\mathbf{A} - \mathbf{BK})\boldsymbol{x}(t)$ is asymptotically stable, i.e., there exists a quadratic Lyapunov function $V(\boldsymbol{x}) = \boldsymbol{x}^\top(t)\mathbf{P}\boldsymbol{x}(t)$, where $\mathbf{P}$ is positive definite, such that $\dot{V}(\boldsymbol{x})$ is negative definite on the trajectories of the closed-loop system, as a consequence of the Lyapunov theorem.

Theorem 6: [Żak03, Theorem 5.2]

The state-feedback control law $\boldsymbol{u}^* = -\mathbf{K}\boldsymbol{x}$ is optimal (minimizes J) if for some $V = \boldsymbol{x}^\top\mathbf{P}\boldsymbol{x}$,

$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} \left(\dot{V} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^\top\mathbf{R}\boldsymbol{u}\right), \tag{20}$$

and

$$\dot{V}\Big|_{\boldsymbol{u}=\boldsymbol{u}^*} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^{*\top}\mathbf{R}\boldsymbol{u}^* = 0. \tag{21}$$

Proof: Given

$$\begin{aligned}
\dot{V}\Big|_{\boldsymbol{u}=\boldsymbol{u}^*} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^{*\top}\mathbf{R}\boldsymbol{u}^* = 0 &\implies \int_0^\infty \dot{V}\Big|_{\boldsymbol{u}=\boldsymbol{u}^*} \mathrm{d}t = -\int_0^\infty \left(\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^{*\top}\mathbf{R}\boldsymbol{u}^*\right) \mathrm{d}t \\
&\implies V(\boldsymbol{x}(\infty)) - V(\boldsymbol{x}(0)) = -J(\boldsymbol{u}^*).
\end{aligned}$$

Above, we explicitly express the performance index as a function of $\boldsymbol{u}^*$. Requiring the closed-loop system to be asymptotically stable, we have $\boldsymbol{x}(\infty) = \mathbf{0} \implies V(\boldsymbol{x}(\infty)) = \mathbf{0}^\top\mathbf{P}\mathbf{0} = 0$, thus

$$V(\boldsymbol{x}(0)) = J(\boldsymbol{u}^*).$$

We need to prove optimality, i.e., there is no $\tilde{\boldsymbol{u}} \neq \boldsymbol{u}^*$ s.t. $J(\tilde{\boldsymbol{u}}) < J(\boldsymbol{u}^*) = V(\boldsymbol{x}(0))$. Let us prove by contradiction, by supposing there is a $\tilde{\boldsymbol{u}}$ s.t. $\tilde{\boldsymbol{u}} \neq \boldsymbol{u}^*$ but $J(\tilde{\boldsymbol{u}}) < J(\boldsymbol{u}^*)$. In other words, since $\tilde{\boldsymbol{u}} \neq \boldsymbol{u}^*$,

$$\begin{aligned}
\dot{V}\Big|_{\boldsymbol{u}=\tilde{\boldsymbol{u}}} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \tilde{\boldsymbol{u}}^\top\mathbf{R}\tilde{\boldsymbol{u}} > 0 &\implies -\int_0^\infty \dot{V}\Big|_{\boldsymbol{u}=\tilde{\boldsymbol{u}}} \mathrm{d}t < \int_0^\infty \left(\boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \tilde{\boldsymbol{u}}^\top\mathbf{R}\tilde{\boldsymbol{u}}\right) \mathrm{d}t \\
&\implies V(\boldsymbol{x}(0)) < J(\tilde{\boldsymbol{u}}) \\
&\implies J(\boldsymbol{u}^*) < J(\tilde{\boldsymbol{u}}),
\end{aligned}$$

giving us the contradiction we need. Therefore, Eqs. (20)–(21) necessarily imply $\boldsymbol{u}^*$ is optimal. □

Note that in the proof above, Eq. (18) is not used, and just the fact that $V = \boldsymbol{x}^\top\mathbf{P}\boldsymbol{x}$ is a Lyapunov function suffices. As a corollary of Theorem 6,

$$\begin{aligned}
&\frac{\partial}{\partial \boldsymbol{u}}\left\{\frac{\mathrm{d}V}{\mathrm{d}t} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^\top\mathbf{R}\boldsymbol{u}\right\}\Bigg|_{\boldsymbol{u}=\boldsymbol{u}^*} = \mathbf{0} \\
\implies &\frac{\partial}{\partial \boldsymbol{u}}\left\{2\boldsymbol{x}^\top\mathbf{P}\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x}\right\}\Bigg|_{\boldsymbol{u}=\boldsymbol{u}^*} + 2\boldsymbol{u}^{*\top}\mathbf{R} = \mathbf{0} \\
\implies &\frac{\partial}{\partial \boldsymbol{u}}\left\{2\boldsymbol{x}^\top\mathbf{P}(\mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u}) + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x}\right\}\Bigg|_{\boldsymbol{u}=\boldsymbol{u}^*} + 2\boldsymbol{u}^{*\top}\mathbf{R} = \mathbf{0} \\
\implies &2\boldsymbol{x}^\top\mathbf{P}\mathbf{B} + 2\boldsymbol{u}^{*\top}\mathbf{R} = \mathbf{0}.
\end{aligned}$$

Note that the matrix calculus performed above is in the *numerator layout* or *Jacobian formulation* (see `https://en.wikipedia.org/wiki/Matrix_calculus`). Therefore, the LQR control law is $\boldsymbol{u}^* = -\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\boldsymbol{x}$, confirming the result in Lecture 8. Up to this point, $\mathbf{P}$ remains to be determined, and for this we revisit Eq. (21):

$$\begin{aligned}
&\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{\boldsymbol{u}=\boldsymbol{u}^*} + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{u}^{*\top}\mathbf{R}\boldsymbol{u}^* = 0 \\
\Longrightarrow\ & 2\boldsymbol{x}^\top\mathbf{P}(\mathbf{A}\boldsymbol{x} - \mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\boldsymbol{x}) + \boldsymbol{x}^\top\mathbf{Q}\boldsymbol{x} + \boldsymbol{x}^\top\mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\boldsymbol{x} = 0 \\
\Longrightarrow\ & \boldsymbol{x}^\top\left\{2\cdot\frac{1}{2}\left[\mathbf{P}\mathbf{A} + (\mathbf{P}\mathbf{A})^\top\right] - 2\mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P} + \mathbf{Q} + \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\right\}\boldsymbol{x} = 0 \\
\Longrightarrow\ & \boldsymbol{x}^\top\left\{\mathbf{A}^\top\mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^\top\mathbf{P}\right\}\boldsymbol{x} = 0.
\end{aligned}$$

Above, we have just recovered the *continuous-time algebraic Riccati equation*.

The following discrete-time counterpart of Theorem 6 can be similarly proven:

Theorem 7: [Żak03, Theorem 5.4]

The state-feedback control law $\boldsymbol{u}^*[k] = -\mathbf{K}\boldsymbol{x}[k]$ is optimal (minimizes J) if for some $V(\boldsymbol{x}[k]) = \boldsymbol{x}^\top[k]\mathbf{P}\boldsymbol{x}[k]$,

$$\boldsymbol{u}^*[k] = \arg\min_{\boldsymbol{u}[k]}\left(\Delta V(\boldsymbol{x}[k]) + \boldsymbol{x}^\top[k]\mathbf{Q}\boldsymbol{x}[k] + \boldsymbol{u}^\top[k]\mathbf{R}\boldsymbol{u}[k]\right), \tag{22}$$

and

$$\left.\Delta V(\boldsymbol{x}[k])\right|_{\boldsymbol{u}=\boldsymbol{u}^*} + \boldsymbol{x}^\top[k]\mathbf{Q}\boldsymbol{x}[k] + \boldsymbol{u}^{*\top}[k]\mathbf{R}\boldsymbol{u}^*[k] = 0. \tag{23}$$

From Theorem 7, the discrete-time LQR control law and discrete-time algebraic Riccati equation can be similarly derived [Żak03, Sect. 5.3.5].

4.2 Adaptive control

All nonlinear control approaches, including neuro-fuzzy control and sliding mode control, rely on Lyapunov stability theory. This little excursion to adaptive control here serves merely as an appetizer for further studies in nonlinear control. Besides being important for aerospace applications as hinted at the beginning of this lecture, adaptive control is one of the five most important features in an industrial loop controller [LAC06].

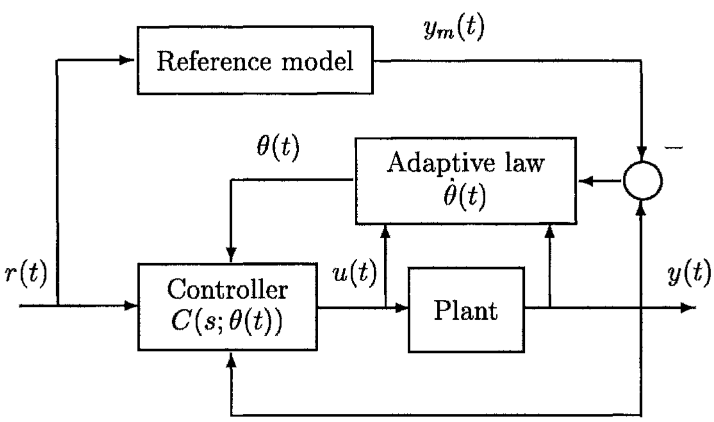


Figure 14: The direct model reference adaptive control system architecture [Tao03, Figure 1.4]. Here, θ denotes a generic controller parameter.

Like a typical control system, an adaptive control system consists of a plant and a controller. However, the plant parameters are potentially unknown, and in the *direct model reference adaptive control* (MRAC) system architecture (see Figure 14), there is an *adaptive law* for updating the controller parameters to achieve some desired system performance.

For simplicity, consider the first-order SISO LTI plant

$$\dot{y}(t) = -a_p y(t) + b_p u(t), \tag{24}$$

where the output $y(t)$ doubles as the state. The control objectives are to ensure *all* closed-loop system signals are bounded, and $y(t)$ tracks asymptotically the output $y_m(t)$ of a reference model:

$$\dot{y}_m(t) = -a_m y_m(t) + b_m r(t), \tag{25}$$

where $a_m > 0$ to ensure asymptotic stability. Note that a_m is user-defined based on performance requirement, and plays the role of the exponential frequency (i.e., inverse of the time constant) of a first-order system. In other words, the performance requirement is embedded in a_m, and not explicitly specified as part of the control objectives. Define the tracking error as

$$e(t) \stackrel{\text{def}}{=} y(t) - y_m(t). \tag{26}$$

The question is can we find suitable gains k_y and k_r for the control law:

$$u(t) = k_y y(t) + k_r r(t), \tag{27}$$

such that $\lim_{t\to\infty} e(t) = 0$? Let us consider the case where a_p and b_p known, and the case where a_p and b_p are unknown, in turn.

Case a_p and b_p are known: By defining

$$k_y = \frac{a_p - a_m}{b_p}, \qquad k_r = \frac{b_m}{b_p}, \tag{28}$$

we can get the closed-loop system:

$$\dot{y}(t) - \dot{y}_m(t) = -a_p y(t) + b_p k_y y(t) + b_p k_r r(t) - (-a_m) y_m(t) - b_m r(t) \implies \dot{e}(t) = -a_m e(t). \tag{29}$$

By the definition of a_m, it is clear that (i) $\lim_{t\to\infty} e(t) = 0$, (ii) all $e(t)$, $y(t)$ and $u(t)$ are bounded, so the control objectives are met.

Case a_p and b_p are unknown: Since a_p is unknown, k_y and k_r are no longer known. This is where the *adaptive law* comes in, which we shall use to estimate k_y and k_r. Rewrite the control law as

$$u(t) = \hat{k}_y(t) y(t) + \hat{k}_r(t) r(t), \tag{30}$$

where $\hat{k}_y(t)$ and $\hat{k}_r(t)$ are estimates of k_y and k_r at time t. The idea of using estimated controller parameters instead of the actual controller parameters is known as the *certainty equivalence principle* [Tao03, p. 32]. Define the gain estimation errors as

$$\tilde{k}_y(t) \stackrel{\text{def}}{=} k_y - \hat{k}_y(t), \qquad \tilde{k}_r(t) \stackrel{\text{def}}{=} k_r - \hat{k}_r(t). \tag{31}$$

Replacing k_y and k_r in Eq. (29) with $\hat{k}_y = k_y - \tilde{k}_y$ and $\hat{k}_y = k_y - \tilde{k}_y$, closed-loop system now becomes:

$$\dot{e}(t) = -a_m e(t) - b_p \left[\tilde{k}_y(t)y(t) + \tilde{k}_r(t)r(t)\right]. \tag{32}$$

As long as $\tilde{k}_y(t)$ and $\tilde{k}_r(t)$ converge to zero, so will $e(t)$ too. The adaptive law is thus a state equation involving $\tilde{k}_y(t)$ and $\tilde{k}_r(t)$ that is asymptotically stable by design. It is exactly this design component that we are going to apply Lyapunov stability theory to — this is why this design component is called a *Lyapunov design* [Tao03, Sect. 1.5.1].

For the Lyapunov design,

- Lyapunov design 1: Suppose we define the Lyapunov function candidate as

$$V = \frac{1}{2}e^2(t) + \frac{1}{2}\tilde{k}_y^2(t) + \frac{1}{2}\tilde{k}_r^2(t), \tag{33}$$

  which is positive definite, then

$$\begin{aligned}
\dot{V} &= e(t)\dot{e}(t) + \tilde{k}_y(t)\dot{\tilde{k}}_y(t) + \tilde{k}_r(t)\dot{\tilde{k}}_r(t) \\
&= -a_m e^2(t) - e(t)b_p \left[\tilde{k}_y(t)y(t) + \tilde{k}_r(t)r(t)\right] + \tilde{k}_y(t)\dot{\tilde{k}}_y(t) + \tilde{k}_r(t)\dot{\tilde{k}}_r(t) \qquad (\because \text{Eq. (32)}) \\
&= -a_m e^2(t) + \tilde{k}_y(t)\left[\dot{\tilde{k}}_y(t) - b_p e(t)y(t)\right] + \tilde{k}_r(t)\left[\dot{\tilde{k}}_r(t) - b_p e(t)r(t)\right].
\end{aligned}$$

  To ensure $\dot{V}$ is at least negative semidefinite, we can define the adaptive laws as

$$\dot{\tilde{k}}_y(t) = b_p e(t)y(t), \qquad \dot{\tilde{k}}_r(t) = b_p e(t)r(t),$$

  but this requires knowledge of b_p.

- Lyapunov design 2: Suppose we define the Lyapunov function candidate as

$$V = \frac{1}{2}e^2(t) + \frac{|b_p|}{2}\tilde{k}_y^2(t) + \frac{|b_p|}{2}\tilde{k}_r^2(t). \tag{34}$$

  Quiz 7

  What is the reason for using $|b_p|$ instead of b_p in the definition above?

  Then,

$$\begin{aligned}
\dot{V} &= e(t)\dot{e}(t) + |b_p|\tilde{k}_y(t)\dot{\tilde{k}}_y(t) + |b_p|\tilde{k}_r(t)\dot{\tilde{k}}_r(t) \\
&= -a_m e^2(t) - e(t)b_p \left[\tilde{k}_y(t)y(t) + \tilde{k}_r(t)r(t)\right] + |b_p|\tilde{k}_y(t)\dot{\tilde{k}}_y(t) + |b_p|\tilde{k}_r(t)\dot{\tilde{k}}_r(t) \qquad (\because \text{Eq. (32)}) \\
&= -a_m e^2(t) + \tilde{k}_y(t)\left[|b_p|\dot{\tilde{k}}_y(t) - b_p e(t)y(t)\right] + \tilde{k}_r(t)\left[|b_p|\dot{\tilde{k}}_r(t) - b_p e(t)r(t)\right].
\end{aligned}$$

  To ensure $\dot{V}$ is at least negative semidefinite, we can define the adaptive laws as

$$\begin{aligned}
\dot{\tilde{k}}_y(t) &= \frac{b_p}{|b_p|}e(t)y(t) = \text{sign}(b_p)e(t)y(t) \implies \dot{\hat{k}}_y(t) = -\,\text{sign}(b_p)e(t)y(t), \\
\dot{\tilde{k}}_r(t) &= \frac{b_p}{|b_p|}e(t)r(t) = \text{sign}(b_p)e(t)r(t) \implies \dot{\hat{k}}_r(t) = -\,\text{sign}(b_p)e(t)r(t).
\end{aligned}$$

  Compared to Lyapunov design 1, Lyapunov design 2 requires only the knowledge of the sign of b_p, which in all likelihood should be available. For implementation,

$$\hat{k}_y(t) = -\operatorname{sign}(b_p)\gamma_y \int_0^t e(\tau)y(\tau)\,\mathrm{d}\tau + \hat{k}_y(0), \tag{35}$$

$$\hat{k}_r(t) = -\operatorname{sign}(b_p)\gamma_r \int_0^t e(\tau)r(\tau)\,\mathrm{d}\tau + \hat{k}_r(0). \tag{36}$$

Above, the so-called *adaptive gains* γ_y and γ_r are added for tuning the rate of convergence of the adaptive laws [Din13, p. 91]. As usual, these gains cannot be too large lest they destabilize the system. Adding the adaptive gains is equivalent to changing the Lyapunov function of Lyapunov design 2 to

$$V = \frac{1}{2}e^2(t) + \frac{|b_p|}{2\gamma_y}\tilde{k}_y^2(t) + \frac{|b_p|}{2\gamma_r}\tilde{k}_r^2(t).$$

Note that since $\dot{V}$ only depends on $e(t)$, and not $\tilde{k}_y(t)$ and $\tilde{k}_r(t)$, we cannot conclude $\dot{V}$ is negative definite, although it is certainly negative semidefinite. However, the fact that V is a Lyapunov function implies $e(t)$, $\hat{k}_y(t) = k_y - \tilde{k}_y(t)$ and $\hat{k}_r(t) = k_r - \tilde{k}_r(t)$ are bounded. By inference, $y(t) = e(t) + y_m(t)$ and $u(t)$ as defined in Eq. (30) are bounded. Showing $e(t)$ converges to zero requires Barbalat's lemma:

Theorem 8: Barbalat's lemma [Din13, Lemma 7.1], [SL91, Lemma 4.2]

If a function $f : \mathbb{R} \mapsto \mathbb{R}$ is uniformly continuous for $t \in [0, \infty)$, and $\int_0^\infty f(t)\,\mathrm{d}t$ exists, then $\lim_{t\to\infty} f(t) = 0$. Note: A sufficient condition for uniform continuity is bounded derivative.

According to Eq. (32), and the fact that $e(t), \tilde{k}_y(t), y(t), \tilde{k}_r(t), r(t)$ are bounded, $\dot{e}(t)$ is bounded, so $e(t)$ and hence $e^2(t)$ are uniformly continuous. Furthermore, since

$$\dot{V} = -a_m e^2(t) \implies \int_0^\infty e^2(\tau)\,\mathrm{d}\tau = \frac{V|_{t=0} - V|_{t=\infty}}{a_m} < \infty,$$

according to Barbalat's lemma, $\lim_{t\to\infty} e(t) = 0$, i.e., the tracking error converges to zero.

In conclusion, for the first-order SISO LTI plant (24) of unknown parameters, the adaptive controller (30) using adaptive laws (35)–(36) can be used to track the output of the reference model (25). A discussion of multivariable MRAC can be found in [Tao03, Sect. 9.2], but the most popular adaptive control approach to date is $\mathcal{L}_1$ adaptive control because of its robustness properties [HC10]. The highly compressed introduction to adaptive control above also contains an introduction to Barbalat's lemma, which is a powerful tool that is often used in conjunction with Lyapunov's direct method to analyze the stability of nonlinear control systems.

References

[BW74] Roxanne Mary Byrne and Edward T. Wall. A synthesis of Lyapunov's first and second methods. *International Journal of Systems Science*, 5(12):1179–1191, 1974.

[Che99] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, 3rd edition, 1999.

[DAL10] Z. T. Dydek, A. M. Annaswamy, and E. Lavretsky. Adaptive Control and the NASA X-15-3 Flight Revisited. *IEEE Control Syst. Mag.*, 30(3):32–48, June 2010.

[DBDN13] Warren E Dixon, Aman Behal, Darren M Dawson, and Siddharth P Nagarkatti. *Nonlinear control of engineering systems: A Lyapunov-based approach*. Springer Science & Business Media, 2013.

[Din13] Zhengtao Ding. *Nonlinear and Adaptive Control Systems*, volume 84 of *IET Control Enginering Series*. The Institution of Engineering and Technology, 2013.

[FPEN15] G.F. Franklin, J.D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Pearson Education Limited, 7th edition, 2015.

[HC10] Naira Hovakimyan and Chengyu Cao. *$\mathcal{L}_1$ Adaptive Control Theory: Guaranteed Robustness with Fast Adaptation*. Society for Industrial and Applied Mathematics, 2010.

[Hes09] João P Hespanha. *Linear systems theory*. Princeton University Press, 2009.

[KA01] Branislav Kisačanin and Gyan C. Agarwal. *Linear Control Systems: With solved problems and MATLAB examples*. The University Series in Mathematics. Springer US, 2001.

[Kai80] Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc., 1980.

[Kha02] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, 3rd edition, 2002.

[Kha15] Hassan K. Khalil. *Nonlinear Control*. Pearson, global edition, 2015.

[LAC06] Yun Li, Kiam Heong Ang, and G. C. Y. Chong. Patents, software, and hardware for PID control: an overview and analysis of the current art. *IEEE Control Systems*, 26(1):42–54, February 2006.

[LW13] E. Lavretsky and K.A. Wise. *Robust and Adaptive Control: with Aerospace Applications*. Springer-Verlag London, 2013.

[Nis15] Norman S. Nise. *Control Systems Engineering*. Wiley, 7th edition, 2015.

[NPWW02] Hung T. Nguyen, Nadipuram R. Prasad, Carol L. Walker, and Ebert A. Walker. *A First Course in Fuzzy and Neural Control*. CRC Press, 2002.

[SL91] Jean-Jacques E. Slotine and Weiping Li. *Applied Nonlinear Control*. Prentice Hall, 1991.

[Son98] Eduardo D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer, 2nd edition, 1998.

[Son08] Eduardo D Sontag. Input to state stability: Basic concepts and results. In *Nonlinear and optimal control theory*, pages 163–220. Springer, 2008.

[Tao03] Gang Tao. *Adaptive control design and analysis*, volume 37 of *Adaptive and Learning Systems for Signal Processing, Communications, and Control*. John Wiley & Sons, Inc., 2003.

[Ter09] W.J. Terrell. *Stability and Stabilization: An Introduction*. Princeton University Press, 2009.

[WL07] Robert L. Williams II and Douglas A. Lawrence. *Linear State-Space Control Systems*. John Wiley & Sons, 2007.

[Żak03] Stanislaw H. Żak. *Systems and Control*. Oxford University Press, 2003.