EE C222/ ME C237 - Spring'18 - Lecture 1 Notes[1]

Murat Arcak

January 17 2018

Nonlinear Systems

$$\dot{x} = Ax + Bu \quad \longrightarrow \quad \dot{x} = f(x, u) \tag{1}$$

Analysis:

$$\begin{array}{lll} \dot{x} = f(x) & f : \mathbb{R}^n \to \mathbb{R}^n & \text{time-invariant (autonomous)} \\ \dot{x} = f(t, x) & f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n & \text{time-varying (non-autonomous)} \end{array}$$

We use the shorthand notation $\dot{x} = f(x)$ for $\frac{d}{dt}x(t) = f(x(t))$.

Design:

$$\dot{x} = f(x, u) \quad u \text{ to be designed as a function of } x.$$

Equilibria

$x = x^*$ is an equilibrium for $\dot{x} = f(x)$ if $f(x^*) = 0$.

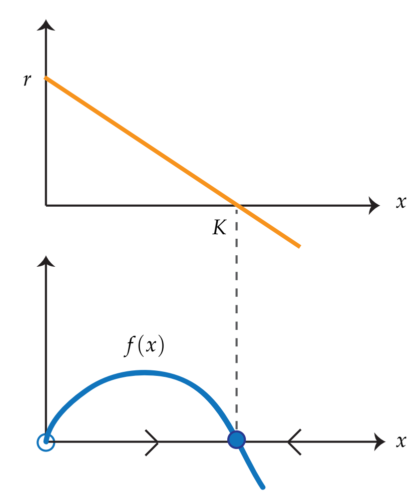
Example: Linear system $\dot{x} = Ax$.

If A is nonsingular, $x^* = 0$ is the unique equilibrium.

If A is singular, the nullspace defines a continuum of equilibria.

Example: *Logistic growth model* in population dynamics

$$\dot{x} = f(x) = \underbrace{r\left(1 - \frac{x}{K}\right)}_{\text{growth rate}} x, \qquad r > 0, \;\; K > 0 \tag{2}$$



$x > 0$ denotes the population and K is called the carrying capacity.

For systems with a scalar state variable $x \in \mathbb{R}$, stability can be determined from the sign of $f(x)$ around the equilibrium. In this example $f(x) > 0$ for $x \in (0, K)$, and $f(x) < 0$ for $x > K$; therefore

$$\begin{array}{ll} x = 0 & \text{unstable equilibrium} \\ x = K & \text{asymptotically stable.} \end{array}$$

Linearization

Local stability properties of x^* can be determined by linearizing the vector field $f(x)$ at x^*:

$$f(x^* + \tilde{x}) = \underbrace{f(x^*)}_{=\ 0} + \underbrace{\left.\frac{\partial f}{\partial x}\right|_{x=x^*}}_{\triangleq\ A} \tilde{x} + \text{higher order terms} \qquad (3)$$

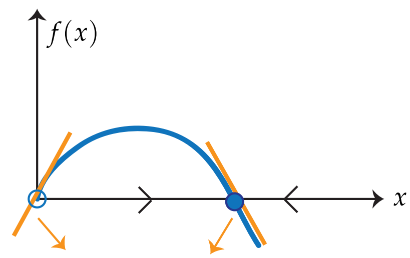
Thus, the linearized model is:

$$\dot{\tilde{x}} = A\tilde{x}. \qquad (4)$$

If $\Re\lambda_i(A) < 0$ for each eigenvalue λ_i of A, then x^* is asymp. stable.

If $\Re\lambda_i(A) > 0$ for some eigenvalue λ_i of A, then x^* is unstable.

Example: Logistic growth model above:



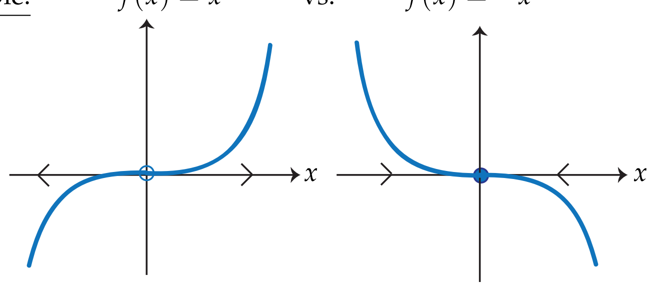
Caveats:

1. Only local properties can be determined from the linearization.

   Example: The logistic growth model linearized at $x = 0$ ($\dot{x} = rx$) would incorrectly predict unbounded growth of $x(t)$. In reality, $x(t) \to K$.

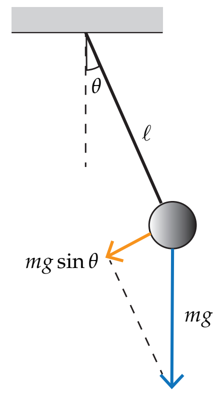
2. If $\Re\lambda_i(A) \leq 0$ with equality for some i, then linearization is inconclusive as a stability test. Higher order terms determine stability.

   Example: $f(x) = x^3$ vs. $f(x) = -x^3$



$f'(0) = 0$ in each case, but one is stable and the other is unstable.

Second order example: Pendulum



$$\ell m\ddot{\theta} = -k\ell\dot{\theta} - mg\sin\theta \tag{5}$$

Define $x = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}$. State space: $S^1 \times \mathbb{R}$.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k}{m}x_2 - \frac{g}{\ell}\sin x_1 \end{aligned} \tag{6}$$

Equilibria: $(0,0)$ and $(\pi,0)$

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{\ell}\cos x_1 & -\frac{k}{\ell} \end{bmatrix} = \begin{cases} \begin{bmatrix} 0 & 1 \\ -\frac{g}{\ell} & -\frac{k}{\ell} \end{bmatrix} & \text{(stable) at } x_1 = 0 \\ \begin{bmatrix} 0 & 1 \\ \frac{g}{\ell} & -\frac{k}{\ell} \end{bmatrix} & \text{(unstable) at } x_1 = \pi \end{cases}$$

Phase portrait: plot of $x_1(t)$ vs. $x_2(t)$ for 2nd order systems

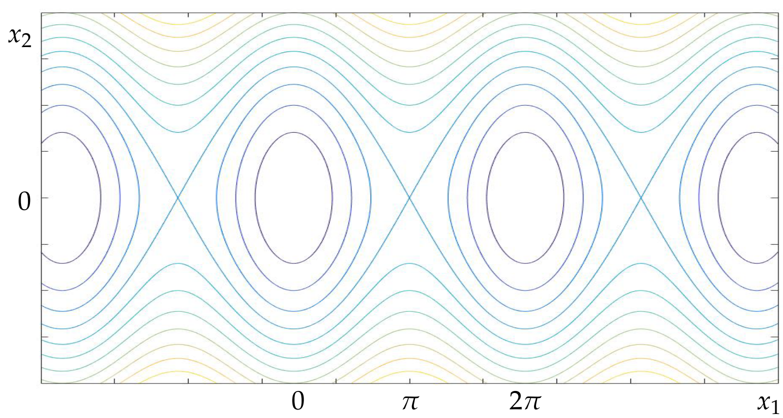


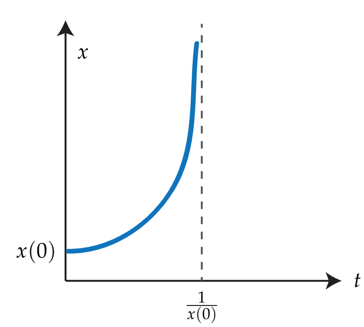
Figure 1: Phase portrait of the pendulum for the undamped case $k = 0$.

Essentially Nonlinear Phenomena

1. Finite Escape Time

   Example: $\dot{x} = x^2$

$$\begin{aligned} &\frac{d}{dt}x^{-1} = -x^{-2}\dot{x} = -1 \\ \Rightarrow\quad &\frac{1}{x(t)} - \frac{1}{x(0)} = -t \\ \Rightarrow\quad &x(t) = \frac{1}{\frac{1}{x(0)} - t} \end{aligned} \tag{7}$$



   For linear systems, $x(t) \to \infty$ cannot happen in finite time.

2. Multiple Isolated Equilibria

   Linear systems: either unique equilibrium or a continuum

   Pendulum: two isolated equilibria (one stable, one unstable)

   "Multi-stable" systems: two or more stable equilibria

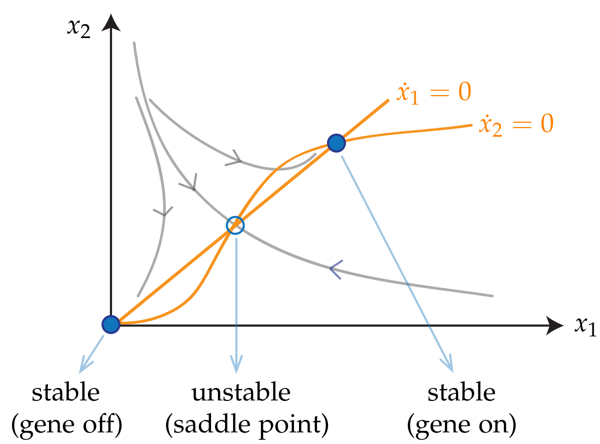
   Example: bistable switch

$$\begin{aligned} \dot{x}_1 &= -ax_1 + x_2 & x_1 &: \text{ concentration of protein} \\ \dot{x}_2 &= \frac{x_1^2}{1+x_1^2} - bx_2 & x_2 &: \text{ concentration of mRNA} \end{aligned} \tag{8}$$

   $a > 0$, $b > 0$ are constants. State space: $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$.

   This model describes a positive feedback where the protein encoded by a gene stimulates more transcription via the term $\frac{x_1^2}{1+x_1^2}$.

   Single equilibrium at the origin when $ab > 0.5$. If $ab < 0.5$, the line where $\dot{x}_1 = 0$ intersects the sigmoidal curve where $\dot{x}_2 = 0$ at two other points, giving rise to a total of three equilibria:



EE C222/ME C237 - Spring'18 - Lecture 2 Notes[1]

Murat Arcak

January 22 2018

Essentially Nonlinear Phenomena Continued

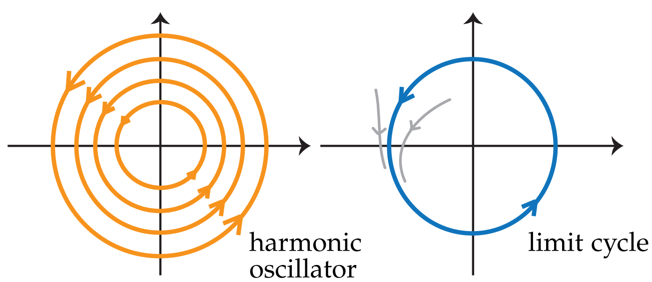
1. Finite escape time

2. Multiple isolated equilibria

3. Limit cycles: Linear oscillators exhibit a continuum of periodic orbits; *e.g.*, every circle is a periodic orbit for $\dot{x} = Ax$ where

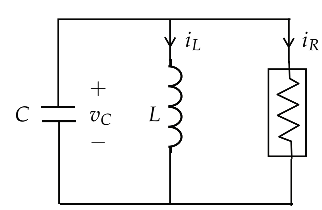
$$A = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix} \qquad (\lambda_{1,2} = \mp j\beta).$$

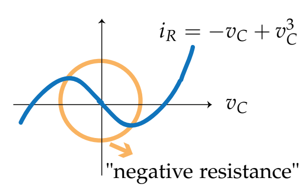
In contrast, a limit cycle is an isolated periodic orbit and can occur only in nonlinear systems.

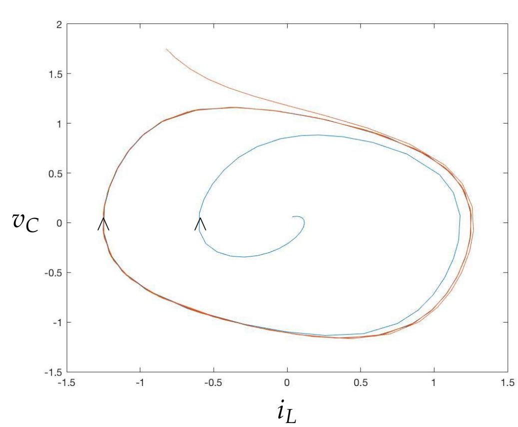


<u>Example:</u> van der Pol oscillator

$$\begin{aligned} C\dot{v}_C &= -i_L + v_C - v_C^3 \\ L\dot{i}_L &= v_C \end{aligned}$$





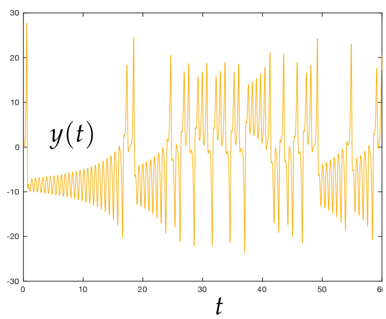


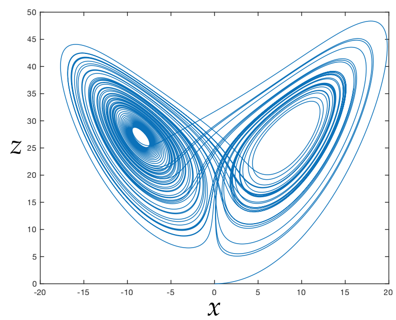
4. Chaos: Irregular oscillations, never exactly repeating.

   Example: Lorenz system (derived by Ed Lorenz in 1963 as a simplified model of convection rolls in the atmosphere):

$$\begin{aligned}\dot{x} &= \sigma(y-x)\\ \dot{y} &= rx - y - xz\\ \dot{z} &= xy - bz.\end{aligned}$$

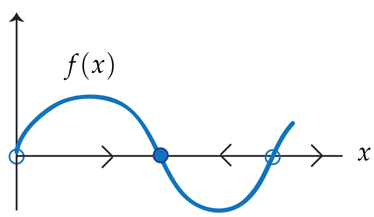
   Chaotic behavior with $\sigma = 10,\ b = 8/3,\ r = 28$:





   - For continuous-time, time-invariant systems, $n \geq 3$ state variables required for chaos.

     $n = 1$: $x(t)$ monotone in t, no oscillations:



     $n = 2$: Poincaré-Bendixson Theorem (to be studied in Lecture 3) guarantees regular behavior.
   - Poincaré-Bendixson does not apply to time-varying systems and $n \geq 2$ is enough for chaos (Homework problem).
   - For discrete-time systems, $n = 1$ is enough (we will see an example in Lecture 5).

Planar (Second Order) Dynamical Systems

Chapter 2 in both Sastry and Khalil

Phase Portraits of Linear Systems: $\dot{x} = Ax$

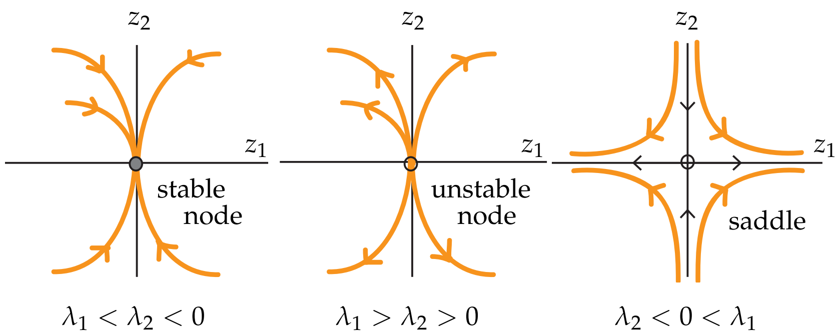
- Distinct real eigenvalues

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

In $z = T^{-1}x$ coordinates:

$$\dot{z}_1 = \lambda_1 z_1, \quad \dot{z}_2 = \lambda_2 z_2.$$

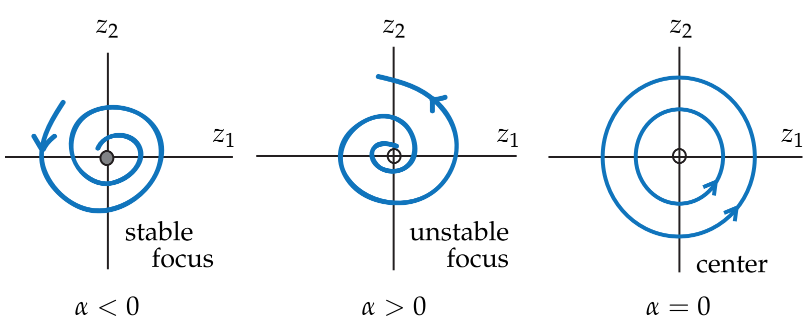
The equilibrium is called a *node* when λ_1 and λ_2 have the same sign (*stable* node when negative and *unstable* when positive). It is called a *saddle point* when λ_1 and λ_2 have opposite signs.



- Complex eigenvalues: $\lambda_{1,2} = \alpha \mp j\beta$

$$T^{-1}AT = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$$

$$\begin{aligned} \dot{z}_1 &= \alpha z_1 - \beta z_2 \\ \dot{z}_2 &= \alpha z_2 + \beta z_1 \end{aligned} \quad \rightarrow \quad \text{polar coordinates} \quad \rightarrow \quad \begin{aligned} \dot{r} &= \alpha r \\ \dot{\theta} &= \beta \end{aligned}$$

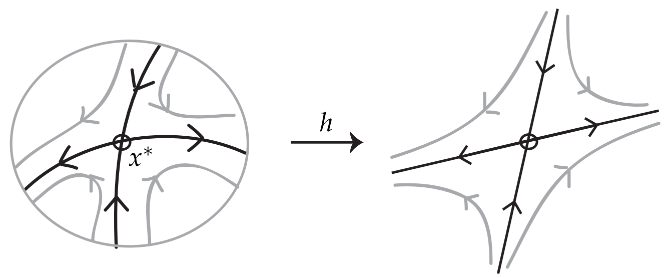


The phase portraits above assume $\beta > 0$ so that the direction of rotation is counter-clockwise: $\dot{\theta} = \beta > 0$.

Phase Portraits of Nonlinear Systems Near Hyperbolic Equilibria

hyperbolic equilibrium: linearization has no eigenvalues on the imaginary axis

Phase portraits of nonlinear systems near hyperbolic equilibria are qualitatively similar to the phase portraits of their linearization. According to the Hartman-Grobman Theorem (below) a "continuous deformation" maps one phase portrait to the other.



Hartman-Grobman Theorem: If x^* is a hyperbolic equilibrium of $\dot{x} = f(x), x \in \mathbb{R}^n$, then there exists a *homeomorphism*[2] $z = h(x)$ defined in a neighborhood of x^* that maps trajectories of $\dot{x} = f(x)$ to those of $\dot{z} = Az$ where $A \triangleq \frac{\partial f}{\partial x}\Big|_{x=x^*}$.

[2] a continuous map with a continuous inverse

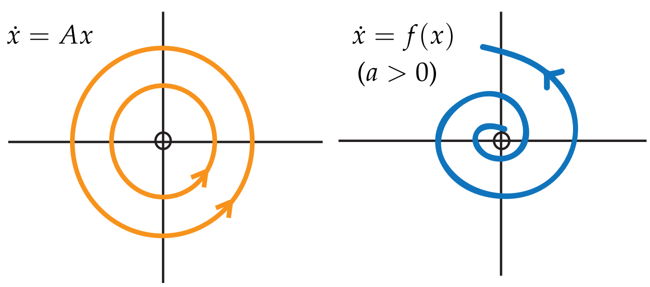
The hyperbolicity condition can't be removed:

Example:

$$\begin{aligned} \dot{x}_1 &= -x_2 + ax_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= x_1 + ax_2(x_1^1 + x_2^2) \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \dot{r} &= ar^3 \\ \dot{\theta} &= 1 \end{aligned}$$

$$x^* = (0,0) \qquad A = \frac{\partial f}{\partial x}\Big|_{x=x^*} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

There is no continuous deformation that maps the phase portrait of the linearization to that of the original nonlinear model:



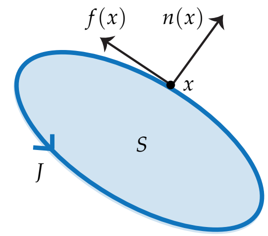
Periodic Orbits in the Plane

Bendixson's Theorem: For a time-invariant planar system

$$\dot{x}_1 = f_1(x_1, x_2) \quad \dot{x}_2 = f_2(x_1, x_2),$$

if $\nabla \cdot f(x) = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2}$ is not identically zero and does not change sign in a simply connected region D, then there are no periodic orbits lying entirely in D.

Proof: By contradiction. Suppose a periodic orbit J lies in D. Let S denote the region enclosed by J and $n(x)$ the normal vector to J at x. Then $f(x) \cdot n(x) = 0$ for all $x \in J$. By the Divergence Theorem:



$$\underbrace{\int_J f(x) \cdot n(x) d\ell}_{=0} = \underbrace{\iint_S \nabla \cdot f(x) dx}_{\neq 0}.$$

Example: $\dot{x} = Ax, x \in \mathbb{R}^2$ can have periodic orbits only if

Trace$(A) = 0$, *e.g.*,

$$A = \begin{bmatrix} 0 & -\beta \\ \beta & 0 \end{bmatrix}.$$

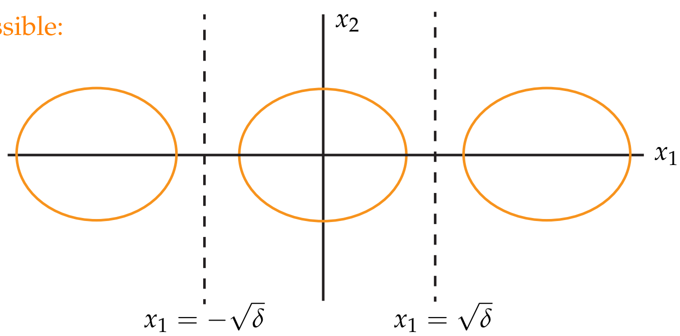
Example:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\delta x_2 + x_1 - x_1^3 + x_1^2 x_2 \quad \delta > 0 \end{aligned}$$

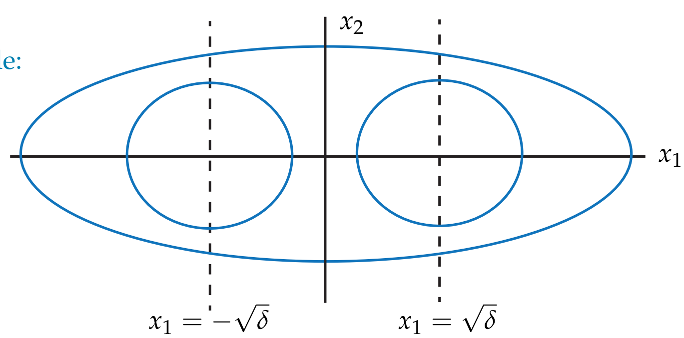
$$\nabla \cdot f(x) = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = x_1^2 - \delta$$

Therefore, no periodic orbit can lie entirely in the region $x_1 \leq -\sqrt{\delta}$ where $\nabla \cdot f(x) \geq 0$, or $-\sqrt{\delta} \leq x_1 \leq \sqrt{\delta}$ where $\nabla \cdot f(x) \leq 0$, or $x_1 \geq \sqrt{\delta}$ where $\nabla \cdot f(x) \geq 0$.

not possible:



possible:



EE C222/ME C237 - Spring'18 - Lecture 3 Notes[1]

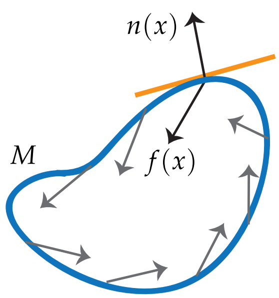
Murat Arcak

January 24 2018

Invariant Sets

Notation: $\phi(t, x_0)$ denotes a trajectory of $\dot{x} = f(x)$ with initial condition $x(0) = x_0$.

Definition: A set $M \subset \mathbb{R}^n$ is positively (negatively) invariant if, for each $x_0 \in M$, $\phi(t, x_0) \in M$ for all $t \geq 0$ ($t \leq 0$).

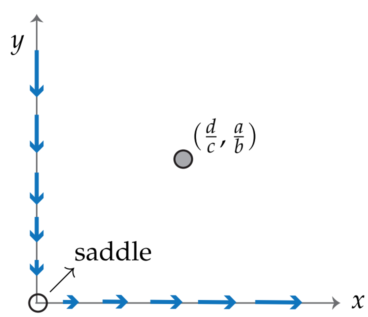


If $f(x) \cdot n(x) \leq 0$ on the boundary then M is positively invariant.

Example 1: A predator-prey model

$$\begin{aligned} \dot{x} &= (a - by)x && \text{Prey (exponential growth when } y = 0) \\ \dot{y} &= (cx - d)y && \text{Predator (exponential decay when } x = 0) \\ & a, b, c, d, > 0 \end{aligned}$$

The nonnegative quadrant is invariant:



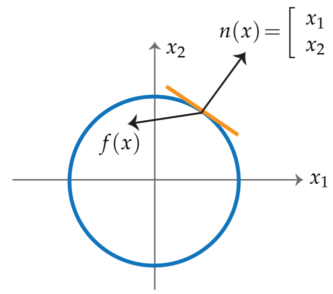
Example 2:

$$\begin{aligned} \dot{x}_1 &= x_1 + x_2 - x_1(x_1^2 + x_2^2) \\ \dot{x}_2 &= -2x_1 + x_2 - x_2(x_1^2 + x_2^2) \end{aligned}$$

Show that $B_r \triangleq \{x | x_1^2 + x_2^2 \leq r^2\}$ is positively invariant for sufficiently large r.

$$f(x) \cdot n(x) = x_1^2 + x_1x_2 - x_1^2(x_1^2 + x_2^2) - 2x_1x_2 + x_2^2 - x_2^2(x_1^2 + x_2^2)$$
$$= -x_1x_2 + (x_1^2 + x_2^2) - (x_1^2 + x_2^2)^2$$
$$-x_1x_2 \leq \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \quad \text{(completion of squares)}$$

Therefore, $f(x) \cdot n(x) \leq \frac{3}{2}r^2 - r^4 \leq 0$ if $r^2 \geq \frac{3}{2}$.



Periodic Orbits in the Plane Continued

Two criteria:

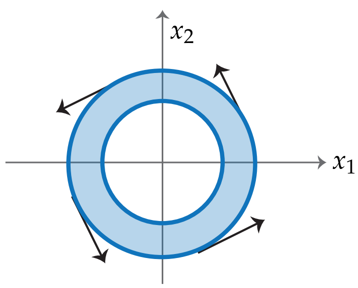
1. Bendixson (absence of periodic orbits)
2. Poincaré-Bendixson (existence of periodic orbits)

Poincaré-Bendixson Theorem: Suppose M is compact[2] and positively invariant for the planar, time invariant system $\dot{x} = f(x), x \in \mathbb{R}^2$. If M contains no equilibrium points, then it contains a periodic orbit.

[2] *i.e.*, closed and bounded

Example 3: Harmonic Oscillator

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= x_1. \end{aligned}$$

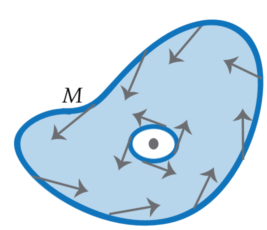


For any $R > r > 0$, the ring $\{x : r^2 \leq x_1^2 + x_2^2 \leq R^2\}$ is compact, invariant and contains no equilibria $\Rightarrow$ at least one periodic orbit. (We know there are infinitely many in this case.)

The "no equilibrium" condition in the PB theorem can be relaxed as:

> "If M contains one equilibrium which is an unstable focus or unstable node"

Proof sketch: Since the equilibrium is an unstable focus or node, we can encircle it with a small closed curve on which $f(x)$ points outward. Then the set obtained from M by carving out the interior of the closed curve is positively invariant and contains no equilibrium.



Example 2 above: B_r is positively invariant for $r \geq \sqrt{\frac{3}{2}}$ but contains the equilibrium $x = 0$.

$$\left.\frac{\partial f}{\partial x}\right|_{x=0} = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} \qquad \lambda_{1,2} = 1 \mp j\sqrt{2} \qquad \text{unstable focus.}$$

Therefore, B_r must contain a periodic orbit.

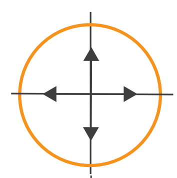
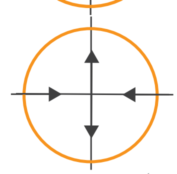
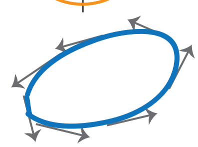
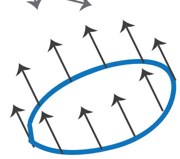
A more general form of the PB Theorem states that, for time invariant, planar systems, bounded trajectories converge to equilibria, periodic orbits, or unions of equilibria connected by trajectories.

Corollary: No chaos for time invariant planar systems.

Index Theory

Again, applicable only to planar systems.

Definition (index): The index of a closed curve is k if, when traversing the curve in one direction, $f(x)$ rotates by $2\pi k$ in the same direction. The index of an equilibrium is defined to be the index of a small curve around it that doesn't enclose another equilibrium.

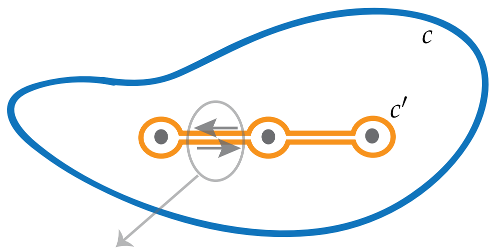
type of equilibrium or curve	index	
node, focus, center	+1	
saddle	-1	
any closed orbit	+1	
a closed curve not encircling any equilibria	0	

The last claim (index $= 0$) follows from the following observations:

- Continuously deforming a closed curve without crossing equilibria leaves its index unchanged.
- A curve not encircling equilibria can be shrunk to an arbitrarily small one, so $f(x)$ can be considered constant.

<u>Theorem:</u> The index of a closed curve is equal to the sum of indices of the equilibria inside.

Graphical proof: Shrinking curve c to c' below without crossing equilibria does not change the index. The index of c' is the sum of the indices of the curves encircling the equilibria because the thin "pipes" connecting these curves do not affect the index of c'.



The following corollary is useful for ruling out periodic orbits (like Bendixson's Theorem studied in the previous lecture):

Corollary: Inside any periodic orbit there must be at least one equilibrium and the indices of the equilibria enclosed must add up to $+1$.

<u>Example (from last lecture):</u>

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\delta x_2 + x_1 - x_1^3 + x_1^2 x_2 \quad \delta > 0 \end{aligned}$$

Bendixson's Criterion: No periodic orbit can lie entirely in one of the regions $x_1 \le -\sqrt{\delta}$, $-\sqrt{\delta} \le x_1 \le \sqrt{\delta}$, or $x_1 \ge \sqrt{\delta}$.

Now apply the corollary above.

Equilibria: $(0,0)$, $(\mp 1, 0)$. To find their indices evaluate the Jacobian:

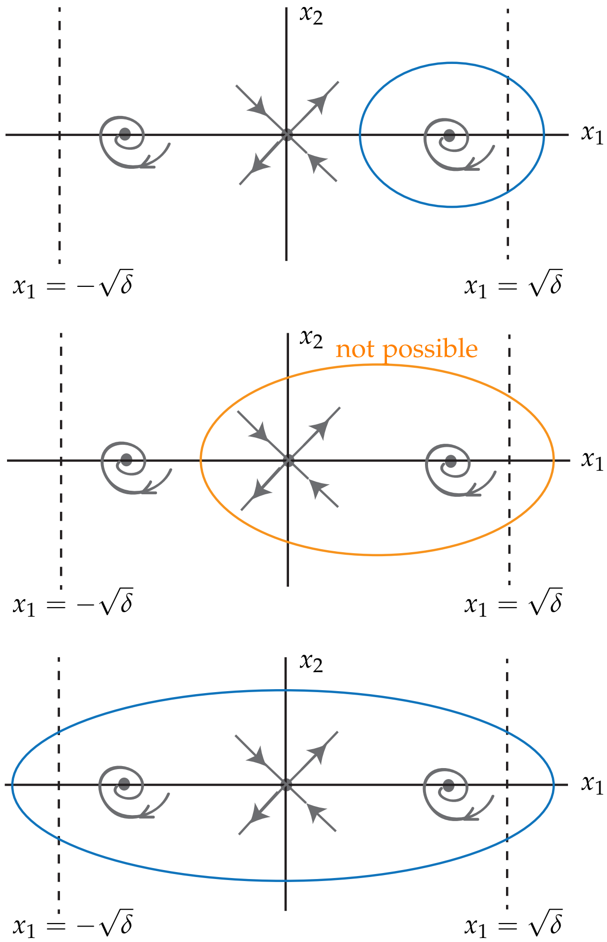
$$\left.\frac{\partial f}{\partial x}\right|_{x=(0,0)} = \begin{bmatrix} 0 & 1 \\ 1 & -\delta \end{bmatrix} \qquad \lambda^2 + \delta\lambda \underbrace{-1}_{<0} = 0.$$

The eigenvalues are real and have opposite signs, therefore $(0,0)$ is a saddle: index $= -1$.

$$\left.\frac{\partial f}{\partial x}\right|_{x=(\mp 1,0)} = \begin{bmatrix} 0 & 1 \\ -2 & 1-\delta \end{bmatrix} \qquad \lambda^2 + (\delta - 1)\lambda \underbrace{+2}_{>0} = 0.$$

The eigenvalues are either real with the same sign (node) or complex conjugates (focus or center), therefore $(\mp 1, 0)$ each has index$= +1$.

Thus, the corollary above rules out the periodic orbit in the middle plot below. It does not rule out the others, but does not prove their existence either. Bendixson's Criterion rules out neither of the three.



x2
x1
x1 = −√δ
x1 = √δ
x2
not possible
x1
x1 = −√δ
x1 = √δ
x2
x1
x1 = −√δ
x1 = √δ

EE C222/ ME C237 - Spring'18 - Lecture 4 Notes[1]

Murat Arcak

January 29 2018

Bifurcations

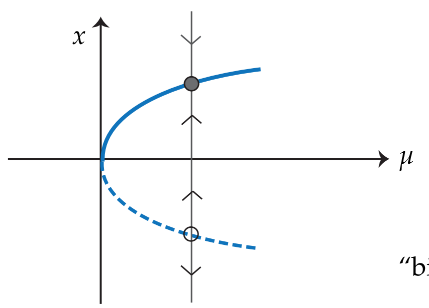
A *bifurcation* is an abrupt change in qualitative behavior as a parameter is varied. Examples: equilibria or limit cycles appearing/disappearing, becoming stable/unstable.

Fold Bifurcation

Also known as "saddle node" or "blue sky" bifurcation.

Example: $\dot{x} = \mu - x^2$

If $\mu > 0$, two equilibria: $x = \mp\sqrt{\mu}$. If $\mu < 0$, no equilibria.



"bifurcation diagram"

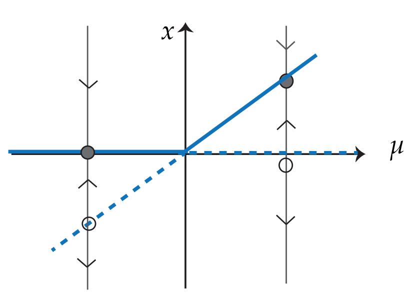
Transcritical Bifurcation

Example: $\dot{x} = \mu x - x^2$

Equilibria: $x = 0$ and $x = \mu$. $\quad \dfrac{\partial f}{\partial x} = \mu - 2x = \begin{cases} \mu & \text{if } x = 0 \\ -\mu & \text{if } x = \mu \end{cases}$

$\mu < 0:\;$ $x = 0$ is stable, $x = \mu$ is unstable

$\mu > 0:\;$ $x = 0$ is unstable, $x = \mu$ is stable

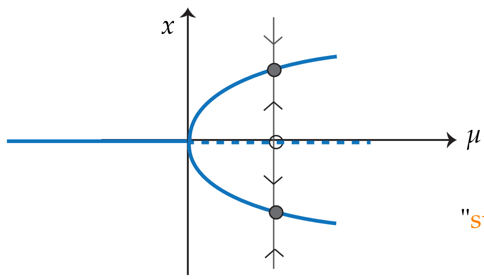


Pitchfork Bifurcation

Example: $\dot{x} = \mu x - x^3$

Equilibria: $x = 0$ for all μ, $x = \mp\sqrt{\mu}$ if $\mu > 0$.

	$\mu < 0$	$\mu > 0$
$\left.\frac{\partial f}{\partial x}\right\|_{x=0} = \mu$	stable	unstable
$\left.\frac{\partial f}{\partial x}\right\|_{x=\mp\sqrt{\mu}} = -2\mu$	N/A	stable

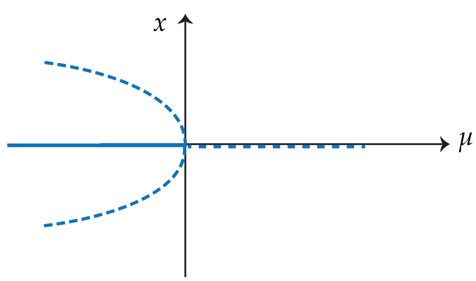


"supercritical pitchfork"

Example: $\dot{x} = \mu x + x^3$

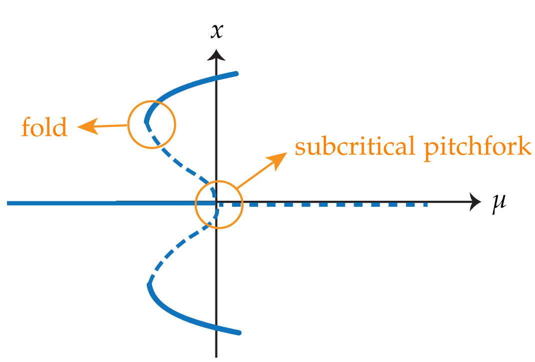
Equilibria: $x = 0$ for all μ, $x = \mp\sqrt{-\mu}$ if $\mu < 0$.

	$\mu < 0$	$\mu > 0$
$\left.\frac{\partial f}{\partial x}\right\|_{x=0} = \mu$	stable	unstable
$\left.\frac{\partial f}{\partial x}\right\|_{x=\mp\sqrt{-\mu}} = -2\mu$	unstable	N/A

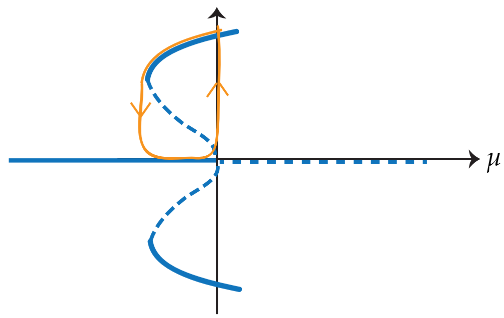


"subcritical pitchfork"

Example: $\dot{x} = \mu x + x^3 - x^5$



Hysteresis arising from a subcritical pitchfork bifurcation:



Bifurcation and hysteresis in perception:

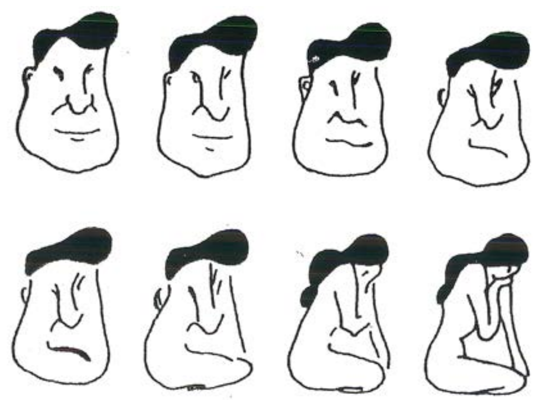


Figure 1: Observe the transition from a man's face to a sitting woman as you trace the figures from left to right, starting with the top row. When does the opposite transition happen as you trace back from the end to the beginning? [Fisher, 1967]

Higher Order Systems

Fold, transcritical, and pitchfork are one-dimensional bifurcations, as evident from the first order examples above. They occur in higher order systems too, but are restricted to a one-dimensional *manifold.*

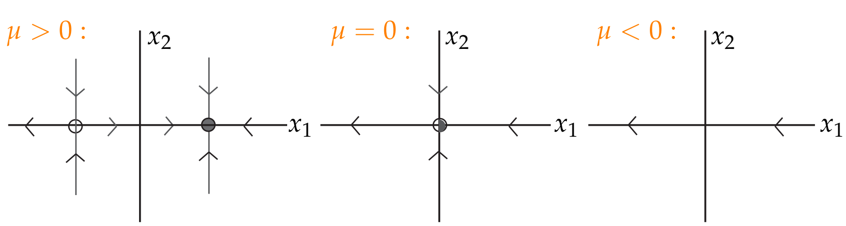
1D subspace: $c_1^T x = \cdots = c_{n-1}^T x = 0$

1D manifold: $g_1(x) = \cdots = g_{n-1}(x) = 0$

Example 1:

$$\begin{aligned} \dot{x}_1 &= \mu - x_1^2 \\ \dot{x}_2 &= -x_2 \end{aligned}$$

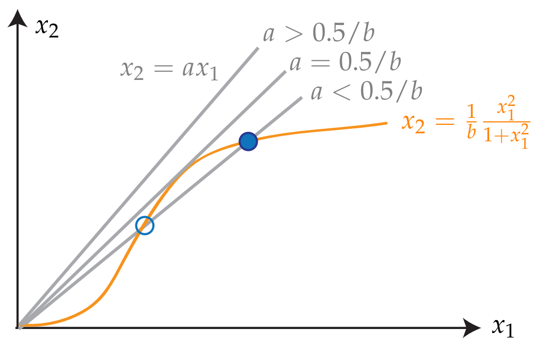
A fold bifurcation occurs on the invariant $x_2 = 0$ subspace:



Example 2: bistable switch (Lecture 1)

$$\dot{x}_1 = -ax_1 + x_2$$
$$\dot{x}_2 = \frac{x_1^2}{1+x_1^2} - bx_2$$

A fold bifurcation occurs at $\mu \triangleq ab = 0.5$:



Characteristic of one-dimensional bifurcations:

$$\boxed{\left.\frac{\partial f}{\partial x}\right|_{\mu=\mu^c,\, x=x^*(\mu^c)} \quad \text{has an eigenvalue at zero}}$$

where $x^*(\mu)$ is the equilibrium point undergoing bifurcation and μ^c is the critical value at which the bifurcation occurs.

Example 1 above:

$$\left.\frac{\partial f}{\partial x}\right|_{\mu=0,x=0} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow \lambda_{1,2} = \boxed{0}, -1$$

Example 2 above:

$$\left.\frac{\partial f}{\partial x}\right|_{\mu=\frac{1}{2},x_1=1,x_2=a} = \begin{bmatrix} -a & 1 \\ \frac{1}{2} & -b \end{bmatrix} \rightarrow \lambda_{1,2} = \boxed{0}, -(a+b)$$

Hopf Bifurcation

Two-dimensional bifurcation unlike the one-dimensional types above.

Example: Supercritical Hopf bifurcation

$$\begin{aligned} \dot{x}_1 &= x_1(\mu - x_1^2 - x_2^2) - x_2 \\ \dot{x}_2 &= x_2(\mu - x_1^2 - x_2^2) + x_1 \end{aligned}$$

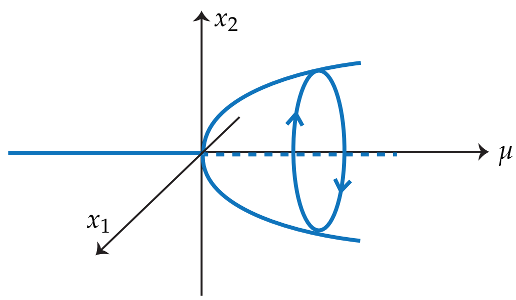
In polar coordinates:

$$\begin{aligned} \dot{r} &= \mu r - r^3 \\ \dot{\theta} &= 1 \end{aligned}$$

Note that a positive equilibrium for the r subsystem means a limit cycle in the (x_1, x_2) plane.

$\mu < 0$: stable equilibrium at $r = 0$

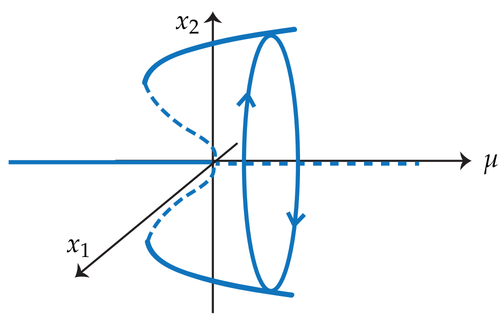
$\mu > 0$: unstable equilibrium at $r = 0$ and stable limit cycle at $r = \sqrt{\mu}$



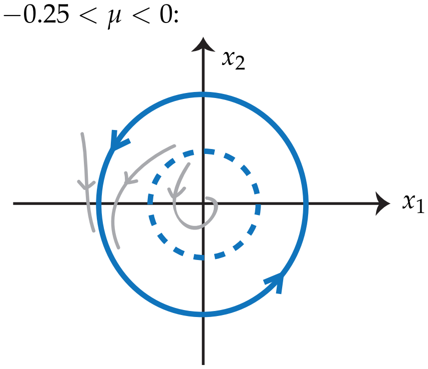
The origin loses stability at $\mu = 0$ and a stable limit cycle emerges.

Example: Subcritical Hopf bifurcation

$$\begin{aligned} \dot{r} &= \mu r + r^3 - r^5 \\ \dot{\theta} &= 1 \end{aligned}$$



Phase portrait for $-0.25 < \mu < 0$:



Characteristic of the Hopf bifurcation:

$$\boxed{\left.\frac{\partial f}{\partial x}\right|_{\mu=\mu^c,\, x=x^*(\mu^c)} \text{ has complex conjugate eigenvalues on the imaginary axis.}}$$

EE C222/ME C237 - Spring'18 - Lecture 5 Notes[1]

Murat Arcak

January 31 2018

Center Manifold Theory

Khalil (Section 8.1), Sastry (Section 7.6.1)

$$\dot{x} = f(x) \quad f(0) = 0 \tag{1}$$

Suppose $A \triangleq \left.\dfrac{\partial f}{\partial x}\right|_{x=0}$ has k eigenvalues will zero real parts, and $m = n - k$ eigenvalues with negative real parts.

Define $\begin{bmatrix} y \\ z \end{bmatrix} = Tx$ such that

$$TAT^{-1} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

where the eigenvalues of A_1 have zero real parts and the eigenvalues of A_2 have negative real parts.

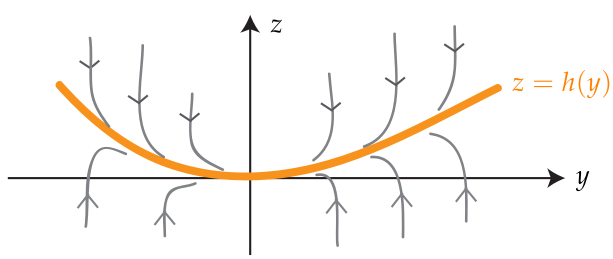
Rewrite $\dot{x} = f(x)$ in the new coordinates:

$$\begin{aligned} \dot{y} &= A_1 y + g_1(y,z) \\ \dot{z} &= A_2 z + g_2(y,z) \end{aligned} \tag{2}$$

$g_i(0,0) = 0$, $\frac{\partial g_i}{\partial y}(0,0) = 0$, $\frac{\partial g_i}{\partial z}(0,0) = 0$, $i = 1,2$.

Theorem 1: There exists an invariant manifold $z = h(y)$ defined in a neighborhood of the origin such that

$$h(0) = 0 \quad \frac{\partial h}{\partial y}(0) = 0.$$



$$\boxed{\text{Reduced System: } \dot{y} = A_1 y + g_1(y, h(y)) \quad y \in \mathbb{R}^k}$$

Theorem 2: If $y = 0$ is asymptotically stable (resp., unstable) for the reduced system, then $x = 0$ is asymptotically stable (resp., unstable) for the full system $\dot{x} = f(x)$.

Characterizing the Center Manifold

Define $w \triangleq z - h(y)$ and note that it satisfies

$$\dot{w} = A_2 z + g_2(y,z) - \frac{\partial h}{\partial y}\Big(A_1 y + g_1(y,z)\Big).$$

The invariance of $z = h(y)$ means that $w = 0$ implies $\dot{w} = 0$. Thus, the expression above must vanish when we substitute $z = h(y)$:

$$\boxed{A_2 h(y) + g_2(y, h(y)) - \frac{\partial h}{\partial y}\Big(A_1 y + g_1(y, h(y))\Big) = 0.}$$

To find $h(y)$ solve this differential equation for h as a function on y.

If the exact solution is unavailable, an approximation is possible. For scalar y, expand $h(y)$ as

$$h(y) = h_2 y^2 + \cdots + h_p y^p + O(y^{p+1})$$

where $h_1 = h_0 = 0$ because $h(0) = \frac{\partial h}{\partial y}(0) = 0$. The notation $O(y^{p+1})$ refers to the higher order terms of power $p+1$ and above.

Example:

$$\begin{aligned}\dot{y} &= yz \\ \dot{z} &= -z + ay^2 \quad a \neq 0\end{aligned}$$

This is of the form (2) with $g_1(y,z) = yz$, $g_2(y,z) = ay^2$, $A_2 = -1$. Thus $h(y)$ must satisfy

$$-h(y) + ay^2 - \frac{\partial h}{\partial y} y h(y) = 0.$$

Try $h(y) = h_2 y^2 + O(y^3)$:

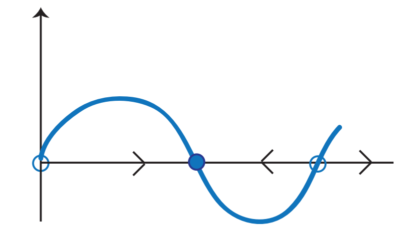
$$\begin{aligned}0 &= -h_2 y^2 + O(y^3) + ay^2 - (2h_2 y + O(y^2))y(h_2^2 + O(y^3)) \\ &= (a - h_2)y^2 + O(y^3) \\ &\implies h_2 = a\end{aligned}$$

Reduced System: $\dot{y} = y(ay^2 + O(y^3)) = ay^3 + O(y^4)$.

If $a < 0$, the full systems is asymptotically stable. If $a > 0$ unstable.

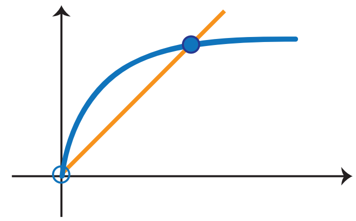
Discrete-Time Models and a Chaos Example

CT: $\dot{x}(t) = f(x(t))$
$f(x^*) = 0$



Asymptotic stability criterion:
$\Re\lambda_i(A) < 0$ where $A \triangleq \left.\frac{\partial f}{\partial x}\right|_{x=x^*}$
$f'(x^*) < 0$ for first order system

DT: $x_{n+1} = f(x_n)$ $n = 0, 1, 2, \ldots$
$f(x^*) = x^*$ ("fixed point")

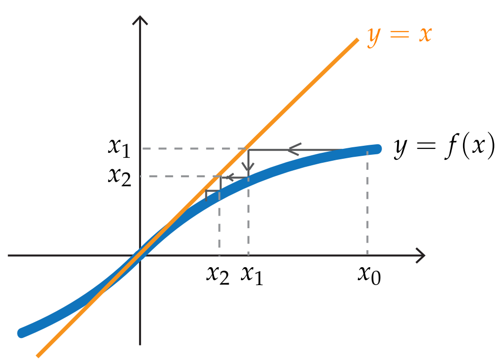


Asymptotic stability criterion:
$|\lambda_i(A)| < 1$ where $A \triangleq \left.\frac{\partial f}{\partial x}\right|_{x=x^*}$
$|f'(x^*)| < 1$ for first order system

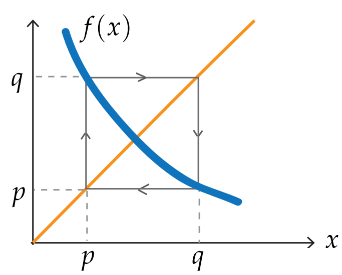
These criteria are inconclusive if the respective inequality is not strict, but for first order systems we can determine stability graphically:

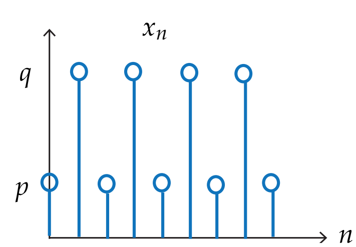
Cobweb Diagrams for First Order Discrete-Time Systems

Example: $x_{n+1} = \sin(x_n)$ has unique fixed point at 0. Stability test above inconclusive since $f'(0) = 1$. However, the "cobweb" diagram below illustrates the convergence of iterations to 0:



In discrete time, even first order systems can exhibit oscillations:





Detecting Cycles Analytically

$$f(p) = q \quad f(q) = p \quad \Longrightarrow \quad f(f(p)) = p \quad f(f(q)) = q$$

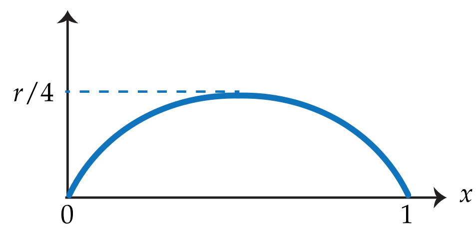
For the existence of a period-2 cycle, the map $f(f(\cdot))$ must have two fixed points in addition to the fixed points of $f(\cdot)$.

Period-3 cycles: fixed points of $f(f(f(\cdot)))$.

Chaos in a Discrete Time Logistic Growth Model

$$x_{n+1} = r(1 - x_n)x_n \tag{3}$$

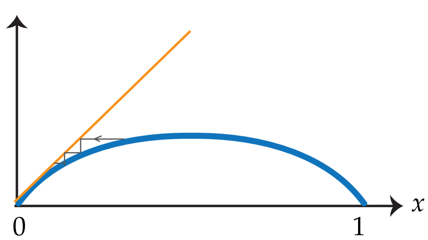
Range of interest: $0 \leq x \leq 1 \quad (x_n > 1 \;\Rightarrow\; x_{n+1} < 0)$



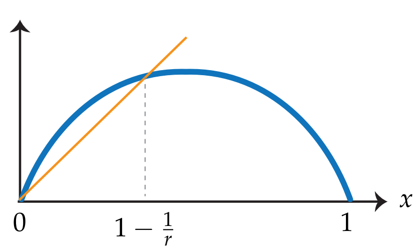
We will study the range $0 \leq r \leq 4$ so that $f(x) = r(1-x)x$ maps $[0,1]$ onto itself.

Fixed points: $x = r(1-x)x \;\Rightarrow\; \begin{cases} x^* = 0 \quad \text{and} \\ x^* = 1 - \frac{1}{r} \text{ if } r > 1. \end{cases}$

$\underline{r \leq 1}$: $x^* = 0$ unique and stable fixed point



$\underline{r > 1}$: $x = 0$ unstable because $f'(0) = r > 1$



Note that a transcritical bifurcation occurred at $r = 1$, creating the new equilibrium

$$x^* = 1 - \frac{1}{r}.$$

Evaluate its stability using $f'(x^*) = r(1-2x^*) = 2 - r$.

$$r < 3 \;\Rightarrow\; |f'(x^*)| < 1 \text{ (stable)}$$
$$r > 3 \;\Rightarrow\; |f'(x^*)| > 1 \text{ (unstable)}.$$

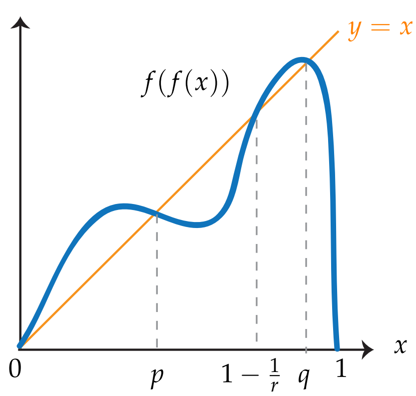
At $r = 3$, a period-2 cycle is born:

$$\begin{aligned} x &= f(f(x)) \\ &= r(1 - f(x))f(x) \\ &= r(1 - r(1-x)x)r(1-x)x \\ &= r^2x(1-x)(1 - r + rx - rx^2) \end{aligned}$$

$$0 = r^2x(1-x)(1 - r + rx - rx^2) - x$$

Factor out x and $(x - 1 + \frac{1}{r})$, find the roots of the quotient:

$$p, q = \frac{r + 1 \mp \sqrt{(r-3)(r+1)}}{2r}$$

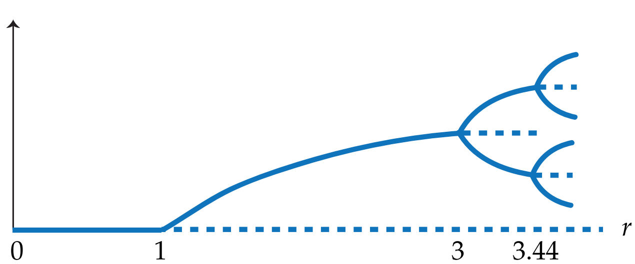


This period-2 cycle is stable when $r < 1 + \sqrt{6} = 3.4494$:

$$\left.\frac{d}{dx} f(f(x))\right|_{x=p} = f'(f(p))f'(p) = f'(p)f'(q) = 4 + 2r - r^2$$

$$|4 + 2r - r^2| < 1 \;\Rightarrow\; 3 < r < 1 + \sqrt{6} = 3.4494$$

At $r = 3.4494$, a period-4 cycle is born!

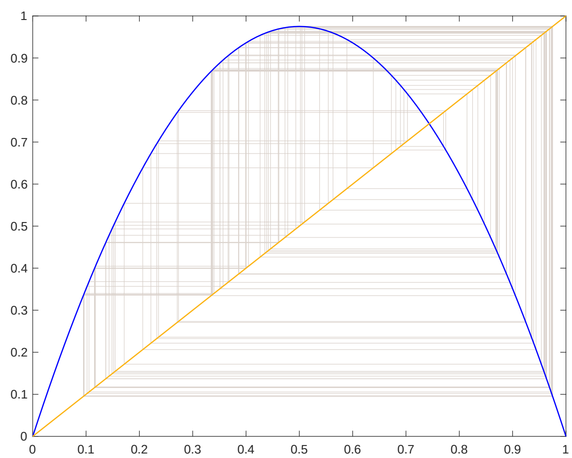


"period doubling bifurcations"

$r_1 = 3$	period-2 cycle born
$r_2 = 3.4494$	period-4 cycle born
$r_3 = 3.544$	period-8 cycle born
$r_4 = 3.564$	period-16 cycle born
$\vdots$	
$r_\infty = 3.5699$	

After $r > r_\infty$, chaotic behavior for a window of r, followed by windows of periodic behavior (*e.g.*, period-3 cycle around $r = 3.83$).

Below is the cobweb diagram for $r = 3.9$ which is in the chaotic regime:



EE C222/ME C237 - Spring'18 - Lecture 6 Notes[1]

Murat Arcak

February 5 2018

Mathematical Background

Sastry, Chapter 3

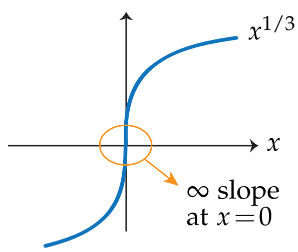
$$\dot{x} = f(x) \quad x(0) = x_0 \tag{1}$$

Do solutions exist? Are they unique?

- If $f(\cdot)$ is continuous (C^0) then a solution exists, but C^0 is not sufficient for uniqueness.

  Example: $\dot{x} = x^{\frac{1}{3}}$ with $x(0) = 0$

  $$x(t) \equiv 0,\ x(t) = \left(\frac{2}{3}t\right)^{\frac{3}{2}} \text{ are both solutions}$$

- Sufficient condition for uniqueness: "Lipschitz continuity" (more restrictive than C^0)

$$|f(x) - f(y)| \leq L|x - y| \tag{2}$$

Definition: $f(\cdot)$ is *locally Lipschitz* if every point x^0 has a neighborhood where (2) holds for all x, y in this neighborhood and for all t for some L.

Example: $(\cdot)^{\frac{1}{3}}$ is NOT locally Lipschitz (due to ∞ slope)

$(\cdot)^3$ *is* locally Lipschitz:

$$x^3 - y^3 = \underbrace{(x^2 + xy + y^2)}_{\substack{\text{in any nbhd} \\ \text{of } x^0\text{, we can} \\ \text{find } L \text{ to upper} \\ \text{bound this}}} (x - y)$$

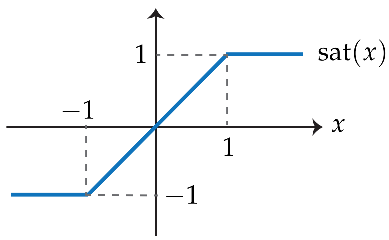
$$\Longrightarrow |x^3 - y^3| \leq L|x - y|$$

- If $f(\cdot)$ is continuously differentiable (C^1), then it is locally Lipschitz.

Examples: x^3, x^2, e^x, etc.

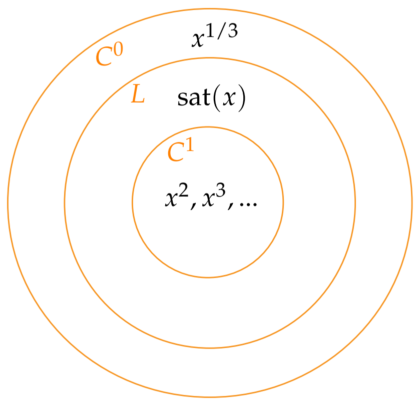
The converse is not true: local Lipschitz $\not\Rightarrow C^1$

Example:



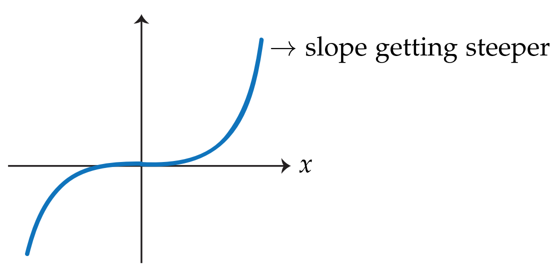
Not differentiable at $x = \mp 1$, but locally Lipschitz:

$$|\operatorname{sat}(x) - \operatorname{sat}(y)| \leq |x - y| \qquad (L = 1).$$



Definition continued: $f(\cdot)$ is *globally Lipschitz* if (2) holds $\forall x, y \in \mathbb{R}^n$ (*i.e.*, the same L works everywhere).

Examples: $\operatorname{sat}(\cdot)$ is globally Lipschitz. $(\cdot)^3$ is not globally Lipschitz:



- Suppose $f(\cdot)$ is C^1. Then it is globally Lipschitz iff $\frac{\partial f}{\partial x}$ is bounded.

$$L = \sup_x |f'(x)|$$

Preview of existence theorems:

1. $f(\cdot)$ is $C^0 \implies$ existence of solution $x(t)$ on finite interval $[0, t_f)$.
2. $f(\cdot)$ locally Lipschitz $\implies$ existence and uniqueness on $[0, t_f)$.
3. $f(\cdot)$ globally Lipschitz $\implies$ existence and uniqueness on $[0, \infty)$.

Examples:

- $\dot{x} = x^2$ (locally Lipschitz) admits unique solution on $[0, t_f)$, but $t_f < \infty$ from Lecture 1 (finite escape).
- $\dot{x} = Ax$ globally Lipschitz, therefore no finite escape

$$|Ax - Ay| \leq L|x - y| \quad \text{with} \quad L = \|A\|$$

The rest of the lecture introduces concepts that are used in proving the existence theorems mentioned above.

Normed Linear Spaces

Definition: $\mathbb{X}$ is a normed linear space if there exists a real-valued norm $|\cdot|$ satisfying:

1. $|x| \geq 0 \ \ \forall x \in \mathbb{X}, \ \ |x| = 0$ iff $x = 0$.
2. $|x + y| \leq |x| + |y| \ \ \forall x, y \in \mathbb{X}$ (triangle inequality)
3. $|\alpha x| = |\alpha| \cdot |x| \ \ \forall \alpha \in \mathbb{R}$ and $x \in \mathbb{X}$.

Definition: A sequence $\{x_k\}$ in $\mathbb{X}$ is said to be a Cauchy sequence if

$$|x_k - x_m| \to 0 \text{ as } k, m \to \infty. \tag{3}$$

Every convergent sequence is Cauchy. The converse is not true.

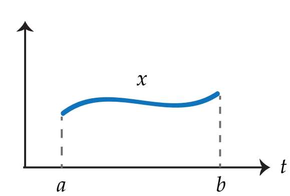
Definition: $\mathbb{X}$ is a Banach space if every Cauchy sequence converges to an element in X.

All Euclidean spaces are Banach spaces.

Example:

$C^n[a, b]$: the set of all continuous functions $[a, b] \to \mathbb{R}^n$ with norm:

$$|x|_C = \max_{t \in [a,b]} |x(t)|$$



1. $|x|_C \geq 0$ and $|x|_C = 0$ iff $x(t) \equiv 0$.

2. $|x+y|_C = \max_{t\in[a,b]} |x(t)+y(t)| \le \max_{t\in[a,b]} \{|x(t)|+|y(t)|\} \le |x|_C + |y|_C$

3. $|\alpha \cdot x|_C = \max_{t\in[a,b]} |\alpha| \cdot |x(t)| = |\alpha| \cdot |x|_C$

It can be shown that $C^n[a,b]$ is a Banach space.

Fixed Point Theorems

$$T(x) = x \tag{4}$$

Brouwer's Theorem (Euclidean spaces):

If U is a closed bounded subset of a Euclidean space and $T : U \to U$ is continuous, then T has a fixed point in U.

Schauder's Theorem (Brouwer's Thm $\to$ Banach spaces):

If U is a closed bounded convex subset of a Banach space $\mathbb{X}$ and $T : U \to U$ is *completely continuous*[2], then T has a fixed point in U.

[2] continuous and for any bounded set $B \subseteq U$ the closure of $T(B)$ is compact

Contraction Mapping Theorem:

If U is a closed subset of a Banach space and $T : U \to U$ is such that

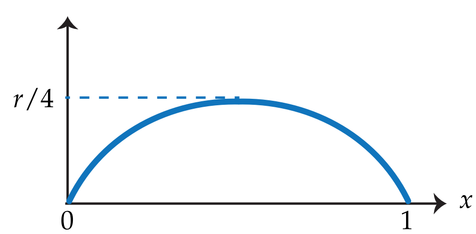
$$|T(x) - T(y)| \le \rho |x-y| \quad \rho < 1 \quad \forall x, y \in U$$

then T has a unique fixed point in U and the solutions of $x_{n+1} = T(x_n)$ converge to this fixed point from any $x_0 \in U$.

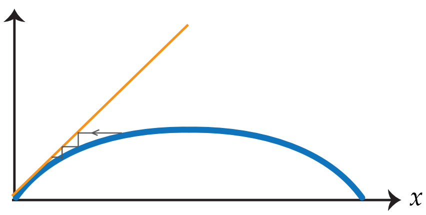
Example: The logistic map (Lecture 5)

$$T(x) = rx(1-x) \tag{5}$$

with $0 \le r \le 4$ maps $U = [0,1]$ to U. $|T'(x)| \le r \quad \forall x \in [0,1]$, so the contraction property holds with $\rho = r$.



If $r < 1$, the contraction mapping theorem predicts a unique fixed point that attracts all solutions starting in $[0,1]$.



Proof steps for the Contraction Mapping Thm:

1. Show that $\{x_n\}$ formed by $x_{n+1} = T(x_n)$ is a Cauchy sequence. Since we are in a Banach space, this implies a limit x^* exists.

2. Show that $x^* = T(x^*)$.

3. Show that x^* is unique.

Details of each step:

1.
$$\begin{aligned}|x_{n+1} - x_n| = |T(x_n) - T(x_{n-1})| &\leq \rho|x_n - x_{n-1}| \\ &\leq \rho^2|x_{n-1} - x_{n-2}| \\ &\vdots \\ &\leq \rho^n|x_1 - x_0|.\end{aligned}$$

$$\begin{aligned}|x_{n+r} - x_n| &\leq |x_{n+r} - x_{n+r-1}| + \cdots + |x_{n+1} - x_n| \\ &\leq (\rho^{n+r} + \cdots + \rho^n)|x_1 - x_0| \\ &= \rho^n(1 + \cdots + \rho^r)|x_1 - x_0| \\ &\leq \rho^n \frac{1}{1-\rho}|x_1 - x_0|\end{aligned}$$

Since $\frac{\rho^n}{1-\rho} \to 0$ as $n \to \infty$, we have $|x_{n+r} - x_n| \to 0$ as $n \to \infty$.

2.
$$\begin{aligned}|x^* - T(x^*)| &= |x^* - x_n + T(x_{n-1}) - T(x^*)| \\ &\leq |x^* - x_n| + |T(x_{n-1}) - T(x^*)| \\ &\leq |x^* - x_n| + \rho|x^* - x_{n-1}|.\end{aligned}$$

Since $\{x_n\}$ converges to x^*, we can make this upper bound arbitrarily small by choosing n sufficiently large. This means that $|x^* - T(x^*)| = 0$, hence $x^* = T(x^*)$.

3. Suppose $y^* = T(y^*)$ $y^* \neq x^*$.

$$|x^* - y^*| = |T(x^*) - T(y^*)| \leq \rho|x^* - y^*| \implies x^* = y^*.$$

Thus we have a contradiction.

EE C222/ME C237 - Spring'18 - Lecture 7 Notes[1]

Murat Arcak

February 7 2018

Existence and Uniqueness Theorems for ODEs

Khalil (Section 3.1), Sastry (Section 3.4)

$$\dot{x} = f(t,x) \quad x(0) = x_0 \tag{1}$$

Theorem 1: $f(t,x)$ locally Lipschitz in x and continuous in t

$\Rightarrow$ existence and uniqueness on some finite interval $[0,\delta]$.

Sketch of the proof: From the local Lipschitz assumption, we can find $r > 0$ and $L > 0$ such that

$$|f(t,x) - f(t,y)| \le L|x-y| \;\; \forall x,y \in \{x \in \mathbb{R}^n : |x - x_0| \le r\}.$$

If $x(t)$ is a solution, then:

$$x(t) = \underbrace{x_0 + \int_0^t f(\tau, x(\tau))d\tau.}_{=:\,T(x)(t)}$$

To apply the Contraction Mapping Theorem:

1. Choose δ small enough that T maps the following subset of $C^n[0,\delta]$ to itself :

$$U = \{x \in C^n[0,\delta] : |x(t) - x_0| \le r \;\forall t \in [0,\delta]\},$$

i.e.

$$|x(t) - x_0| \le r \quad \forall t \in [0,\delta] \quad \Rightarrow \quad |T(x)(t) - x_0| \le r \quad \forall t \in [0,\delta]. \tag{2}$$

To find such a δ note that

$$T(x)(t) - x_0 = \int_0^t f(\tau, x(\tau))d\tau = \int_0^t \Big(f(\tau,x(\tau)) - f(\tau,x_0) + f(\tau,x_0)\Big)d\tau$$

$$\begin{aligned}|T(x)(t) - x_0| &\le \int_0^\delta |f(\tau,x(\tau)) - f(\tau,x_0)|d\tau + \int_0^\delta |f(\tau,x_0)|d\tau \\ &\le \int_0^\delta L|x(\tau) - x_0|d\tau + \int_0^\delta h d\tau \quad \text{where } h \text{ is a bound on } |f(\tau,x_0)| \\ &\le (Lr+h)\delta.\end{aligned}$$

Thus, by choosing $\delta \le \frac{r}{Lr+h}$ we ensure that the implication (2) holds.

2. Show that T is a contraction in U, *i.e.*, there exists $\rho < 1$ s.t.

$$x,y \in U \implies |T(x) - T(y)|_C \le \rho|x-y|_C.$$

Note that, for all $t \in [0, \delta]$,

$$
\begin{aligned}
|T(x)(t) - T(y)(t)| &= \int_0^t |f(\tau, x(\tau)) - f(\tau, y(\tau))| d\tau \\
&\leq L \int_0^t |x(\tau) - y(\tau)| d\tau \\
&\leq \underbrace{L\delta}_{=:\rho} \max_{\tau \in [0,\delta]} |x(\tau) - y(\tau)| = \rho |x - y|_C.
\end{aligned}
$$

Therefore,

$$
|T(x) - T(y)|_C = \max_{t \in [0,\delta]} |T(x)(t) - T(y)(t)| \leq \rho |x - y|_C
$$

and $\rho < 1$ if $\delta \leq \frac{r}{Lr+h}$ as prescribed above.

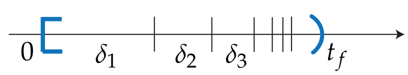
Theorem 2: $f(t, x)$ *globally* Lipschitz in x uniformly[2] in t, and continuous in t $\Longrightarrow$ existence and uniqueness on $[0, \infty)$.

[2] same L works for all t

Proof: Choose a δ that doesn't depend on x_0 and apply Theorem 1 repeatedly to cover $[0, \infty)$. This is possible because L works everywhere and we can pick r as large as we wish. Indeed, for any $\delta < \frac{1}{L}$, we can choose r large enough that $\delta \leq \frac{r}{Lr+h}$.

Q: Why can't we do this in Theorem 1?

A: δ depends on x_0 (no universal L) and x_0 changes at the next iteration. We can't use the same δ in every iteration:



- The theorems above are sufficient only, and can be conservative:

Example: $\dot{x} = -x^3$ is not globally Lipschitz but

$$
x(t) = \text{sgn}(x_0) \sqrt{\frac{x_0^2}{1 + 2tx_0^2}}
$$

is defined on $[0, \infty)$.

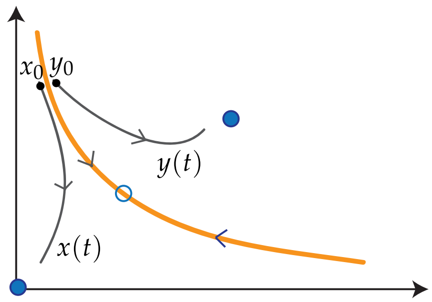
Continuous Dependence on Initial Conditions and Parameters

Theorem 3: (Continuous dependence on initial conditions) Let $x(t), y(t)$ be two solutions of $\dot{x} = f(t, x)$ starting from x_0 and y_0, and remaining in a set with Lipschitz constant L on $[0, \tau]$. Then, for any $\epsilon > 0$, there exists $\delta(\epsilon, \tau) > 0$ such that

$$
|x_0 - y_0| \leq \delta \Longrightarrow |x(t) - y(t)| \leq \epsilon \quad \forall t \in [0, \tau].
$$

- This conclusion does not hold on infinite time intervals (even if f is globally Lipschitz).

Example: bistable system



If ϵ is smaller than the distance between the two stable equilibria, no choice of δ guarantees $|x(t) - y(t)| \leq \epsilon \ \ \forall t \geq 0$.

- Theorem 3 also shows continuous dependence on parameter μ in $f(t, x, \mu)$ if we rewrite the system equations as:

$$\begin{aligned} \dot{x} &= f(t, x, \mu) \\ \dot{\mu} &= 0 \end{aligned} \qquad X = \begin{bmatrix} x \\ \mu \end{bmatrix} \qquad \dot{X} = F(t, X) \triangleq \begin{bmatrix} f(t, x, \mu) \\ 0 \end{bmatrix},$$

  where μ appears as a state variable with initial condition $\mu(0) = \mu$.

Q: How do you reconcile bifurcations with continuous dependence on parameters? We could pick two values of the bifurcation parameter arbitrarily close, but one below and one above the critical value, thereby expecting a drastic difference in the solutions.

A: The two solutions are close in the short term (Theorem 3 holds on finite time intervals); the drastic difference builds up over time.

Sensitivity to Parameters

Consider the system

$$\dot{x} = f(t, x, \mu) \quad x \in \mathbb{R}^n, \mu \in \mathbb{R}^p \tag{3}$$

where μ is a vector of p parameters, and let $\phi(t, x_0, \mu)$ denote the trajectories starting at the initial condition x_0.

To determine to what extent this trajectory depends on the parameters we define the $n \times p$ *sensitivity matrix*:

$$S(t, x_0, \mu) := \frac{\partial \phi(t, x_0, \mu)}{\partial \mu} = \left[\frac{\partial \phi(t, x_0, \mu)}{\partial \mu_1} \cdots \frac{\partial \phi(t, x_0, \mu)}{\partial \mu_p} \right], \tag{4}$$

where each column is the sensitivity with respect to a particular parameter.

To see how $S(t, x_0, \mu)$ can be computed numerically, first note that $\phi(t, x_0, \mu)$ satisfies the equation (3), that is,

$$\frac{\partial \phi(t, x_0, \mu)}{\partial t} = f(t, \phi(t, x_0, \mu), \mu).$$

Next, differentiate both sides with respect to μ:

$$\frac{\partial^2 \phi(t, x_0, \mu)}{\partial t \partial \mu} = \frac{\partial f}{\partial x}(t, \phi(t, x_0, \mu), \mu) \frac{\partial \phi(t, x_0, \mu)}{\partial \mu} + \frac{\partial f}{\partial \mu}(t, \phi(t, x_0, \mu), \mu)$$

and use the definition of the sensitivity matrix to rewrite this as

$$\frac{\partial S(t, x_0, \mu)}{\partial t} = \frac{\partial f}{\partial x}(t, \phi(t, x_0, \mu), \mu) S(t, x_0, \mu) + \frac{\partial f}{\partial \mu}(t, \phi(t, x_0, \mu), \mu).$$

Thus, S can be computed by numerical integration of (3) simultaneously with

$$\dot{S} = \frac{\partial f}{\partial x}(t, x, \mu) S + \frac{\partial f}{\partial \mu}(t, x, \mu).$$

The initial condition for S is $\frac{\partial x_0}{\partial \mu} = 0$, assuming that x_0 is independent of the parameters.

Example: For the harmonic oscillator

$$\begin{aligned} \dot{x}_1 &= -\mu x_2 \\ \dot{x}_2 &= \mu x_1 \end{aligned}$$

we have

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 0 & -\mu \\ \mu & 0 \end{bmatrix} \qquad \frac{\partial f}{\partial \mu} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

Thus the sensitivity equation is

$$\dot{S} = \begin{bmatrix} 0 & -\mu \\ \mu & 0 \end{bmatrix} S + \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$

Logarithmic Sensitivity

To compare the sensitivity with respect to multiple parameters $\mu_1, \ldots, \mu_p$ it is preferable to use the *logarithmic sensitivity*

$$\frac{\partial \phi(t, x_0, \mu)}{\partial \mu_i / \mu_i} = \frac{\partial \phi(t, x_0, \mu)}{\partial \ln \mu_i}$$

so that the denominator is dimensionless and represents the change in the parameter μ_i relative to its nominal value. This means that the ith column of the sensitivity matrix S in (4) must be multiplied with the nominal parameter μ_i, $i = 1, \ldots, p$, before these columns are compared for the relative significance of the parameters.

Application to Parameter Tuning and Identification

Sensitivity equations are useful for solving a class of optimization problems of the form

$$\begin{aligned} &\min_\mu J(\mu) = \textstyle\int_{t_0}^{t_1} q(t, x(t), \mu) dt \\ &\text{subject to } \dot{x} = f(t, x, \mu), \; x(t_0) = x_0. \end{aligned}$$

For example, one may take $q(t, x) = |h(x(t)) - r(t)|^2$ to penalize the error between the output $y(t) = h(x(t))$ of a control system and a reference trajectory $r(t)$ to be followed. In this example $\dot{x} = f(t, x, \mu)$ represents the closed loop model with tunable control parameters μ.

In other applications $\dot{x} = f(t, x, \mu)$ may represent the model of a physical process with unknown parameters, and $q(t, x) = |h(x(t)) - r(t)|^2$ penalizes the error between the model prediction for a variable, $y(t) = h(x(t))$, and the experimental observation $r(t)$. Then the optimization problem above aims to find parameters that best fit the experimental data.

A typical optimization algorithm requires the gradient $\frac{\partial J(\mu)}{\partial \mu}$, which can be obtained with the help of the chain rule and the sensitivity equations:

$$\frac{\partial J(\mu)}{\partial \mu} = \int_{t_0}^{t_1} \left(\frac{\partial q}{\partial x}(t, x(t), \mu) S(t, x_0, \mu) + \frac{\partial q}{\partial \mu}(t, x(t), \mu) \right) dt.$$

EE C222/ME C237 - Spring'18 - Lecture 8 Notes[1]

Murat Arcak

February 12 2018

Lyapunov Stability Theory

Khalil Chapter 4, Sastry Chapter 5

Consider a time invariant system

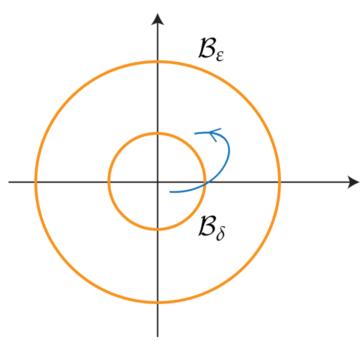
$$\dot{x} = f(x)$$

and assume equilibrium at $x = 0$, *i.e.* $f(0) = 0$. If the equilibrium of interest is $x^* \neq 0$, let $\tilde{x} = x - x^*$:

$$\dot{\tilde{x}} = f(x) = f(\tilde{x} + x^*) \triangleq \tilde{f}(\tilde{x}) \Longrightarrow \tilde{f}(0) = 0.$$

Definition: The equilibrium $x = 0$ is stable if for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|x(0)| \leq \delta \implies |x(t)| \leq \varepsilon \quad \forall t \geq 0. \tag{1}$$

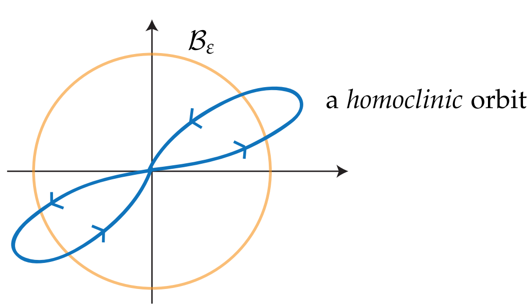


It is unstable if not stable.

Asymptotically stable if stable and $x(t) \to 0$ for all $x(0)$ in a neighborhood of $x = 0$.

Globally asymptotically stable if stable and $x(t) \to 0$ for every $x(0)$.

Note that $x(t) \to 0$ does not necessarily imply stability: one can construct an example where trajectories converge to the origin, but only after a large detour that violates the stability definition.



Lyapunov's Stability Theorem

Aleksandr Lyapunov (1857-1918)

1. Let D be an open, connected subset of $\mathbb{R}^n$ that includes $x = 0$. If there exists a C^1 function $V : D \to \mathbb{R}$ such that

$$V(0) = 0 \text{ and } V(x) > 0 \quad \forall x \in D - \{0\} \qquad \text{(positive definite)}$$

   and

$$\dot{V}(x) := \nabla V(x)^T f(x) \leq 0 \quad \forall x \in D \qquad \text{(negative semidefinite)}$$

   then $x = 0$ is stable.

2. If $\dot{V}(x) < 0 \quad \forall x \in D - \{0\}$ (negative definite) then $x = 0$ is asymptotically stable.

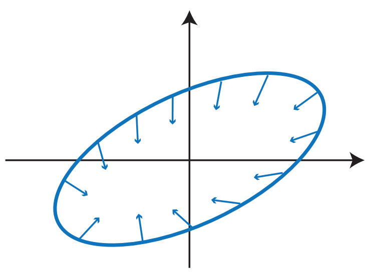
3. If, in addition, $D = \mathbb{R}^n$ and

$$|x| \to \infty \implies V(x) \to \infty \qquad \text{(radially unbounded)}$$

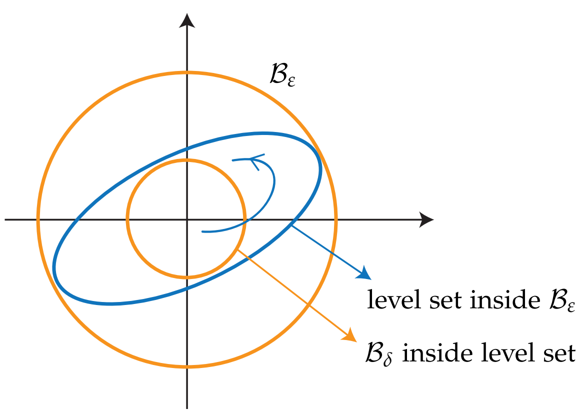
   then $x = 0$ is globally asymptotically stable.

<u>Sketch of the proof:</u>

The sets $\Omega_c \triangleq \{x : V(x) \leq c\}$ for constants c are called *level sets* of V and are positively invariant because $\nabla V(x)^T f(x) \leq 0$.



Stability follows from this property: choose a level set inside the ball of radius ε, and a ball of radius δ inside this level set. Trajectories starting in $\mathcal{B}_\delta$ can't leave $\mathcal{B}_\varepsilon$ since they remain inside the level set.

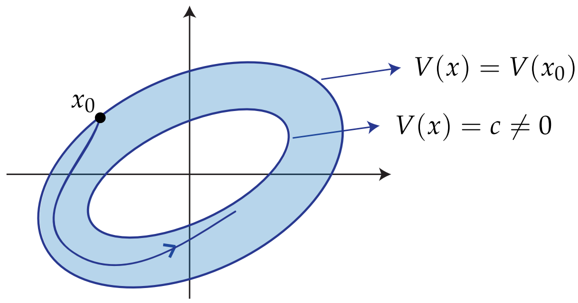


Asymptotic stability:

Since $V(x(t))$ is decreasing and bounded below by 0, we conclude

$$V(x(t)) \to c \geq 0.$$

We will show $c = 0$ (*i.e.*, $x(t) \to 0$) by contradiction. Suppose $c \neq 0$:



Let

$$\gamma \triangleq \max_{\{x:\, c \leq V(x) \leq V(x_0)\}} -\dot{V}(x) \;> 0$$

where the maximum exists because it is evaluated over a bounded[2] set, and is positive because $\dot{V}(x) < 0$ away from $x = 0$. Then,

$$\dot{V}(x) \leq -\gamma \implies V(x(t)) \leq V(x_0) - \gamma t,$$

which implies $V(x(t)) < 0$ for $t > \frac{V(x_0)}{\gamma}$ – a contradiction because $V \geq 0$. Therefore, $c = 0$ which implies $x(t) \to 0$.

[2] By positive definiteness of V, the level sets $\{x : V(x) \leq \text{constant}\}$ are bounded when the constant is sufficiently small. Since we are proving *local* asymptotic stability we can assume x_0 is close enough to the origin that the constant $V(x_0)$ is sufficiently small.

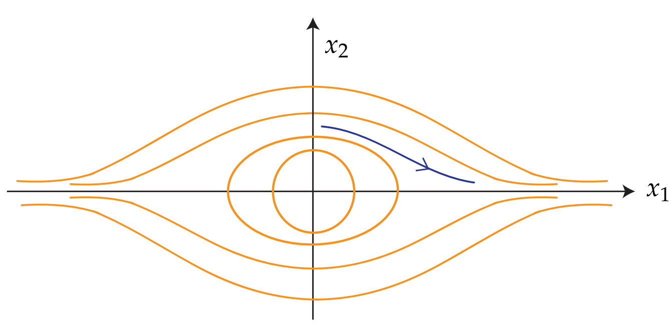
Global asymptotic stability:

Why do we need radial unboundedness?

Example:

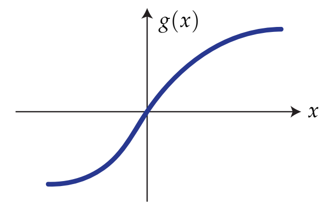
$$V(x) = \frac{x_1^2}{1 + x_1^2} + x_2^2 \tag{2}$$

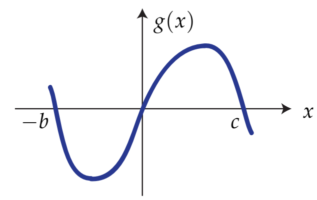
Set $x_2 = 0$, let $x_1 \to \infty$: $V(x) \to 1$ (not radially unbounded). Then Ω_c is not a bounded set for $c \geq 1$:



Therefore, $x_1(t)$ may grow unbounded while $V(x(t))$ is decreasing.

Finding Lyapunov Functions





Example:

$$\dot{x} = -g(x) \quad x \in \mathbb{R},\ xg(x) > 0 \ \ \forall x \neq 0 \tag{3}$$

$V(x) = \frac{1}{2}x^2$ is positive definite and radially unbounded.

$\dot{V}(x) = -xg(x)$ is negative definite. Therefore $x = 0$ is globally asymptotically stable.

If $xg(x) > 0$ only in $(-b, c) - \{0\}$, then take $D = (-b, c)$

$\implies x = 0$ is locally asymptotically stable.

There are other equilibria where $g(x) = 0$, so we know global asymptotic stability is not possible.

Example:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -ax_2 - g(x_1) \quad a \geq 0,\ xg(x) > 0 \ \ \forall x \in (-b, c) - \{0\} \end{aligned} \tag{4}$$

The pendulum is a special case with $g(x) = \sin(x)$.

The choice $V(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$ doesn't work because $\dot{V}(x)$ is sign indefinite (show this).

The function

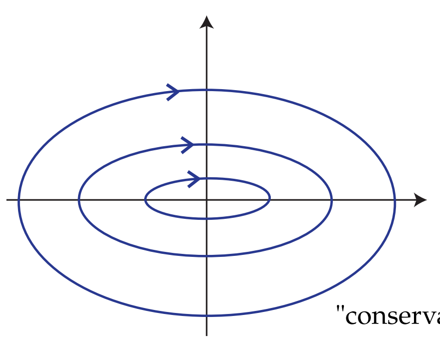
$$V(x) = \int_0^{x_1} g(y)dy + \frac{1}{2}x_2^2$$

is positive definite on $D = (-b, c) - \{0\}$ and

$$\dot{V}(x) = g(x_1)x_2 - ax_2^2 - x_2 g(x_1) = -ax_2^2$$

is negative semidefinite $\implies$ stable.

If $a = 0$, no asymptotic stability because $\dot{V}(x) = 0 \implies V(x(t)) = V(x(0))$.



"conservative system"

If $a > 0$, (4) is asymptotically stable but the Lyapunov function above doesn't allow us to reach that conclusion. We need either another V with negative definite $\dot{V}$, or the Lasalle-Krasovskii Invariance Principle to be discussed in the next lecture.

EE C222/ME C237 - Spring'18 - Lecture 9 Notes[1]

Murat Arcak

February 14 2018

LaSalle-Krasovskii Invariance Principle

- Applicable to time-invariant systems.
- Allows us to conclude asymptotic stability from $\dot{V}(x) \leq 0$ if additional conditions hold:

Suppose $\Omega_c = \{x : V(x) \leq c\}$ is bounded and $\dot{V}(x) \leq 0$ in Ω_c. Define $S = \{x \in \Omega_c : \dot{V}(x) = 0\}$ and let M be the largest invariant set in S. Then, for every $x(0) \in \Omega_c$, $x(t) \to M$.

Corollary: If no solution other than $x(t) \equiv 0$ can stay identically in S then $M = \{0\}$ and we conclude asymptotic stability.

Example (from last lecture):

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -ax_2 - g(x_1) \quad a > 0,\ xg(x) > 0 \ \ \forall x \neq 0 \end{aligned} \tag{1}$$

$$V(x) = \int_0^{x_1} g(y)dy + \frac{1}{2}x_2^2 \quad \Longrightarrow \quad \dot{V}(x) = -ax_2^2$$

$$S = \{x \in \Omega_c | x_2 = 0\}$$

If $x(t)$ stays identically in S, then $x_2(t) \equiv 0 \Longrightarrow \dot{x}_2(t) \equiv 0 \Longrightarrow g(x_1(t)) \equiv 0 \Longrightarrow x_1(t) \equiv 0 \Longrightarrow$ asymptotic stability from Corollary.

Example (linear system): Same system above with $g(x_1) = bx_1$:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -ax_2 - bx_1 \quad a > 0, b > 0 \end{aligned} \tag{2}$$

$V(x) = \frac{b}{2}x_1^2 + \frac{1}{2}x_2^2 \Longrightarrow \dot{V}(x) = -ax_2^2 \Longrightarrow$ Invariance Principle works as in the example above.

Alternatively, construct another Lyapunov function with negative definite $\dot{V}(x)$. Try $V(x) = x^T P x$ where $P = P^T > 0$ is to be selected.

$$\dot{V}(x) = x^T P \dot{x} + \dot{P} x = x^T (A^T P + PA) x \text{ where } A = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix}$$

Let $P = \frac{1}{2}\begin{bmatrix} b & \epsilon \\ \epsilon & 1 \end{bmatrix}$, that is $V(x) = \frac{b}{2}x_1^2 + \epsilon x_1 x_2 + \frac{1}{2}x_2^2$.
Note that $P > 0$ if $\epsilon^2 < b$.

$$A^TP + PA = \begin{bmatrix} -\epsilon b & -\epsilon a/2 \\ -\epsilon a/2 & \epsilon - a \end{bmatrix} \begin{array}{l} \leq 0 \text{ if } \epsilon = 0 \\ < 0 \text{ if } \underbrace{0 < \epsilon < a \text{ and } \epsilon b(a-\epsilon) > \frac{\epsilon^2 a^2}{4}}_{0<\epsilon<\frac{ba}{b+\frac{a^2}{4}}} \end{array}$$

Linear Systems

Sastry (Sec. 5.7-5.8), Khalil (Sec. 4.3)

$$\dot{x} = Ax \quad x \in \mathbb{R}^n \tag{3}$$

$x = 0$ is stable if $\Re\{\lambda_i(A)\} \leq 0$ for all $i = 1, \cdots, n$ and eigenvalues on the imaginary axis have Jordan blocks of order one.[2] It is asymptotically stable if $\Re\{\lambda_i(A)\} < 0$ for all i, *i.e.*, A is "Hurwitz."

[2] *i.e.*, if λ is an eigenvalue of multiplicity q then $\lambda I - A$ must have rank $n - q$.

Example:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ is unstable: } \quad \left.\begin{array}{l} \dot{x}_1 = x_2 \\ \dot{x}_2 = 0 \end{array}\right\} x_1(t) = x_1(0) + x_2(0)t$$

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ is stable.}$$

Lyapunov Functions for Linear Systems

$$\begin{aligned} V(x) &= x^T P x \qquad P = P^T > 0 \\ \dot{V}(x) &= x^T (A^T P + PA) x \end{aligned} \tag{4}$$

If $\exists P = P^T > 0$ such that $A^TP + PA = -Q < 0$, then A is Hurwitz. The converse is also true:

Theorem: A is Hurwitz if and only if for any $Q = Q^T > 0$, there exists $P = P^T > 0$ such that

$$A^TP + PA = -Q. \tag{5}$$

Moreover, the solution P is unique.

(5) is known as the Lyapunov Equation. The Matlab command `lyap(A',Q)` returns the solution P.

Proof:

(if) From (4) above, the Lyapunov function $V(x) = x^T P x$ proves asymptotic stability which means A is Hurwitz.

(only if) Assume $\Re\{\lambda_i(A)\} < 0\ \forall i$. Show $\exists P = P^T > 0$ such that $A^TP + PA = -Q$.

Candidate:

$$P = \int_0^\infty e^{A^T t} Q e^{At} dt. \tag{6}$$

- The integral exists because $\|e^{At}\| \leq \kappa e^{-\alpha t}$.
- $P = P^T$
- $P > 0$ because $x^T P x = \int_0^\infty (e^{At}x)^T Q \underbrace{(e^{At}x)}_{\triangleq \phi(t,x)} dt \geq 0$ and

  $x^T P x = 0 \Longrightarrow \phi(t,x) \equiv 0 \Longrightarrow x = 0$ because e^{At} is nonsingular.
- $A^T P + PA = \int_0^\infty \underbrace{\left(A^T e^{A^T t} Q e^{At} + e^{A^T t} Q e^{At} A\right)}_{= \frac{d}{dt}\left(e^{At} Q e^{At}\right)} dt$

  $= e^{A^T t} Q e^{At}\Big|_0^\infty = 0 - Q = -Q$

Uniqueness:

Suppose there is another $\hat{P} = \hat{P}^T > 0$ satisfying $\hat{P} \neq P$, and $A^T \hat{P} + \hat{P} A = -Q$.

$\Longrightarrow (P - \hat{P})A + A^T(P - \hat{P}) = 0$

Define $W(x) = x^T (P - \hat{P}) x$.

$\frac{d}{dt} W(x(t)) = 0 \Longrightarrow W(x(t)) = W(x(0)) \quad \forall t.$

Since A is Hurwitz, $x(t) \to 0$ and $W(x(t)) \to 0$.

Combining the two statements above, we conclude $W(x(0)) = 0$ for any $x(0)$. This is possible only if $P - \hat{P} = 0$ which contradicts $\hat{P} \neq P$.

Invariance Principle Applied to Linear Systems

$$A^T P + PA = -Q \leq 0 \tag{7}$$

Can we conclude that A is Hurwitz if Q is only semidefinite?

Decompose Q as $Q = C^T C$ where $C \in \mathbb{R}^{r \times n}$, r is the rank of Q.

$$\dot{V}(x) = -x^T Q x = -x^T C^T C x = -y^T y$$

where $y \triangleq Cx$. The invariance principle guarantees asymptotic stability if

$$y(t) = Cx(t) \equiv 0 \implies x(t) \equiv 0.$$

This implications is true if the pair (C, A) is observable.

<u>Example</u> (beginning of the lecture):

$$A = \begin{bmatrix} 0 & 1 \\ -b & a \end{bmatrix} \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix} \quad \Longrightarrow \quad C = [0 \quad \sqrt{a}]$$

(C, A) is observable if $b \neq 0$.

EE C222/ME C237 - Spring'18 - Lecture 10 Notes[1]

Murat Arcak

February 21 2018

Lyapunov's Linearization Method

$$\dot{x} = f(x) \quad f(0) = 0$$

Define $A = \frac{\partial f(x)}{\partial x}\Big|_{x=0}$ and decompose $f(x)$ as

$$f(x) = Ax + g(x) \quad \text{where} \quad \frac{|g(x)|}{|x|} \to 0 \text{ as } |x| \to 0.$$

<u>Theorem:</u> The origin is asymptotically stable if $\Re\{\lambda_i(A)\} < 0$ for each eigenvalue, and unstable if $\Re\{\lambda_i(A)\} > 0$ for some eigenvalue.

<u>Note:</u> We can conclude only *local* asymptotic stability from this linearization. Inconclusive if A has eigenvalues on the imaginary axis.

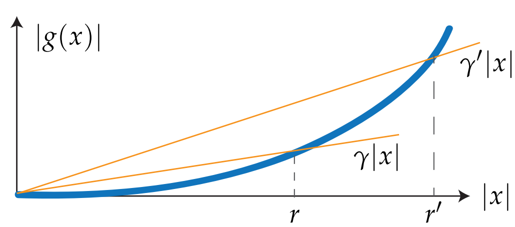
<u>Proof:</u> Find $P = P^T > 0$ such that $A^T P + PA = -Q < 0$. Use $V(x) = x^T P x$ as a Lyapunov function for the nonlinear system $\dot{x} = Ax + g(x)$.

$$\begin{aligned}\dot{V}(x) &= x^T P(Ax + g(x)) + (Ax + g(x))^T P x \\ &= x^T (PA + A^T P)x + 2x^T P g(x) \\ &\leq -x^T Q x + 2|x|\|P\||g(x)|\end{aligned}$$

$$\boxed{\lambda_{\min}(Q)|x|^2 \leq x^T Q x \leq \lambda_{\max}(Q)|x|^2}$$

$$\dot{V}(x) \leq -\lambda_{\min}(Q)|x|^2 + 2\|P\||x||g(x)|$$

Since $\dfrac{|g(x)|}{|x|} \to 0$ as $x \to 0$, for any $\gamma > 0$ we can find $r > 0$ such that $|x| \leq r \Rightarrow |g(x)| \leq \gamma|x|$; see the illustration below for the case $x \in \mathbb{R}$.



Thus, $|x| \leq r(\gamma) \;\Rightarrow\; \dot{V}(x) \leq -\lambda_{\min}(Q)|x|^2 + 2\gamma\|P\||x|^2$.

Choose $\gamma < \frac{\lambda_{\min}(Q)}{2\|P\|}$ so that $\dot{V}$ is negative definite in a ball of radius $r(\gamma)$ around the origin, and appeal to Lyapunov's Stability Theorem (Lecture 8) to conclude (local) asymptotic stability.

Region of Attraction

$$R_A = \{x : \phi(t, x) \to 0\} \tag{1}$$

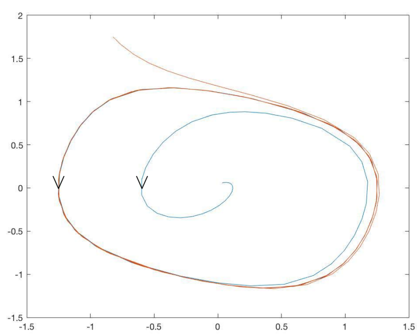
"Quantifies" local asymptotic stability. Global asymptotic stability: $R_A = \mathbb{R}^n$.

Proposition: If $x = 0$ is asymptotically stable, then its region of attraction is an open, connected, invariant set. Moreover, the boundary is formed by trajectories.

Example: van der Pol system in reverse time:

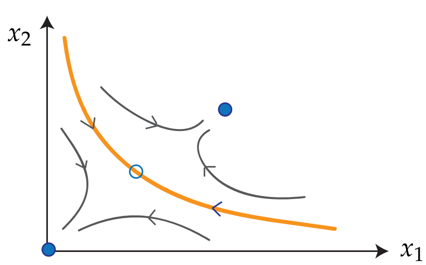
$$\begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= x_1 - x_2 + x_2^3 \end{aligned} \tag{2}$$

The boundary is the (unstable) limit cycle. Trajectories starting within the limit cycle converge to the origin.



Example: bistable switch:

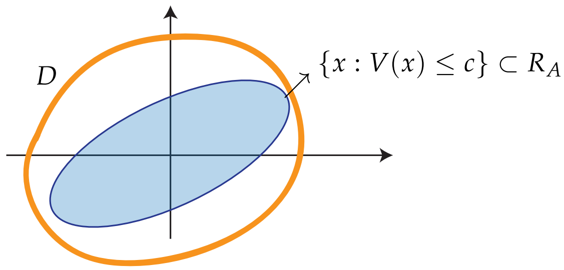
$$\begin{aligned} \dot{x}_1 &= -ax_1 + x_2 \\ \dot{x}_2 &= \frac{x_1^2}{1 + x_1^2} - bx_2 \end{aligned} \tag{3}$$



Estimating the Region of Attraction with a Lyapunov Function

Suppose $\dot{V}(x) < 0$ in $D - \{0\}$. The level sets of V inside D are invariant and trajectories starting in them converge to the origin.

Therefore we can use the largest levet set of V that fits into D as an (under)approximation of the region of attraction.



This estimate depends on the choice of Lyapunov function. A simple (but often conservative) choice is: $V(x) = x^T P x$ where P is selected for the linearization (see p.1).

Time-Varying Systems

Khalil (Sec. 4.5), Sastry (Sec. 5.2)

$$\dot{x} = f(t, x) \quad f(t, 0) \equiv 0 \tag{4}$$

To simplify the definitions of stability and asymptotic stability for the equilibrium $x = 0$, we first define a class of functions known as "comparison functions."

Comparison Functions

Definition: A continuous function $\alpha : [0, \infty) \to [0, \infty)$ is class-$\mathcal{K}$ if it is zero at zero and strictly increasing. It is class-$\mathcal{K}_\infty$ if, in addition, $\alpha(r) \to \infty$ as $r \to \infty$.

A continuous function $\beta : [0, \infty) \times [0, \infty) \to [0, \infty)$ is class-$\mathcal{KL}$ if:

1. $\beta(\cdot, s)$ is class-$\mathcal{K}$ for every fixed s,
2. $\beta(r, \cdot)$ is decreasing and $\beta(r, s) \to 0$ as $s \to \infty$, for every fixed r.

Example: $\alpha(r) = \tan^{-1}(r)$ is class-$\mathcal{K}$, $\alpha(r) = r^c, c > 0$ is class-$\mathcal{K}_\infty$, $\beta(r, s) = r^c e^{-s}$ is class-$\mathcal{KL}$.

Proposition: If $V(\cdot)$ is positive definite, then we can find class-$\mathcal{K}$ functions $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ such that

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|). \tag{5}$$

If $V(\cdot)$ is radially unbounded, we can choose $\alpha_1(\cdot)$ to be class-$\mathcal{K}_\infty$.

Example: $\quad V(x) = x^T P x \quad P = P^T > 0$

$$\alpha_1(|x|) = \lambda_{\min}(P)|x|^2 \quad \alpha_2(|x|) = \lambda_{\max}(P)|x|^2.$$

Stability Definitions

Definition: $x = 0$ is stable if for every $\epsilon > 0$ and t_0, there exists $\delta > 0$ such that

$$|x(t_0)| \leq \delta(t_0, \epsilon) \implies |x(t)| \leq \epsilon \quad \forall t \geq t_0.$$

If the same δ works for all t_0, *i.e.* $\delta = \delta(\epsilon)$, then $x = 0$ is uniformly stable.

It is easier to define uniform stability and uniform asymptotic stability using comparison functions:

- $x = 0$ is uniformly stable if there exists a class-$\mathcal{K}$ function $\alpha(\cdot)$ and a constant $c > 0$ such that

  $$|x(t)| \leq \alpha(|x(t_0)|)$$

  for all $t \geq t_0$ and for every initial condition such that $|x(t_0)| \leq c$.
- uniformly asymptotically stable if there exists a class-$\mathcal{KL}$ $\beta(\cdot, \cdot)$ s.t.

  $$|x(t)| \leq \beta(|x(t_0)|, t - t_0)$$

  for all $t \geq t_0$ and for every initial condition such that $|x(t_0)| \leq c$.
- globally uniformly asymptotically stable if $c = \infty$.
- uniformly exponentially stable if $\beta(r, s) = kre^{-\lambda s}$ for some $k, \lambda > 0$:

  $$|x(t)| \leq k|x(t_0)|e^{-\lambda(t-t_0)}$$

  for all $t \geq t_0$ and for every initial condition such that $|x(t_0)| \leq c$.

Example: Consider the following system, defined for $t > -1$:

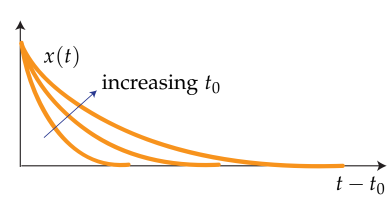
$$\dot{x} = \frac{-x}{1+t} \tag{6}$$

$$\begin{aligned} x(t) &= x(t_0)e^{\int_{t_0}^{t} \frac{-1}{1+s} ds} = x(t_0)e^{\log(1+s)|_t^{t_0}} \\ &= x(t_0)e^{\log \frac{1+t_0}{1+t}} = x(t_0)\frac{1+t_0}{1+t} \end{aligned}$$

$|x(t)| \leq |x(t_0)| \implies$ the origin is uniformly stable with $\alpha(r) = r$.

The origin is also asymptotically stable, but not uniformly, because the convergence rate depends on t_0:

$$x(t) = x(t_0)\frac{1+t_0}{1+t_0+(t-t_0)} = \frac{x(t_0)}{1+\frac{t-t_0}{1+t_0}}.$$



EE C222/ME C237 - Spring'18 - Lecture 11 Notes[1]

Murat Arcak

February 26 2018

[1]

Time-Varying Systems Continued

Uniform stability: There exists a class $\mathcal{K}$ function $\alpha(\cdot)$ and a constant $c > 0$, both independent of t_0, such that

$$|x(t)| \leq \alpha(|x(t_0)|) \quad \forall t \geq t_0 \quad \text{when} \quad |x(t_0)| \leq c.$$

Uniform asymptotic stability: There exists a class $\mathcal{KL}$ function $\beta(\cdot,\cdot)$ and a constant $c > 0$ such that

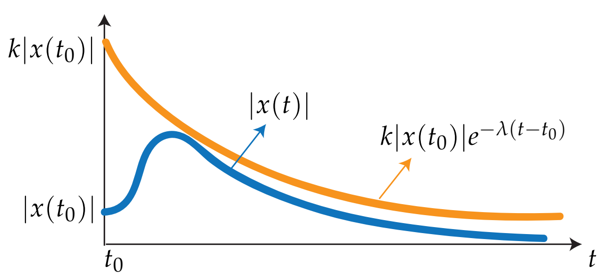
$$|x(t)| \leq \beta(|x(t_0)|, t - t_0) \quad \forall t \geq t_0 \quad \text{when} \quad |x(t_0)| \leq c.$$

Uniform exponential stability: There exist constants $k, \lambda, c > 0$ s.t.

$$|x(t)| \leq k|x(t_0)|e^{-\lambda(t-t_0)} \quad \forall t \geq t_0 \quad \text{when} \quad |x(t_0)| \leq c,$$

that is $\beta(r,s) = kre^{-\lambda s}$.

$k > 1$ allows for overshoot:



Example:

$$\dot{x} = -x^3 \quad \Rightarrow \quad x(t) = \operatorname{sgn}(x(t_0))\sqrt{\frac{x_0^2}{1 + 2(t-t_0)x_0^2}}$$

$x = 0$ is asymptotically stable but not exponentially stable.

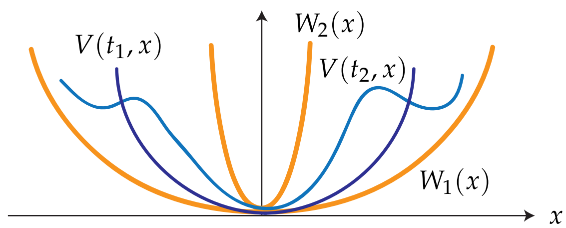
Proposition: $x = 0$ is exponentially stable for $\dot{x} = f(x)$, $f(0) = 0$, if and only if $A \triangleq \left.\frac{\partial f}{\partial x}\right|_{x=0}$ is Hurwitz, that is $\Re\lambda_i(A) < 0\ \forall i$.

Although strict inequality in $\Re\lambda_i(A) < 0$ is not necessary for asymptotic stability (see example above where $A = 0$), it *is* necessary for exponential stability.

Lyapunov's Stability Theorem for Time-Varying Systems

Khalil, Section 4.5

1. If $W_1(x) \leq V(t,x) \leq W_2(x)$ and $\dot{V}(t,x) \triangleq \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t,x) \leq 0$ for some positive definite functions $W_1(\cdot)$, $W_2(\cdot)$ on a domain D that includes the origin, then $x = 0$ is uniformly stable.



2. If, further, $\dot{V}(t,x) \leq -W_3(x)\ \forall x \in D$ for some positive definite $W_3(\cdot)$, then $x = 0$ is uniformly asymptotically stable.
3. If $D = \mathbb{R}^n$ and $W_1(\cdot)$ is radially unbounded, then $x = 0$ is globally uniformly asymptotically stable.
4. If $W_i(x) = k_i|x|^a$, $i = 1,2,3$, for some constants $k_1, k_2, k_3, a > 0$, then $x = 0$ is uniformly exponentially stable.

Proof:

1. $\alpha_1(|x|) \leq W_1(x) \leq V(t,x) \leq W_2(x) \leq \alpha_2(|x|)$

$$\begin{aligned} \dot{V} \leq 0 &\Rightarrow V(x(t),t) \leq V(x(t_0),t_0) \\ &\Rightarrow \alpha_1(|x(t)|) \leq \alpha_2(|x(t_0)|) \\ &\Rightarrow |x(t)| \leq \alpha(|x(t_0)|) \triangleq (\alpha_1^{-1} \circ \alpha_2)(|x(t_0)|). \end{aligned}$$

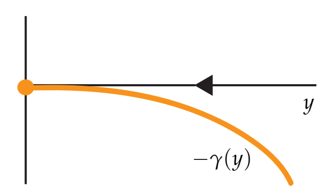
Note: The inverse of a class-$\mathcal{K}$ function is well defined locally (globally if $\mathcal{K}_\infty$) and is class-$\mathcal{K}$. The composition of two class-$\mathcal{K}$ functions is also class-$\mathcal{K}$.

2. $\dot{V} \leq -W_3(x) \leq -\alpha_3(|x|) \leq -\alpha_3(\alpha_2^{-1}(V)) \triangleq -\gamma(V)$

$$\frac{d}{dt}V(t,x(t)) \leq -\gamma(V(t,x(t)))$$

Let $y(t)$ be the solution of $\dot{y} = -\gamma(y)$, $y(t_0) = V(t_0, x(t_0))$. Then,

$$V(t,x(t)) \leq y(t).$$



Since $\dot{y} = -\gamma(y)$ is a first order differential equation and $-\gamma(y) < 0$ when $y > 0$, we conclude monotone convergence of $y(t)$ to 0:

$$y(t) = \beta(y(t_0), t - t_0) \implies V(t,x(t)) \leq \beta(\underbrace{V(t_0,x(t_0))}_{\leq \alpha_2(|x(t_0)|)}, t - t_0)$$

$$\begin{aligned} &\Rightarrow\ \alpha_1(|x(t)|) \leq \beta(\alpha_2(|x(t_0)|), t - t_0) \\ &\Rightarrow\ |x(t)| \leq \tilde{\beta}(|x(t_0)|, t - t_0) \triangleq \alpha_1^{-1}(\beta(\alpha_2(|x(t_0)|), t - t_0)) \end{aligned}$$

3. If $\alpha_1(\cdot)$ is class $\mathcal{K}_\infty$ then $\alpha_1^{-1}(\cdot)$ exists globally above.

4. $\alpha_3(|x|) = k_3|x|^a$, $\alpha_2(|x|) = k_2|x|^a$

$$\Rightarrow \gamma(V) = \alpha_3(\alpha_2^{-1}(V)) = k_3\left(\left(\frac{V}{k_2}\right)^{\frac{1}{a}}\right)^a = \frac{k_3}{k_2}V$$

$$\dot{y} = -\frac{k_3}{k_2}y \;\Rightarrow\; y(t) = y(t_0)e^{-(k_2/k_2)(t-t_0)}$$

$$\beta(r,s) = re^{-(k_3/k_2)s} \;\Rightarrow\; \tilde{\beta}(r,s) = \left(\frac{k_2}{k_1}r^a e^{-(k_3/k_2)s}\right)^{\frac{1}{a}} = \left(\frac{k_2}{k_1}\right)^{\frac{1}{a}} r^{-\frac{k_3 a}{k_2}s}.$$

<u>Example:</u>

$$\dot{x} = -g(t)x^3 \quad \text{where} \quad g(t) \geq 1 \quad \text{for all} \quad t$$

$$V(x) = \frac{1}{2}x^2 \quad \Rightarrow \quad \dot{V}(t,x) = -g(t)x^4 \leq -x^4 \triangleq W_3(x)$$

Globally uniformly asymptotically stable but not exponentially stable. Take $g(t) \equiv 1$ as a special case:

$$\dot{x} = -x^3 \quad \Rightarrow \quad x(t) = \mathrm{sgn}(x(t_0))\sqrt{\frac{x_0^2}{1+2(t-t_0)x_0^2}}$$

which converges slower than exponentially.

<u>Example:</u> $\dot{x} = A(t)x$. Take $V(x) = x^T P(t)x$:

$$\begin{aligned}\dot{V}(x) &= x^T\dot{P}(t)x + \dot{x}^T P(t)x + x^T P(t)\dot{x} \\ &= x^T\underbrace{(\dot{P} + A^T P + PA)}_{\triangleq -Q(t)}x\end{aligned}$$

If $k_1 I \leq P(t) \leq k_2 I$ and $k_3 I \leq Q(t)$, $k_1, k_2, k_3 > 0$, then

$$k_1|x|^2 \leq V(t,x) \leq k_2|x|^2 \quad \text{and} \quad \dot{V}(t,x) \leq -k_3|x|^2$$

$\Rightarrow$ global uniform exponential stability.

<u>What if $W_3(\cdot)$ is only semidefinite?</u>

Khalil, Section 8.3

Lasalle-Krasovskii Invariance Principle is <u>not</u> applicable to time-varying systems. Instead, use the following (weaker) result:

<u>Theorem:</u> Suppose $W_1(x) \leq V(t,x) \leq W_2(x)$

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}f(t,x) \leq -W_3(x),$$

where $W_1(\cdot), W_2(\cdot)$ are positive definite and $W_3(\cdot)$ is positive semidefinite. Suppose, further, $W_1(\cdot)$ is radially unbounded, $f(t,x)$ is locally Lipschitz in x and bounded in t, and $W_3(\cdot)$ is C^1. Then

$$\boxed{W_3(x(t)) \to 0 \text{ as } t \to \infty.}$$

Note: This proves convergence to $S = \{x : W_3(x) = 0\}$ whereas the Invariance Principle, when applicable, guarantees convergence to the largest invariant set within S.

Example:

$$\begin{aligned}\dot{x}_1 &= -x_1 + w(t)x_2 \\ \dot{x}_2 &= -w(t)x_1\end{aligned}$$

$V(t,x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \;\Rightarrow\; \dot{V}(t,x) = -x_1^2$. If $w(t)$ is bounded in t then the theorem above implies $x_1(t) \to 0$ as $t \to \infty$, but no guarantee about the convergence of $x_2(t)$ to zero.

By contrast, if $w(t) \equiv w \neq 0$, then we can use the Invariance Principle and conclude $x_2(t) \to 0$ (show this).

Barbalat's Lemma (used in proving the theorem above):

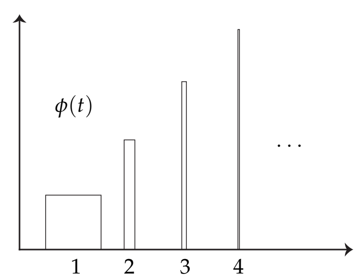
If $\lim_{t\to\infty} \int_0^t \phi(\tau)d\tau$ exists and is finite, and $\phi(\cdot)$ is *uniformly continuous*[2] then $\phi(t) \to 0$ as $t \to \infty$.

[2] For every $\epsilon > 0$ there exists $\delta > 0$ such that $\forall t_1, t_2\ |t_1 - t_2| \leq \delta \Rightarrow |\phi(t_1) - \phi(t_2)| \leq \epsilon$. Boundedness of the derivative $\dot{\phi}(t)$ implies uniform continuity.

Uniform continuity in Barbalat's Lemma can't be relaxed:

Example: Let $\phi(t)$ be a sequence of pulses centered at $k = 1, 2, 3, \ldots$ with amplitude $= k$, width $= 1/k^3$, then

$$\int_0^\infty \phi(t)dt = \sum_{k=1}^\infty \frac{1}{k^2} < \infty \quad \text{but} \quad \phi(t) \not\to 0.$$



Proof of the theorem:

$$\begin{aligned}&\alpha_1(|x|) \leq V(t,x) \leq \alpha_2(|x|) \quad \alpha_1 \in \mathcal{K}_\infty \\ &\Rightarrow |x(t)| \leq \alpha_1^{-1}(\alpha_2(|x(t_0)|))\end{aligned}$$

$x(t)$ bounded $\Rightarrow \dot{x}(t) = f(t, x(t))$ is bounded $\Rightarrow x(t)$ is uniformly continuous.

$$\begin{aligned}&\dot{V}(t,x) \leq -W_3(x(t)) \\ &\Rightarrow V(x(T)) - V(x(t_0), t_0) \leq -\int_{t_0}^T W_3(x(t))dt \\ &\Rightarrow \int_{t_0}^\infty W_3(x(t))dt \leq V(x(t_0), t_0) < \infty.\end{aligned}$$

Since $W_3(\cdot)$ is C^1, it is uniformly continuous on the bounded domain where $x(t)$ resides. So, by Barbalat's Lemma, $W_3(x(t)) \to 0$ as $t \to \infty$.

EE C222/ME C237 - Spring'18 - Lecture 12 Notes[1]

Murat Arcak

February 28 2018

Linear Time-Varying Systems

Khalil Section 4.6, Sastry Section 5.7

$$\dot{x} = A(t)x \qquad x(t) = \Phi(t,t_0)x(t_0) \tag{1}$$

- The state transition matrix $\Phi(t,t_0)$ satisfies the equations:

$$\frac{\partial}{\partial t}\Phi(t,t_0) = A(t)\Phi(t,t_0) \tag{2}$$

$$\frac{\partial}{\partial t_0}\Phi(t,t_0) = -\Phi(t,t_0)A(t_0) \tag{3}$$

- No eigenvalue test for stability in the time-varying case:

$$A(t) = \begin{bmatrix} -1+1.5\cos^2 t & 1-1.5\sin t\cos t \\ -1-1.5\sin t\cos t & -1+1.5\sin^2 t \end{bmatrix}$$

  eigenvalues: $-0.25 \mp i0.25\sqrt{7}$ for all t, but unstable:

$$\Phi(t,0) = \begin{bmatrix} e^{0.5t}\cos t & e^{-t}\sin t \\ e^{-0.5t}\sin t & e^{-t}\cos t \end{bmatrix}$$

- For linear systems uniform asymptotic stability is equivalent to uniform exponential stability:

  Theorem[2]: $x = 0$ is uniformly asymptotically stable if and only if

$$\|\Phi(t,t_0)\| \le ke^{-\lambda(t-t_0)} \text{ for some } k>0,\ \lambda>0.$$

[2] Khalil Thm. 4.11, Sastry Thm. 5.33

- Last lecture: $V(t,x) = x^T P(t)x$ proves uniform exp. stability if

$$\begin{array}{ll} \text{(i)} & \dot{P}(t) + A^T(t)P(t) + P(t)A(t) = -Q(t) \\ \text{(ii)} & 0 < k_1 I \le P(t) \le k_2 I \\ \text{(iii)} & 0 < k_3 I \le Q(t) \text{ for all } t. \end{array}$$

  The converse is also true:

  Theorem: Suppose $x = 0$ is uniformly exponentially stable, $A(t)$ is continuous and bounded, $Q(t)$ is continuous and symmetric, and there exist $k_3, k_4 > 0$ such that

$$0 < k_3 I \le Q(t) \le k_4 I \text{ for all } t.$$

  Then, there exists a symmetric $P(t)$ satisfying (i)–(ii) above.

Proof:

Time-invariant: $P = \int_0^\infty e^{A^T\tau} Q e^{A\tau} d\tau$

Time-varying: $P(t) = \int_t^\infty \Phi^T(\tau,t) Q(\tau) \Phi(\tau,t) d\tau$

Using the Leibniz rule, property (3), and $\Phi(t,t) = I$ we obtain:

$$\begin{aligned}
\dot{P}(t) &= \int_t^\infty \left(\frac{\partial}{\partial t}\Phi^T(\tau,t)Q(\tau)\Phi(\tau,t) + \Phi^T(\tau,t)Q(\tau)\frac{\partial}{\partial t}\Phi(\tau,t) \right) d\tau \\
&\qquad\qquad - \Phi^T(t,t)Q(t)\Phi(t,t) \\
&= \int_t^\infty \left(-A^T(t)\Phi^T(\tau,t)Q(\tau)\Phi(\tau,t) - \Phi^T(\tau,t)Q(\tau)\Phi(\tau,t)A(t) \right) d\tau \\
&\qquad\qquad - \Phi^T(t,t)Q(t)\Phi(t,t) \\
&= -A^T(t)P(t) - P(t)A(t) - Q(t).
\end{aligned}$$

Lyapunov-based Feedback Design Examples

Model Reference Adaptive Control

Illustrated on a first order system:

$$\dot{y} = a^* y + u \tag{4}$$

where a^* is unknown.

Reference model:

$$\dot{y}_m = -a y_m + r(t) \quad a > 0,\ r(t) : \text{reference signal.} \tag{5}$$

Goal: Design a controller that guarantees $y(t) - y_m(t) \to 0$ without the knowledge of a^*.

If we knew a^*, we would choose:

$$u = -\underbrace{(a^* + a)}_{=:k^*} y + r(t) \quad \Rightarrow \quad \dot{y} = -ay + r(t).$$

The tracking error $\boxed{e(t) := y(t) - y_m(t)}$ then satisfies:

$$\dot{e} = -ae \ \Rightarrow \ e(t) \to 0 \text{ exponentially.}$$

Adaptive design when a^* (therefore, k^*) is unknown:

$$u = -k(t)y + r(t)$$

where $\dot{k}(t)$ is to be designed. Then: $\dot{e} = -ae - \underbrace{(k(t) - k^*)}_{=:\tilde{k}(t)} y.$

Use the Lyapunov function: $V = \frac{1}{2}e^2 + \frac{1}{2}\tilde{k}^2$:

$$\begin{aligned} \dot{V} &= -ae^2 - \tilde{k}ey + \tilde{k}\dot{\tilde{k}} \\ &= -ae^2 + \tilde{k}(\dot{\tilde{k}} - ey). \end{aligned}$$

Note $\dot{\tilde{k}} = \dot{k}$ and choose $\boxed{\dot{k} = ey}$ so that $\dot{V} = -ae^2$.

This guarantees stability of $(e, \tilde{k}) = (0, 0)$ and boundedness of $(e(t), \tilde{k}(t))$ since the level sets of $V = \frac{1}{2}e^2 + \frac{1}{2}\tilde{k}^2$ are positively invariant. In addition, if $r(t)$ is bounded, then $y_m(t)$ in (5) is bounded, and so is $y(t) = y_m(t) + e(t)$. Then we can apply the Theorem from Lecture 11, page 3, to the time-varying model

$$\dot{e} = -ae - y(t)\tilde{k}, \quad \dot{\tilde{k}} = y(t)e,$$

and conclude from $\dot{V} = -ae^2$ that $e(t) \to 0$.

Whether $\tilde{k}(t) \to 0$ ($k(t) \to k^*$) depends on further properties of the reference signal $r(\cdot)$ that are beyond the scope of this lecture.

Backstepping

Khalil (Sec. 14.3), Sastry (Sec. 6.8)

Feedback stabilization: Given the system

$$\dot{x} = f(x) + g(x)u \tag{6}$$

with input u, design a control law $u = \alpha(x)$ such that $x = 0$ is asymptotically stable for the closed-loop system:

$$\dot{x} = f(x) + g(x)\alpha(x).$$

Backstepping is a technique that simplifies this task for a class of systems.

Suppose a stabilizing feedback $u = \alpha(X)$ is available for:

$$\dot{X} = F(X) + G(X)u \quad X \in \mathbb{R}^n, u \in \mathbb{R}$$

and suppose the closed-loop system admits a Lyapunov function $V(X)$ such that

$$\frac{\partial V}{\partial X}\Big(F(X) + G(X)\alpha(X)\Big) \leq -W(X) < 0 \quad \forall X \neq 0.$$

Can we modify $\alpha(X)$ to stabilize the augmented system below?

$$\boxed{\begin{aligned} \dot{X} &= F(X) + G(X)x \\ \dot{x} &= u. \end{aligned}}$$

Define the error variable $\boxed{z = x - \alpha(X)}$ and change variables:

$(X, x) \to (X, z)$:

$$\begin{aligned}\dot{X} &= F(X) + G(X)\alpha(X) + G(X)z \\ \dot{z} &= u - \dot{\alpha}(X, z)\end{aligned}$$

where $\dot{\alpha}(X,z) = \frac{\partial \alpha}{\partial X}\Big(F(X) + G(X)\alpha(X) + G(X)z\Big)$. Take the new Lyapunov function:

$$V_+(X,z) = V(X) + \frac{1}{2}z^2.$$

$$\dot{V}_+ = \underbrace{\frac{\partial V}{\partial X}\Big(F(X) + G(X)\alpha(X)\Big)}_{\leq -W(X)} + \underbrace{\frac{\partial V}{\partial X}G(X)z + z(u - \dot{\alpha})}_{= z\left(u - \dot{\alpha} + \frac{\partial V}{\partial X}G(X)\right)}$$

Let: $\boxed{u = \dot{\alpha} - \frac{\partial V}{\partial X}G(X) - kz, \quad k > 0.}$

Then, $\dot{V}_+ \leq -W(X) - kz^2 \;\Rightarrow\; (X,z) = 0$ is asymptotically stable.

<u>Example:</u>

$$\begin{aligned}\dot{x}_1 &= x_1^2 + x_2 \\ \dot{x}_2 &= u.\end{aligned} \tag{7}$$

Treat x_2 as "virtual" control input for the x_1-subsystem:

$$\begin{aligned}\alpha(x_1) &= -k_1 x_1 - x_1^2 \quad k_1 > 0 \\ V_1(x_1) &= \frac{1}{2}x_1^2.\end{aligned}$$

Apply backstepping:

$$\begin{aligned}z_2 &= x_2 - \alpha(x_1) = x_2 + k_1 x_1 + x_1^2 \\ \dot{z}_2 &= u - \dot{\alpha} \\ u &= \dot{\alpha} - \frac{\partial V_1}{\partial x_1} - k_2 z_2, \quad k_2 > 0 \\ &= \underbrace{-(k_1 + 2x_1)(x_1^2 + x_2)}_{= \dot{\alpha}} - \underbrace{x_1}_{= \frac{\partial V_1}{\partial x_1}} - k_2\underbrace{(x_2 + k_1 x_1 + x_1^2)}_{= z_2}.\end{aligned}$$

EE C222/ME C237 - Spring'18 - Lecture 13 Notes[1]

Murat Arcak

March 5 2018

Backstepping

Khalil (Sec. 14.3), Sastry (Sec. 6.8)

Suppose a stabilizing feedback $u = \alpha(X)$, $\alpha(0) = 0$, is available for:

$$\dot{X} = F(X) + G(X)u \quad X \in \mathbb{R}^n, u \in \mathbb{R},\ F(0) = 0,$$

along with a Lyapunov function V such that

$$\frac{\partial V}{\partial X}\Big(F(X) + G(X)\alpha(X)\Big) \leq -W(X) < 0 \quad \forall X \neq 0.$$

Can we modify $\alpha(X)$ to stabilize the augmented system below?

$$\boxed{\begin{aligned} \dot{X} &= F(X) + G(X)x \\ \dot{x} &= u. \end{aligned}}$$

Define the error variable $\boxed{z = x - \alpha(X)}$ and change variables: $(X, x) \to (X, z)$:

$$\begin{aligned} \dot{X} &= F(X) + G(X)\alpha(X) + G(X)z \\ \dot{z} &= u - \dot{\alpha}(X, z) \end{aligned}$$

where $\dot{\alpha}(X,z) = \frac{\partial \alpha}{\partial X}\Big(F(X) + G(X)\alpha(X) + G(X)z\Big)$. Take the new Lyapunov function:

$$V_+(X, z) = V(X) + \frac{1}{2}z^2.$$

$$\dot{V}_+ = \underbrace{\frac{\partial V}{\partial X}\Big(F(X) + G(X)\alpha(X)\Big)}_{\leq -W(X)} + \underbrace{\frac{\partial V}{\partial X}G(X)z + z(u - \dot{\alpha})}_{= z\left(u - \dot{\alpha} + \frac{\partial V}{\partial X}G(X)\right)}$$

Let: $\boxed{u = \dot{\alpha} - \frac{\partial V}{\partial X}G(X) - kz, \quad k > 0.}$

Then, $\dot{V}_+ \leq -W(X) - kz^2 \;\Rightarrow\; (X, z) = 0$ is asymptotically stable.

Example 1:

$$\begin{aligned} \dot{x}_1 &= x_1^2 + x_2 \\ \dot{x}_2 &= u. \end{aligned} \tag{1}$$

Treat x_2 as "virtual" control input for the x_1-subsystem:

$$\begin{aligned} \alpha(x_1) &= -k_1 x_1 - x_1^2 \quad k_1 > 0 \\ V_1(x_1) &= \frac{1}{2}x_1^2. \end{aligned}$$

Apply backstepping:

$$
\begin{aligned}
z_2 &= x_2 - \alpha(x_1) = x_2 + k_1 x_1 + x_1^2 \\
\dot{z}_2 &= u - \dot{\alpha} \\
u &= \dot{\alpha} - \frac{\partial V_1}{\partial x_1} - k_2 z_2, \quad k_2 > 0 \\
&= \underbrace{-(k_1 + 2x_1)(x_1^2 + x_2)}_{=\dot{\alpha}} - \underbrace{x_1}_{=\frac{\partial V_1}{\partial x_1}} - k_2 \underbrace{(x_2 + k_1 x_1 + x_1^2)}_{=z_2}.
\end{aligned}
$$

- Above we discussed backstepping over a pure integrator. The main idea generalizes trivially to:

$$
\boxed{\begin{aligned}
\dot{X} &= F(X) + G(X)x \\
\dot{x} &= f(X, x) + g(X, x)u
\end{aligned}}
$$

where $X \in \mathbb{R}^n$, $x \in \mathbb{R}$, and $g(X, x) \neq 0$ for all $(X, x) \in \mathbb{R}^{n+1}$.

With the preliminary feedback

$$
u = \frac{1}{g(X, x)}(-f(X, x) + v) \tag{2}
$$

the x-subsystem becomes a pure integrator: $\dot{x} = v$. Substituting the backstepping control law from above:

$$
v = \dot{\alpha} - \frac{\partial V}{\partial X} G(X) - kz, \quad z \triangleq x - \alpha(X), \quad k > 0
$$

into (2), we get:

$$
\boxed{u = \frac{1}{g(X, x)} \left(-f(X, x) + \dot{\alpha} - \frac{\partial V}{\partial X} G(X) - kz \right).}
$$

- Backstepping can be applied recursively to systems of the form:[2]

$$
\begin{aligned}
\dot{x}_1 &= f_1(x_1) + g_1(x_1)x_2 \\
\dot{x}_2 &= f_2(x_1, x_2) + g_2(x_1, x_2)x_3 \\
\dot{x}_3 &= f_3(x_1, x_2, x_3) + g_3(x_1, x_2, x_3)x_4 \\
&\vdots \\
\dot{x}_n &= f_n(x) + g_n(x)u
\end{aligned} \tag{3}
$$

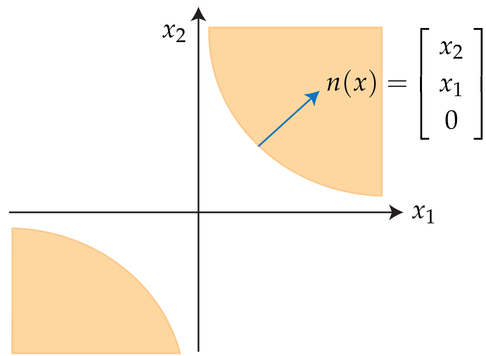
[2] Systems of this form are called "strict feedback systems."

where $g_i(x_1, \ldots, x_i) \neq 0$ for all $x \in \mathbb{R}^n$, $i = 1, 2, \cdots, n$.

<u>Example 2:</u>

$$
\begin{aligned}
\dot{x}_1 &= (x_1 x_2 - 1)x_1^3 + (x_1 x_2 + x_3^2 - 1)x_1 \\
\dot{x}_2 &= x_3 \\
\dot{x}_3 &= u.
\end{aligned} \tag{4}
$$

Not in strict feedback form because x_3 appears too soon. In fact, this system is not globally stabilizable because the set $x_1x_2 \geq 2$ is positively invariant regardless of u:



To see this, note that

$$n(x) \cdot f(x,u) = [(x_1x_2 - 1)x_1^3 + (x_1x_2 + x_3^2 - 1)x_1]x_2 + x_3x_1$$

and substitute $x_1x_2 = 2$:

$$\begin{aligned}
&= \left(x_1^3 + (1 + x_3^2)x_1\right)x_2 + x_3x_1 \\
&= \left(x_1^2 + (1 + x_3^2)\right)x_1x_2 + x_3x_1 \\
&= 2x_1^2 + 2(1 + x_3^2) + x_3x_1 \\
&= \underbrace{2x_1^2 + x_3x_1 + 2x_3^2}_{\geq 0} + 2 > 0.
\end{aligned}$$

- The condition $g_i(x_1, \ldots, x_i) \neq 0$ in (3) can be relaxed in some cases:

Example 3:

$$\begin{aligned}
\dot{x}_1 &= x_1^2 x_2 \\
\dot{x}_2 &= u
\end{aligned} \tag{5}$$

Treat x_2 as virtual control and let $\boxed{\alpha_1(x_1) = -x_1}$ which stabilizes the x_1-subsystem, as verified with Lyapunov function $V_1(x_1) = \frac{1}{2}x_1^2$. Then $z_2 := x_2 - \alpha_1(x_1)$ satisfies $\dot{z}_2 = u - \dot{\alpha}_1$, and

$$u = \dot{\alpha}_1 - \frac{\partial V_1}{\partial x_1}x_1^2 - k_2z_2 = -x_1^2x_2 - x_1^3 - k_2(x_2 + x_1)$$

achieves global asymptotic stability:

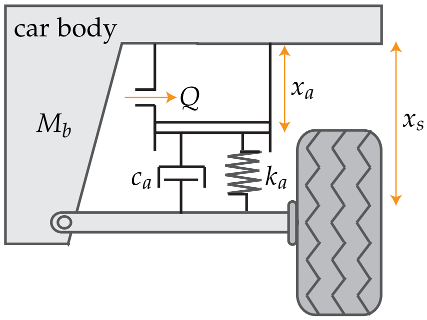
$$V = \frac{1}{2}x_1^2 + \frac{1}{2}z_2^2 \quad \Rightarrow \quad \dot{V} = -x_1{}^4 - k_2z_2^2.$$

Note that we can't conclude exponential stability due to the quartic term x_1^4 above (recall the Lyapunov sufficient condition for exponential stability in Lecture 11, p.2). In fact, the linearization of the closed-loop system proves the lack of exponential stability:

$$\begin{bmatrix} 0 & 0 \\ 0 & -k_2 \end{bmatrix} \rightarrow \lambda_{1,2} = 0, -k_2.$$

Design example: Active suspension

Krstić et al., Nonlinear and Adaptive Control Design, Section 2.2.2.



$$M_b \ddot{x}_s = -k_a(x_s - x_a) - c_a(\dot{x}_s - \dot{x}_a)$$

$$\dot{x}_a = \frac{1}{A} Q \qquad A\text{: effective piston surface}$$

$$\text{Flow: } \dot{Q} = -c_f Q + k_f u \qquad u\text{: current applied to the solenoid valve (control input)}$$

Define state variables: $x_1 = x_s$, $x_2 = \dot{x}_s$, $x_3 = x_a$, $x_4 = Q$:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k_a}{M_b}(x_1 - x_3) - \frac{c_a}{M_b}(x_2 - \frac{1}{A} x_4) \\ \dot{x}_3 &= \frac{1}{A} x_4 \\ \dot{x}_4 &= -c_f x_4 + k_f u. \end{aligned} \tag{6}$$

This system is not in strict feedback form due to the x_4 term in $\dot{x}_2$. To overcome this problem define:

$$\begin{aligned} \bar{x}_3 &\triangleq \frac{k_a}{M_b} x_3 + \frac{c_a}{M_b A} x_4 \\ \xi &\triangleq x_3 \end{aligned}$$

and change variables to $(x_1, x_2, \bar{x}_3, \xi)$:

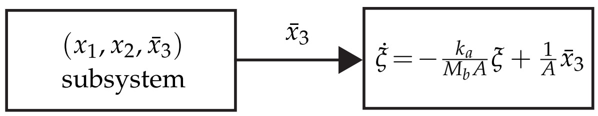
$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k_a}{M_b} x_1 - \frac{c_a}{M_b} x_2 + \bar{x}_3 \\ \dot{\bar{x}}_3 &= \frac{k_a - c_a c_f}{M_b A} x_4 + \frac{c_a k_f}{M_b A} u. \end{aligned}$$

Two steps of backstepping starting with the virtual control law:

The stiff nonlinearity $k_1 x_1^3$ prevents large excursions of x_1.

$$\alpha_1(x_1) = -c_1 x_1 - k_1 x_1^3$$

will stabilize the $(x_1, x_2, \bar{x}_3)$ subsystem. Full $(x_1, x_2, \bar{x}_3, \xi)$ system:



The ξ-subsystem is an asymptotically stable linear system driven by $\bar{x}_3$; therefore the full system is stabilized.

EE C222/ME C237 - Spring'18 - Lecture 14 Notes[1]

Murat Arcak

March 7 2018

Input-to-State Stability

Khalil, Section 4.9

$$\dot{x} = f(x,u) \quad u\text{: exogenous input}$$

For linear systems, asymp. stability of the zero-input model $\dot{x} = Ax$ implies a bounded-input bounded-state property for $\dot{x} = Ax + Bu$:

$$\begin{aligned} x(t) &= e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau \\ \Longrightarrow |x(t)| &\leq \|e^{At}\||x_0| + \int_0^t \|e^{A(t-\tau)}\|\|B\||u(\tau)|d\tau \\ &\leq \kappa e^{-\alpha t}|x_0| + \|B\| \sup_{0\leq\tau\leq t} |u(\tau)| \int_0^t \kappa e^{-\alpha(t-\tau)}d\tau \\ &\leq \underbrace{\kappa e^{-\alpha t}|x_0|}_{\text{effect of initial condition}} + \underbrace{\frac{\kappa}{\alpha}\|B\| \sup_{0\leq\tau\leq t} |u(\tau)|}_{\text{effect of input}}. \end{aligned}$$

For nonlinear systems $\dot{x} = f(x,u)$, asymp. stability of the origin for the zero-input model $\dot{x} = f(x,0)$ does <u>not</u> guarantee boundedness of states under bounded inputs.

<u>Example 1:</u> $\dot{x} = -x + xu$

$u(t) \equiv \text{constant} > 1 \Longrightarrow$ exponential growth of $x(t)$.

A precise formulation of the bounded-input bounded-state property for nonlinear systems:

<u>Definition:</u> The system $\dot{x} = f(x,u)$, $f(0,0) = 0$ is said to be input-to-state stable (ISS) if:

$$\boxed{|x(t)| \leq \beta(|x(0)|, t) + \gamma\left(\sup_{0\leq\tau\leq t} |u(\tau)|\right)}$$

for some class-$\mathcal{KL}$ function β and class-$\mathcal{K}$ function γ, called an ISS gain function.

<u>Example:</u> For the linear system above, $\gamma(s) = \frac{\kappa}{\alpha}\|B\|s$.

Implications of ISS

1. $\dot{x} = f(x,u)$ ISS $\Longrightarrow$ $\dot{x} = f(x,0)$ globally asymptotically stable

Proof:

Substitute $u(t) \equiv 0$ in the definition above: $|x(t)| \leq \beta(|x(0)|, t)$.

2. $u(t) \to 0$ as $t \to \infty \;\Rightarrow\; x(t) \to 0$ as $t \to \infty$.

Proof:

Need to show that for any $\epsilon > 0$, there exists T such that

$$|x(t)| \leq \epsilon \quad \forall t \geq T.$$

Since $u(t) \to 0$, we can find T_1 such that $\gamma(|u(t)|) \leq \epsilon/2$ for all $t \geq T_1$. Choose $t_0 = T_1$ and apply ISS definition:

$$|x(t)| \leq \beta(|x(T_1)|, t - T_1) + \epsilon/2 \quad \forall t \geq T_1.$$

Choose T_2 such that

$$\beta(|x(T_1)|, T_2) \leq \epsilon/2.$$

Then, $|x(t)| \leq \epsilon$ for all $t \geq T_1 + T_2 \triangleq T$.

A Lyapunov Characterization of ISS

The system $\dot{x} = f(x, u)$ is ISS if there exist class-$\mathcal{K}_\infty$ functions $\alpha_i,\ i = 1, 2, 3, 4$, and a C^1 function V such that

$$\begin{aligned} \alpha_1(|x|) \leq V(x) &\leq \alpha_2(|x|) \\ \frac{\partial V}{\partial x} f(x, u) &\leq -\alpha_3(|x|) + \alpha_4(|u|). \end{aligned}$$

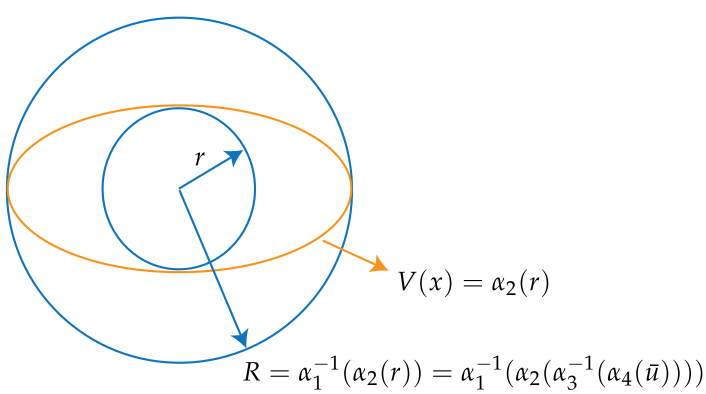
V is called an "ISS Lyapunov function."

Sketch of the proof:

Let $\bar{u} \triangleq \sup_{\tau \geq 0} |u(\tau)|$. Then:

$$|x| \geq r \triangleq \alpha_3^{-1}(\alpha_4(\bar{u})) \quad \Rightarrow \quad \frac{\partial V}{\partial x} f(x, u(t)) \leq 0 \quad \forall t \geq 0.$$

This implies that the level set $\{x : V(x) \leq \alpha_2(r)\}$ is invariant and attractive. Thus, all trajectories converge to this level set which is enclosed in the outer ball $|x| \leq R \triangleq \alpha_1^{-1}(\alpha_2(r))$.



Example 2: $\dot{x} = -x^r + x^s u$, r: odd integer, is ISS if $r > s$. Take:

$$V(x) = \frac{1}{2}x^2$$
$$\dot{V}(x) = -x^{r+1} + x^{s+1}u.$$

Young's inequality (recall from homework):

$$yz \leq \frac{\lambda^p}{p}|y|^p + \frac{1}{q\lambda^q}|z|^q$$

for any $\lambda > 0$, and $p > 1, q > 1$ satisfying $(p-1)(q-1) = 1$. Apply to:

$$x^{s+1}u \leq \frac{\lambda^p}{p}|x|^{(s+1)p} + \frac{1}{q\lambda^q}|u|^q$$

and choose

$$p = \frac{r+1}{s+1} \quad q = 1 + \frac{1}{p-1} \text{ and } \lambda \text{ such that } \frac{\lambda^p}{p} = \frac{1}{2}$$

$$\Rightarrow \; x^{s+1}u \leq \frac{1}{2}|x|^{r+1} + \frac{1}{q\lambda^q}|u|^q$$
$$\Rightarrow \; \dot{V}(x) \leq -|x|^{r+1} + \frac{1}{2}|x|^{r+1} + \frac{1}{q\lambda^q}|u|^q$$
$$\leq \underbrace{-\frac{1}{2}|x|^{r+1}}_{-\alpha_3(|x|)} + \underbrace{\frac{1}{q\lambda^q}|u|^q}_{-\alpha_4(|u|)}.$$

Note:

- $\dot{x} = -x + xu$ $(r = s = 1)$ is not ISS as shown in Example 1.
- $\dot{x} = -x + x^2u$ $(r = 1, s = 2)$ is not ISS: it exhibits finite time escape for $u(t) \equiv$ constant $\neq 0$, even with an exponentially decaying $u(t)$.
- $\dot{x} = -x^3 + u$ $(r = 3, s = 0)$ is ISS.

Example 3:

$$\dot{x}_1 = -x_1 + x_2^2$$
$$\dot{x}_2 = -x_2 + u.$$

Let $V(x) = \frac{1}{2}x_1^2 + \frac{a}{4}x_2^4$, $a > 0$ to be determined.[2]

[2] Why not quadratic V?

$$\dot{V}(x) = -x_1^2 + x_1x_2^2 + a(-x_2^4 + x_2^3u)$$

Apply the Young Inequalities:

$$x_1x_2^2 \leq \frac{1}{2}x_1^2 + \frac{1}{2}x_2^4$$
$$x_2^3u \leq \frac{\lambda^{4/3}}{4/3}x_2^4 + \frac{1}{4\lambda^4}u^4$$

Choose λ such that $\frac{\lambda^{4/3}}{4/3} = \frac{1}{2}$.

$$\dot{V}(x) \leq -\frac{1}{2}x_1^2 + \frac{1}{2}x_2^4 + a\left(-\frac{1}{2}x_2^4 + \frac{1}{4\lambda^4}u^4\right)$$

Let $a = 2$:

$$\dot{V}(x) \leq \underbrace{-\frac{1}{2}x_1^2 - \frac{1}{2}x_2^4}_{\leq -\alpha_3(|x|)} + \underbrace{\frac{1}{2\lambda^4}u^4}_{=\alpha_4(|u|)}$$

for an appropriate choice of α_3. Thus, the system is ISS.

Stability of Series Interconnections

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \quad && x_1 \in \mathbb{R}^{n_1} \\ \dot{x}_2 &= f_2(x_2) && x_2 \in \mathbb{R}^{n_2} \end{aligned} \tag{1}$$

$$\boxed{\dot{x}_2 = f_2(x_2)} \xrightarrow{x_2} \boxed{\dot{x}_1 = f_1(x_1, x_2)}$$

Suppose $x_2 = 0$ is globally asymptotically stable for $\dot{x}_2 = f_2(x_2)$ and $x_1 = 0$ is globally asymptotically stable for $\dot{x}_1 = f_1(x_1, 0)$. Is $(x_1, x_2) = 0$ globally asymptotically stable for the interconnection?

Answer: No.

Example 4:

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_1^2 x_2 \\ \dot{x}_2 &= -x_2 \end{aligned}$$

exhibits finite time escape.

Proposition: Consider the series interconnection:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2) \\ \dot{x}_2 &= f_2(x_2, u). \end{aligned}$$

If the x_1 subsystem is ISS with x_2 viewed as an input, and the x_2 subsystem is ISS with input u, then the interconnection is ISS.

Example 3 revisited:

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2^2 \quad && \text{is ISS with respect to } x_2 \\ \dot{x}_2 &= -x_2 + u && \text{is ISS with input } u \end{aligned}$$

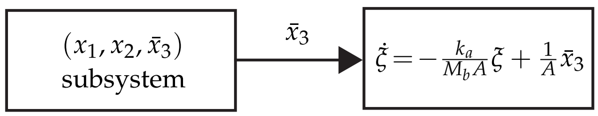
$\Rightarrow$ the interconnection is ISS — an alternative to the proof in Ex. 3.

Corollary: $(x_1, x_2) = 0$ is globally asymptotically stable when $u \equiv 0$.

$\boxed{\text{GAS}} \rightarrow \boxed{\text{ISS}} \equiv \text{GAS}$

Note that Example 4 fails the ISS condition for the x_1 subsystem.

Example: Active suspension design example in Lecture 13:



The $(x_1, x_2, \bar{x}_3)$-subsystem globally asymptotically stabilized by backstepping. The ξ-subsystem is an asymptotically stable linear system, therefore ISS with respect to the input $\bar{x}_3$.

EE C222/ME C237 - Spring'18 - Lecture 15 Notes[1]

Murat Arcak

March 12 2018

Reachable Sets and Safety Certification

Reachable sets with unit peak inputs

$$\boxed{R_T \triangleq \{x(T) \mid \dot{x} = f(x,u),\ x(0) = 0,\ |u| \leq 1\}} \tag{1}$$

The set of points that can be reached from $x(0) = 0$ with inputs not exceeding unit magnitude. Difficult to find exactly, but methods exist to find overapproximations.

ISS gives a very conservative bound:

$$|x(T)| \leq \underbrace{\beta(|x(0)|, T)}_{=0} + \gamma\left(\sup_{0\leq t\leq T} \underbrace{|u(t)|}_{\leq 1}\right) \leq \gamma(1).$$

A less conservative estimate with level sets:

Find positive definite $V(\cdot)$ and a constant $c > 0$ such that

$$|u| \leq 1 \quad \text{and} \quad V(x) \geq c \quad \Rightarrow \quad \nabla V(x) \cdot f(x,u) \leq 0.$$

Then, the level set $\Omega_c \triangleq \{x : V(x) \leq c\}$ contains the reachable set:

$$R_T \subset \Omega_c \quad \forall T \geq 0.$$

Example: Linear system $\dot{x} = Ax + Bu$. Use $V(x) = x^T P x$. If there exists $P = P^T > 0$ such that

$$\boxed{u^T u \leq 1 \text{ and } x^T P x \geq 1 \quad \Rightarrow \quad x^T(A^T P + PA)x + x^T PBu + u^T B^T P x \leq 0}$$

then the ellipsoid $\{x : x^T P x \leq 1\}$ is an overapproximation of R_T.

Rewrite the above implication as:

$$\left\{\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + 1 \geq 0\right\} \wedge \left\{\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 1 \geq 0\right\}$$

$$\Longrightarrow \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \leq 0.$$

Note that this statement is verified if we can find $\alpha \geq 0$, $\beta \geq 0$ such that, for all x and u,

$$\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^TP+PA & PB \\ B^TP & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \alpha\left(\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 1\right)$$
$$+\beta\left(\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + 1\right) \leq 0 \tag{2}$$

or, equivalently:

$$\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^TP+PA+\alpha P & PB \\ B^TP & -\beta I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \leq \alpha - \beta.$$

This inequality holds for all x and u if and only if

$$\begin{bmatrix} A^TP+PA+\alpha P & PB \\ B^TP & -\beta I \end{bmatrix} \leq 0 \tag{3}$$
$$\beta - \alpha \leq 0. \tag{4}$$

Let $\beta = \alpha$ which is the best choice to satisfy (3) without violating (4):

$$\begin{bmatrix} A^TP+PA+\alpha P & PB \\ B^TP & -\alpha I \end{bmatrix} \leq 0. \tag{5}$$

Summary: procedure to overapproximate the reachable set

Look for $P = P^T > 0$ and $\alpha > 0$ satisfying the matrix inequality (5). This in not a linear matrix inequality (LMI) in α and P, but it is an LMI in P if α is fixed. The resulting ellipsoid $\{x : x^TPx \leq 1\}$ is a superset of R_T.

Additional objectives can be incorporated, such as minimizing the volume of the ellipsoid, which is proportional to $\sqrt{\det P^{-1}}$:

$$\text{minimize} \log(\det P^{-1}) \text{ which is convex in } P.$$

S-procedure

The principle used to obtain (2) is known as the S-procedure in control theory. To show that:

$$q_0(\xi) \geq 0 \quad \text{whenever} \quad q_i(\xi) \geq 0 \quad i = 1, 2, \ldots, p$$

look for $\tau_1, \tau_2, \ldots, \tau_p \geq 0$ such that

$$q_0(\xi) - \sum_{i=1}^{p} \tau_i q_i(\xi) \geq 0.$$

In (2), $q_i(\cdot)$, $i = 0, 1, 2$, are quadratic functions of $\xi = \begin{bmatrix} x \\ u \end{bmatrix}$.

Reachable sets with unit energy inputs

$$\boxed{R_T \triangleq \{x(T) \mid \dot{x} = f(x,u),\ x(0) = 0,\ \int_0^T u^T(t)u(t)dt \le 1\}} \qquad (6)$$

For an overapproximation, find positive definite $V(\cdot)$ such that

$$\nabla V(x) \cdot f(x,u) \le u^T u.$$

$$\begin{aligned} \frac{d}{dt}V(x(t)) \le u^T u \ &\Rightarrow\ V(x(T)) - V(x(0)) \le \int_0^T u^T(t)u(t)dt \le 1 \\ &\Rightarrow\ V(x(T)) \le 1. \end{aligned}$$

Therefore, $x \in R_T$ implies $V(x) \le 1$, *i.e.*, the level set contains the reachable set:

$$R_T \subset \{x : V(x) \le 1\}.$$

Example:

$$\dot{x} = Ax + Bu \quad V(x) = x^T P x.$$

Find $P = P^T > 0$ such that

$$x^T(A^T P + PA)x + x^T PBu + u^T B^T Px \le u^T u$$

or, written more compactly:

$$\begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \le \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.$$

This means

$$\begin{bmatrix} A^T P + PA & PB \\ B^T P & -I \end{bmatrix} \le 0$$

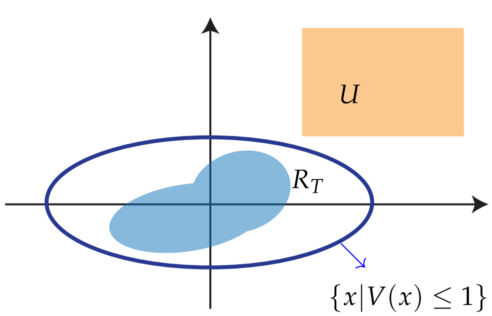
which is a LMI in P.

Safety Certification

Given an "unsafe" set U, show that

$$R_T \cap U = \emptyset.$$

The level set overapproximations above can be used to prove safety:



Look for a V with the additional property that $x \in U \Rightarrow V(x) > 1$. Such functions V are sometimes called "barrier functions."

Example: Suppose the unsafe set is the half-space:

$$U = \{x \,:\, a^T x > 1\}.$$

Let $V(x) = x^T P x$. From the S-procedure, if there exists $\tau > 0$ such that

$$(x^T P x - 1) - \tau(a^T x - 1) \geq 0, \tag{7}$$

then $x \in U \Rightarrow V(x) > 1$.

Exercise: Show that (7) is equivalent to: $P \geq aa^T$.

Thus, the LMIs in the previous examples can be augmented with this additional constraint to certify safety.

EE C222/ME C237 - Spring'18 - Lecture 16 Notes[1]

Murat Arcak

March 19 2018

Sum of Squares Programming

Establishing nonnegativity of functions is critical in nonlinear system analysis, *e.g.*, a Lyapunov function V for $\dot{x} = f(x)$ must satisfy

$$V(x) > 0 \quad \forall x \neq 0 \tag{1}$$

$$-\nabla V(x)^T f(x) \geq 0 \quad \forall x. \tag{2}$$

For $f(x) = Ax$ and $V(x) = x^T P x$, the conditions above are simple matrix inequalities:

$$P > 0, \quad -A^T P - PA \geq 0.$$

How can we check nonnegativity when f and V are more general polynomials?

Sum of Squares (SOS) Polynomials

A *monomial* is a product of powers of variables (*e.g.*, $m(x) = x_1^2 x_2$) and its degree is the sum of its exponents (*e.g.*, 3 for $m(x) = x_1^2 x_2$).

A *polynomial* is a finite linear combination of monomials and its degree is the maximum degree of these monomials.

Example 1: The polynomial

$$q(x_1, x_2) = x_1^2 - 2x_1 x_2^2 + 2x_1^4 + 2x_1^3 x_2 - x_1^2 x_2^2 + 6x_2^4 \tag{3}$$

has degree 4.

Definition: A polynomial p is a *sum of squares* (SOS) if there exist polynomials $g_1, \cdots, g_r$ such that

$$p = \sum_{i=1}^{r} g_i^2. \tag{4}$$

A SOS polynomial $p(x)$ is nonnegative for all x. The converse is not true: there exist nonnegative polynomials that are not SOS.

The polynomial $q(x_1, x_2)$ in (3) is SOS because it can be rewritten as:

$$(x_1 - x_2^2)^2 + \frac{1}{2}\left(2x_1^2 + x_1 x_2 - 3x_2^2\right)^2 + \frac{1}{2}\left(3x_1 x_2 + x_2^2\right)^2. \tag{5}$$

You can verify the equivalence of (3) and (5) by multiplying out terms in (5) and matching them to those in (3).

How a SOS decomposition like (5) can be obtained is discussed next.

SOS Decomposition

Let $z(x)$ be the vector of all monomials of degree $\leq d$ in n variables[2]:

$$z(x) \triangleq [1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^d]^T.$$

Then any polynomial with degree $\leq 2d$ can be rewritten as

$$p(x) = z(x)^T Q\, z(x) \tag{6}$$

where Q is a symmetric matrix.

[2] The length of this vector is $l_{[n,d]} := \begin{pmatrix} n+d \\ d \end{pmatrix}$.

Example 2: Let $p(x_1, x_2) = 2x_1^2x_2^2$ which has degree 4. With $n = 2$ and $d = 2$,

$$z(x) = [1, x_1, x_2, x_1^2, x_1x_2, x_2^2]^T, \tag{7}$$

and (6) holds with either

$$Q_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{or} \quad Q_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Thus, the choice of Q is not unique.

Theorem: A polynomial p with degree $\leq 2d$ is SOS if and only if there exists $Q = Q^T \geq 0$ satisfying (6).

Proof: (only if) If p is SOS then, by definition, $p = \sum_{i=1}^r g_i^2$ for some polynomials g_i, $i = 1, \cdots, r$. Write g_i as:

$$g_i(x) = C_i z(x) \tag{8}$$

where C_i is a row vector of coefficients. Then $g_i^2 = z^T C_i^T C_i z$ and

$$p = \sum_{i=1}^r g_i^2 = z^T \underbrace{\left(\sum_{i=1}^r C_i^T C_i\right)}_{Q \geq 0} z.$$

(if) Given $Q = Q^T \geq 0$ satisfying (6), decompose Q as $Q = C^T C$ where C has as many rows as the rank of Q, say r. Then,

$$Q = C^T C = \sum_{i=1}^r C_i^T C_i$$

where C_i is the ith row. If we define g_i as in (8), then $z^T Q z = \sum_{i=1}^r g_i^2$. $\square$

Since Q is not unique, not all Q satisfying (6) will certify SOS. In Example 2 above, $Q_1 \geq 0$ but Q_2 is indefinite. We need to characterize the set of all Q satisfying (6) and search for a $Q \geq 0$ in this set.

Parameterization of all matrices Q satisfying (6):

Find a particular solution Q_0 such that

$$p(x) = z(x)^T Q_0 z(x),$$

and find a basis of symmetric matrices N_j, $j = 1, 2, \cdots, K$, such that[3]

$$z(x)^T N_j z(x) = 0 \quad \text{for all } x. \tag{9}$$

[3] There are $K = \frac{l_{[n,d]}(l_{[n,d]}+1)}{2} - l_{[n,2d]}$ such matrices.

Then we can parameterize the set of all Q satisfying (6) as

$$Q = Q_0 + \sum_{j=1}^{K} \lambda_j N_j \quad \lambda_j \in \mathbb{R},$$

and p is SOS if and only if there exist $\lambda_1, \cdots, \lambda_K$ such that $Q \geq 0$.

For $n = d = 2$, $z(x)$ is as defined in (7) and a basis as in (9) is:

$$N_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad N_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$N_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad N_4 = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$N_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad N_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}.$$

<u>Example 1 revisited:</u> For $q(x_1, x_2)$ in (3), a suitable choice for Q_0 is

$$Q_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 6 \end{bmatrix}.$$

Note that $Q_0 \not\geq 0$, but $Q_0 + 6N_6 \geq 0$. Moreover, $Q_0 + 6N_6$ can be decomposed as

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{bmatrix}$$

which explains how the SOS form (5) was obtained.

Synthesizing SOS Polynomials

With the method above we can numerically check whether a given polynomial function V satisfies (1)-(2). However, in practice, it is more important to be able to *search* for a V satisfying (1)-(2). This is accomplished by synthesizing V as a weighted sum of basis polynomials with weights left as decision variables.

This leads to the following SOS synthesis problem:

Given basis polynomials p_i, $i = 0, 1, \cdots, m$, each with degree $\leq 2d$, find parameters $a_1, \cdots, a_m$ such that $p_0 + a_1 p_1 + \cdots + a_m p_m$ is SOS.

To solve this problem, find a matrix Q_i satisfying $p_i = z^T Q_i z$ for each $i = 0, 1, \cdots, m$. Then search for $a_1, \cdots, a_m$ and $\lambda_1, \cdots, \lambda_K$ satisfying

$$Q_0 + \sum_{i=1}^{m} a_i Q_i + \sum_{j=1}^{K} \lambda_j N_j \geq 0. \tag{10}$$

This is a linear matrix inequality (LMI) and can be solved numerically with standard semidefinite program (SDP) solvers.

There are also software packages[4] that follow the procedures above to automatically convert SOS programs to LMIs, such as (10).

[4] *e.g.*, SOSOPT

EE C222/ME C237 - Spring'18 - Lecture 17 Notes[1]

Murat Arcak

March 21 2018

[1]

Review of Sum of Squares (SOS) Polynomials

Checking whether a polynomial is SOS

A polynomial p with degree $\leq 2d$ is a sum of squares if and only if there exists $Q = Q^T \geq 0$ s.t.

$$p(x) = z(x)^T Q\, z(x) \tag{1}$$

where $z(x)$ is the vector of all monomials of degree $\leq d$:

$$z(x) \triangleq [1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^d]^T.$$

Find a particular solution Q_0 such that

$$p(x) = z(x)^T Q_0 z(x),$$

and find a basis of symmetric matrices N_j, $j = 1, 2, \cdots, K$, such that

$$z(x)^T N_j z(x) = 0 \quad \text{for all } x. \tag{2}$$

Then p is SOS if and only if there exist reals $\lambda_1, \cdots, \lambda_K$ such that

$$Q = Q_0 + \sum_{j=1}^{K} \lambda_j N_j \geq 0. \tag{3}$$

This is a linear matrix inequality (LMI) and can be solved numerically with standard semidefinite program (SDP) solvers.

Synthesizing SOS Polynomials

Given p_i, $i = 0, 1, \cdots, m$, each with degree $\leq 2d$, find reals $a_1, \cdots, a_m$ s.t. $p_0 + a_1 p_1 + \cdots + a_m p_m$ is SOS.

Find a particular Q_i satisfying $p_i = z^T Q_i z$ for each $i = 0, 1, \cdots, m$. Then search for $a_1, \cdots, a_m$ and $\lambda_1, \cdots, \lambda_K$ satisfying the LMI

$$Q_0 + \sum_{i=1}^{m} a_i Q_i + \sum_{j=1}^{K} \lambda_j N_j \geq 0. \tag{4}$$

Applications

Searching for a Lyapunov Function

Given $\dot{x} = f(x)$, $f(0) = 0$, where f is a vector of polynomials, search for a Lyapunov function of the form

$$V(x) = p_0(x) + a_1 p_1(x) + \cdots + a_m p_m(x) \tag{5}$$

where p_i, $i = 0, 1, \cdots, m$ are basis polynomials selected ahead of time, and a_i, $i = 1, \cdots, m$ are weights to be determined.

To ensure V is positive definite, pick a positive definite polynomial ℓ (e.g., $\ell(x) = \varepsilon x^T x$ for some small ε) and impose the constraint:

$$V(x) - \ell(x) \text{ is SOS.} \tag{6}$$

To ensure $\nabla V(x)^T f(x)$ is negative semidef., impose the constraint:

$$-\nabla V(x)^T f(x) \text{ is SOS.} \tag{7}$$

Constraints (6) and (7) can be brought to the LMI form (4) and feasible a_i, $i = 1, \cdots, m$ can be determined numerically (if they exist).

Overapproximating Reachable Sets

Recall from Lecture 15 that

$$R_T \triangleq \left\{ x(T) \mid \dot{x} = f(x, u),\ x(0) = 0,\ \int_0^T u^T(t) u(t) dt \leq 1 \right\} \tag{8}$$

defines the reachable set from $x(0) = 0$ under unit energy inputs and, if we can find a positive definite V such that

$$\nabla V(x)^T f(x, u) \leq u^T u, \tag{9}$$

then we can overapproximate R_T by:

$$R_T \subset \{x : V(x) \leq 1\}.$$

This follows because, from (9),

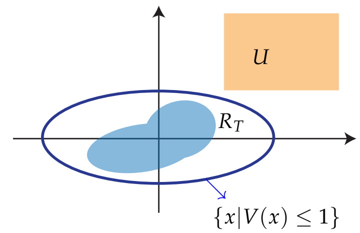
$$\begin{aligned} \frac{d}{dt} V(x(t)) \leq u^T u \ &\Rightarrow\ V(x(T)) - V(x(0)) \leq \int_0^T u^T(t) u(t) dt \leq 1 \\ &\Rightarrow\ V(x(T)) \leq 1. \end{aligned}$$

If $f(x, u)$ is a vector of polynomials in x and u, we can search for a polynomial V of the form (5), and encode (9) with the constraint:

$$-\nabla V(x)^T f(x, u) + u^T u \text{ is SOS in } x \text{ and } u. \tag{10}$$

This can then be combined with (6) and brought to the LMI form (4).

Certifying Safety



If unsafe set U does not intersect the overapproximation above, then it can't intersect the actual reachable set. Thus, we can certify safety by proving the implication:

$$x \in U \quad \Rightarrow \quad V(x) \geq 1 + \varepsilon \tag{11}$$

for some $\varepsilon > 0$.

Suppose the unsafe set can be expressed as

$$U = \{x : q_i(x) \geq 0, \, i = 1, \cdots, p\}$$

where q_i are polynomials. Then we can encode (11) with the constraints:

$$V(x) - (1 + \varepsilon) - \sum_{i=1}^{p} s_i(x) q_i(x) \text{ is SOS} \tag{12}$$

$$s_i(x), \; i = 1, \cdots, p \text{ are SOS.} \tag{13}$$

We can parameterize the search space for s_i as we did for V in (5), and combine (6), (10), (12)-(13) into a LMI.

Above we implicitly used a generalization of the S-procedure from Lecture 15. Specifically, to prove that

$$q_0(x) \geq 0 \quad \text{whenever} \quad q_i(x) \geq 0, \; i = 1, 2, \ldots, p$$

we look for nonnegative functions $s_1, s_2, \ldots, s_p$ (rather than constants as in Lecture 15) such that

$$q_0(x) - \sum_{i=1}^{p} s_i(x) q_i(x) \geq 0.$$

Underapproximating the Region of Attraction

Given system $\dot{x} = f(x)$ with asymptotically stable equilibrium at the origin $x = 0$, the region of attraction, denoted R_A, is the set of initial conditions from which the trajectories converge to the origin.

Recall from Lecture 10 that, if V is positive definite and

$$\nabla V(x)^T f(x) < 0 \quad \text{whenever } x \neq 0 \text{ and } V(x) \leq \gamma, \tag{14}$$

then $\Omega_\gamma \triangleq \{x : V(x) \leq \gamma\} \subset R_A$.

Let ℓ be a positive definite polynomial. If there exists a SOS polynomial s such that

$$-[\ell(x) + \nabla V(x)^T f(x)] - s(x)[\gamma - V(x)] \text{ is SOS,} \tag{15}$$

then $V(x) \le \gamma$ implies $\nabla V(x)^T f(x) \le -\ell(x)$ as stipulated in (14).

To obtain a LMI from (15), one option is to fix the Lyapunov function[2] V and to parameterize the search space for s. We can further maximize γ subject to (15) by incrementing γ until the the resulting LMI is infeasible.

[2] choose, *e.g.*, a quadratic Lyapunov function for the linearized model at $x = 0$

Alternatively s can be fixed and V parameterized. If we parameterize both s and V, however, (15) is no longer affine in the parameters because the term $s(x)V(x)$ contains the products of these parameters.

Below is a procedure that alternates between first fixing V, varying s, and next fixing s, varying V. When a new V is obtained, however, the shape of the level set changes and it may be ambiguous whether the new one is bigger. To remove this ambiguity we define a "shape function" p and use its level sets to judge the size of the region of attraction estimate.

Step 1: Let $V_0(x)$ be an initial choice for a Lyapunov function, *e.g.*, a quadratic function for the linearized model at the origin. Find

$$\gamma^* := \max \gamma \quad \text{s.t.} \quad \nabla V_0(x)^T f(x) < 0 \text{ whenever } x \neq 0 \text{ and } V_0(x) \le \gamma.$$

To satisfy the constraint look for a SOS multiplier $s_1(x)$ that satisfies

$$-[\ell(x) + \nabla V_0(x)^T f(x)] - s_1(x)[\gamma - V_0(x)] \text{ is SOS}$$

where ℓ is positive definite, *e.g.*, $\ell(x) := \epsilon(x_1^2 + x_2^2)$ for some $\epsilon > 0$.

Step 2: Let $p(x)$ be some fixed, positive definite convex polynomial (*e.g.*, $p(x) = x_1^2 + x_2^2$), and let $V_0(x)$ and γ^* be as in Step 1. Find

$$\beta^* := \max \beta \quad \text{s.t.} \quad V_0(x) \le \gamma^* \text{ whenever } p(x) \le \beta.$$

To satisfy the constraint look for a SOS multiplier $s_2(x)$ such that

$$[\gamma^* - V_0(x)] - s_2(x)[\beta - p(x)] \text{ is SOS.}$$

This means that $\{x : p(x) \le \beta\}$ is contained in $\{x : V_0(x) \le \gamma^*\}$.

Step 3: Given $\gamma^*, s_1(x)$ from Step 1 and $p(x), s_2(x)$ from Step 2, search for $V(x)$ to solve:

$$\begin{aligned} \max_{\substack{\beta>0 \\ \text{4th-order } V(x)}} \quad & \beta \\ \text{subject to} \quad & V(x) - \ell(x) \text{ is SOS} \\ & -[\ell(x) + \nabla V(x)^T f(x)] - s_1(x)[\gamma^* - V(x)] \text{ is SOS} \\ & [\gamma^* - V(x)] - s_2(x)[\beta - p(x)] \text{ is SOS.} \end{aligned}$$

The first constraint ensures V is positive definite. The second implies that the level set $\{x : V(x) \leq \gamma^*\}$ is invariant, hence a valid approximation for the region of attraction. The third constraint and the maximization of β ensure that V is selected such that the level set $\{x : V(x) \leq \gamma^*\}$ is as large as possible, as measured by function p.

To proceed, replace $V_0(x)$ in Step 1 with the function $V(x)$ from Step 3, and repeat the steps above for several iterations, until the change in β^* in Step 2 is sufficiently small. The final approximation of the ROA is the set where $V(x) \leq \gamma^*$.

EE C222/ME C237 - Spring'18 - Lecture 18 Notes[1]

Murat Arcak

April 2 2018

Feedback Linearization

Sastry, Chapter 9; Khalil, Chapter 13

Today: Relative degree, input-output linearization, zero dynamics

Consider the single-input single-output (SISO) nonlinear system:

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x). \end{aligned} \tag{1}$$

Relative degree (informal definition): Number of times we need to take the time derivative of the output to see the input:

$$\dot{y} = \underbrace{\frac{\partial h}{\partial x} f(x)}_{=:\, L_f h(x)} + \underbrace{\frac{\partial h}{\partial x} g(x)}_{=:\, L_g h(x)} u$$

$L_f h$ is called the *Lie derivative* of h along the vector field f

If $L_g h(x) \neq 0$ in an open set containing the equilibrium, then the relative degree is equal to 1. If $L_g h(x) \equiv 0$, continue taking derivatives:

$$\ddot{y} = \underbrace{L_f L_f h(x)}_{=:\, L_f^2 h(x)} + L_g L_f h(x) u.$$

If $L_g L_f h(x) \neq 0$, then relative degree is 2. If $L_g L_f h(x) \equiv 0$, continue.

Definition: The system (1) has *relative degree* r if, in a neighbourhood of the equilibrium,

$$\begin{aligned} &L_g L_f^{i-1} h(x) = 0 \quad i = 1, 2, \ldots, r-1 \\ &L_g L_f^{r-1} h(x) \neq 0. \end{aligned} \tag{2}$$

Examples:

1. $$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1^3 + u \\ y &= x_1 \end{aligned} \tag{3}$$

   has relative degree = 2.

2. SISO linear system:

$$\dot{x} = Ax + Bu \quad y = Cx$$

$L_g h(x) = CB,\;\; L_g L_f h(x) = CAB,\;\; \dots,\;\; L_g L_f^{r-1} = CA^{r-1}B.$

$CB \neq 0 \;\Rightarrow\;$ relative degree = 1

$CB = 0,\;\; CAB \neq 0 \;\Rightarrow\;$ relative degree = 2

$CB = \cdots = CA^{r-2}B = 0,\;\; CA^{r-1}B \neq 0 \;\Rightarrow\;$ relative degree = r

The parameters $CA^{i-1}B \;\; i = 1, 2, 3, \dots$ are called *Markov parameters* and are invariant under similarity transformations.

3.
$$\begin{aligned} \dot{x}_1 &= x_2 + x_3^3 & \qquad y &= x_1 \\ \dot{x}_2 &= x_3 & \qquad \dot{y} &= \dot{x}_1 = x_2 + x_3^3 \\ \dot{x}_3 &= u & \qquad \ddot{y} &= \dot{x}_2 + 3x_3^2\dot{x}_3 = x_3 + 3x_3^2 u \end{aligned}$$

$L_g L_f h(x) = 3x_3^2 = 0$ when $x_3 = 0$, and $\neq 0$ elsewhere. Thus, this system does not have a well-defined relative degree around $x = 0$.

Input-Output Linearization

If a system has a well-defined relative degree then it is input-output linearizable:

$$y^{(r)} = L_f^r h(x) + \underbrace{L_g L_f^{r-1} h(x)}_{\neq 0} u$$

Apply preliminary feedback:

$$\boxed{u = \frac{1}{L_g L_f^{r-1} h(x)} \Big(- L_f^r h(x) + v \Big)} \tag{4}$$

where v is a new input to be designed. Then, $y^{(r)} = v$ is a linear system in the form of an integrator chain:

$$\begin{aligned} \dot{\zeta}_1 &= \zeta_2 \\ \dot{\zeta}_2 &= \zeta_3 \\ &\vdots \\ \dot{\zeta}_r &= v \end{aligned}$$

where $\zeta_1 =: y = h(x)$, $\zeta_2 =: \dot{y} = L_f h(x), \dots, \zeta_r =: y^{(r-1)} = L_f^{r-1} h(x)$. To ensure $y(t) \to 0$ as $t \to \infty$, apply the feedback:

$$\boxed{\begin{aligned} v &= -k_1\zeta_1 - k_2\zeta_2 - \cdots - k_r\zeta_r \\ &= -k_1 h(x) - k_2 L_f h(x) - \cdots - k_r L_f^{r-1} h(x) \end{aligned}} \tag{5}$$

where $k_1, \dots, k_r$ are such that $s^r + k_r s^{r-1} + \cdots + k_2 s + k_1$ has all roots in the open left half-plane.

Does the controller (4)-(5) achieve asymptotic stability of $x = 0$?

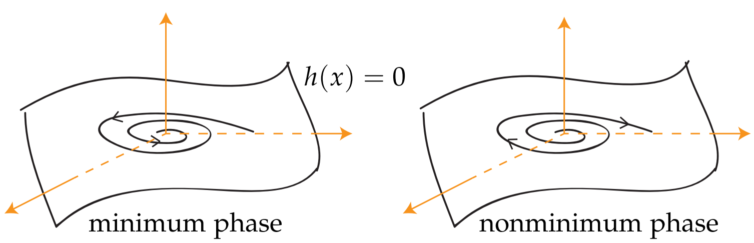
Not necessarily! It renders the $(n-r)$-dimensional manifold:

$$h(x) = L_f h(x) = \cdots = L_f^{r-1} h(x) = 0$$

invariant and attractive. The dynamics restricted to this manifold are called zero dynamics and determine whether or not $x = 0$ is stable.

If the origin of the zero dynamics is asymptotically stable, the system is called minimum phase. If unstable, it is called nonminimum phase.

Example: $n = 3, \quad r = 1$



Finding the Zero Dynamics

Set $y = \dot{y} = \cdots = y^{(r-1)} = 0$ and substitute (4) with $v = 0$, that is:

$$u^* = \frac{-L_f^r h(x)}{L_g L_f^{r-1} h(x)}.$$

The remaining dynamical equations describe the zero dynamics.

Example:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \alpha x_3 + u \\ \dot{x}_3 &= \beta x_3 - u \\ y &= x_1 \end{aligned} \tag{6}$$

This system has relative degree 2. With $x_1 = x_2 = 0$ and $u^* = -\alpha x_3$, the remaining dynamical equation is

$$\dot{x}_3 = (\alpha + \beta) x_3.$$

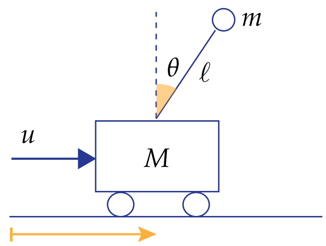
Thus this system is minimum phase if $\alpha + \beta < 0$.

For a linear SISO system, *relative degree* is the difference between the degrees of the denominator and the numerator of the transfer function, and *zeros* are the roots of the numerator. The definitions of relative degree and zero dynamics above generalize these concepts to nonlinear systems. As an example, the transfer function for (6) is

$$\frac{s - (\alpha + \beta)}{s^2 (s - \beta)},$$

which has relative degree two and a zero at $s = \alpha + \beta$ as expected.

Example: Cart/Pole



$$\ddot{y} = \frac{1}{\frac{M}{m} + \sin^2\theta}\left(\frac{u}{m} + \dot{\theta}^2\ell\sin\theta - g\sin\theta\cos\theta\right)$$
$$\ddot{\theta} = \frac{1}{\ell(\frac{M}{m} + \sin^2\theta)}\left(-\frac{u}{m}\cos\theta - \dot{\theta}^2\ell\cos\theta\sin\theta + \frac{M+m}{m}g\sin\theta\right) \tag{7}$$

Relative degree = 2.

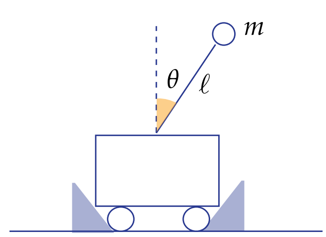
To find the zero dynamics, substitute $y = \dot{y} = 0$, and

$$u^* = -m(\dot{\theta}^2\ell\sin\theta - g\sin\theta\cos\theta)$$

in the $\ddot{\theta}$ equation:

$$\boxed{\ddot{\theta} = \frac{g}{\ell}\sin\theta.}$$

Same as the dynamics of the pole when the cart is held still:



Nonminimum phase because $\theta = 0$ is unstable for the zero dynamics.

EE222 - Spring'18 - Lecture 19 Notes[1]

Murat Arcak

April 4 2018

Feedback Linearization (continued)

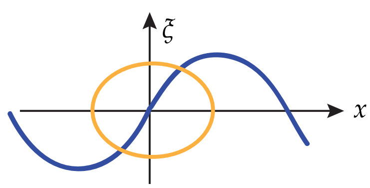
Nonlinear Changes of Variables

$T : \mathbb{R}^n \to \mathbb{R}^n$ is called a *diffeomorphism* if its inverse T^{-1} exists, and both T and T^{-1} are continuously differentiable (C^1).

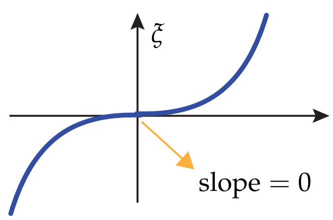
Examples:

1. $\xi = Tx$ is a diffeomorphism if T is a nonsingular matrix

2. $\xi = \sin x$ is a local diffeomorphism around $x = 0$, but not global



3. $\xi = x^3$ is not a diffeomorphism because $T^{-1}(\cdot)$ is not C^1 at $\xi = 0$



How to check if $\xi = T(x)$ is a local diffeomorphism?

Implicit Function Theorem

Suppose $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is C^1 and there exists $x_0 \in \mathbb{R}^n$, $\xi_0 \in \mathbb{R}^m$ such that

$$f(x_0, \xi_0) = 0.$$

If $\frac{\partial f}{\partial x}(x_0, \xi_0)$ is nonsingular, then in a neighborhood of (x_0, ξ_0),

$$f(x, \xi) = 0$$

has a unique solution $x = g(\xi)$ where g is C^1 at $\xi = \xi_0$.

Corollary: Let $f(x, \xi) = T(x) - \xi$. If $\frac{\partial T}{\partial x}$ is nonsingular at x_0, then $T(\cdot)$ is a local diffeomorphism around x_0.

A "Normal Form" that Explicitly Displays the Zero Dynamics

Theorem: If $\dot{x} = f(x) + g(x)u$, $y = h(x)$ has a well-defined relative degree $r \leq n$, then there exist a diffeomorphism $T : x \to \begin{bmatrix} z \\ \zeta \end{bmatrix}$, $z \in \mathbb{R}^{n-r}$, $\zeta \in \mathbb{R}^r$, that transforms the system to the form:

$$\boxed{\begin{aligned} \dot{z} &= f_0(z, \zeta) \\ \dot{\zeta}_1 &= \zeta_2 \\ &\vdots \\ \dot{\zeta}_r &= b(z, \zeta) + a(z, \zeta)u, \quad y = \zeta_1. \end{aligned}} \tag{1}$$

In particular, $\dot{z} = f_0(z, 0)$ represents the zero dynamics. $\square$

To obtain this form, let $\zeta = [h(x) \quad L_f h(x) \quad \ldots \quad L_f^{r-1}h(x)]^T$, and find $n - r$ independent variables z such that $\dot{z}$ does not contain u.

Note that the terms $b(z, \zeta)$ and $a(z, \zeta)$ correspond to $L_f^r(x)$ and $L_g L_f^{r-1}h(x)$ in the original coordinates.

Example:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \alpha x_3 + u \\ \dot{x}_3 &= \beta x_3 - u \\ y &= x_1. \end{aligned}$$

Let $\zeta_1 = x_1$, $\zeta_2 = x_2$, and note that $z = x_2 + x_3$ is independent of ζ_1, ζ_2, and $\dot{z}$ does not contain u. Thus, the normal form is:

$$\begin{aligned} \dot{z} &= (\alpha + \beta)x_3 = (\alpha + \beta)z - (\alpha + \beta)\zeta_2 \\ \dot{\zeta}_1 &= \zeta_2 \\ \dot{\zeta}_2 &= \alpha x_3 + u = \alpha z - \alpha\zeta_2 + u. \end{aligned}$$

I/O Linearizing Controller in the new coordinates (1):

$$u = \frac{1}{a(z, \zeta)}\Big(- b(z, \zeta) + v\Big) \tag{2}$$

$$v = -k_1\zeta_1 \cdots - k_r\zeta_r \tag{3}$$

where $k_1, \cdots, k_r$ are such that all roots of $s^r + k_r s^{r-1} + \cdots + k_2 s + k_1$ have negative real parts.

Theorem: If $z = 0$ is locally exponentially stable for the zero dynamics $\dot{z} = f_0(z, 0)$, then (2)–(3) locally exponentially stabilizes $x = 0$.

Proof: Closed-loop system:

$$\begin{aligned} \dot{z} &= f_0(z, \zeta) \\ \dot{\zeta} &= A\zeta \end{aligned}$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & \\ 0 & 0 & 1 & \dots & \\ & & & \ddots & \\ & & & & 1 \\ -k_1 & -k_2 & -k_3 & \dots & -k_r \end{bmatrix}$$

is Hurwitz. The Jacobian linearization at $(z, \zeta) = 0$ is:

$$J = \begin{bmatrix} \frac{\partial f_0}{\partial z}(0,0) & \frac{\partial f_0}{\partial \zeta}(0,0) \\ 0 & A \end{bmatrix}$$

where $\frac{\partial f_0}{\partial z}(0,0)$ is Hurwitz since $\dot{z} = f_0(z,0)$ is exponentially stable by the proposition in Lecture 11, page 1. Since A is also Hurwitz, all eigenvalues of J have negative real parts $\Rightarrow$ exponential stability.

Global asymptotic stability can be guaranteed with additional assumptions on the zero dynamics, such as ISS of

$$\dot{z} = f_0(z, \zeta)$$

with respect to the input ξ:

$$\boxed{\dot{\zeta} = A\zeta} \xrightarrow{\ \zeta\ } \boxed{\dot{z} = f_0(z,\zeta)}$$

<u>Example:</u> $\quad \dot{z} = -z + z^2\zeta, \quad \dot{\zeta} = -k\zeta$

Note: the z subsystem is not ISS

$(z, \zeta) = 0$ is locally exponentially stable, but not globally: solutions escape in finite time for large $z(0)$.

I/O Linearizing Controller for Tracking

For the output $y(t)$ to track a reference signal[2] $y_d(t)$, replace (3) with:

[2] assumed to be r times differentiable

$$v = -k_1(\zeta_1 - y_d(t)) - k_2(\zeta_2 - \dot{y}_d(t)) \cdots - k_r(\zeta_r - y_d^{(r-1)}(t)) + y_d^{(r)}(t)$$

Let $e_1 \triangleq \zeta_1 - y_d(t)$, $e_2 \triangleq \zeta_2 - \dot{y}_d(t)$, $\dots$, $e_r \triangleq \zeta_r - y_d^{(r-1)}(t)$. Then:

$$\left.\begin{aligned} \dot{e}_1 &= e_2 \\ \dot{e}_2 &= e_3 \\ &\vdots \\ \dot{e}_r &= v - y_d^{(r)}(t) = -k_1 e_1 - \cdots - k_r e_r \end{aligned}\right\} \quad \dot{e} = Ae.$$

Thus $e(t) \to 0$, that is $y(t) - y_d(t) \to 0$.

If $y_d(t)$ and its derivatives are bounded, then $\zeta(t)$ is bounded. If the zero dynamics $\dot{z} = f_0(z, \zeta)$ is ISS with respect to ζ, then $z(t)$ is also bounded. Thus, all internal signals are bounded.

EE C222/ME C237 - Spring'18 - Lecture 20 Notes[1]

Murat Arcak

April 16 2018

[1] Licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Full-State Feedback Linearization

The system $\dot{x} = f(x) + g(x)u$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, is (full state) feedback linearizable if a function $h : \mathbb{R}^n \mapsto \mathbb{R}$ exists such that the relative degree from u to $y = h(x)$ is n.

Since $r = n$, the normal form in Lecture 19 has no zero dynamics and

$$x \to \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{bmatrix} = \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{bmatrix}$$

is a diffeomorphism that transforms the system to the form:

$$\begin{aligned} \dot{\zeta}_1 &= \zeta_2 \\ \dot{\zeta}_2 &= \zeta_3 \\ &\vdots \\ \dot{\zeta}_n &= L_f^n h(x) + L_g L_f^{n-1} h(x) u. \end{aligned}$$

Then, the feedback linearizing controller

$$u = \frac{1}{L_g L_f^{n-1} h(x)} \Big(-L_f^n h(x) + v \Big), \quad v = -k_1 \zeta_1 \cdots - k_n \zeta_n,$$

yields the closed-loop system:

$$\dot{\zeta} = A\zeta \quad \text{where} \quad A = \begin{bmatrix} 0 & 1 & 0 & \dots & \\ 0 & 0 & 1 & \dots & \\ & & & \ddots & \\ & & & & 1 \\ -k_1 & -k_2 & -k_3 & \dots & -k_n \end{bmatrix}.$$

<u>Example:</u>

$$\begin{aligned} \dot{x}_1 &= x_2 + 2x_1^2 \\ \dot{x}_2 &= x_3 + u \\ \dot{x}_3 &= x_1 - x_3 \end{aligned}$$

The choice $y = x_3$ gives relative degree $r = n = 3$.

Let $\zeta_1 = x_3$, $\zeta_2 = \dot{x}_3 = x_1 - x_3$, $\zeta_3 = \ddot{x}_3 = \dot{x}_1 - \dot{x}_3 = x_2 + 2x_1^2 - x_1 + x_3$.

$$\begin{aligned} \dot{\zeta}_1 &= \zeta_2 \\ \dot{\zeta}_2 &= \zeta_3 \\ \dot{\zeta}_3 &= (4x_1 - 1)(x_2 + 2x_1^2) + x_1 + u. \end{aligned}$$

Feedback linearizing controller:

$$u = -(4x_1 - 1)(x_2 + 2x_1^2) - x_1 - k_1\zeta_1 - k_2\zeta_2 - k_3\zeta_3.$$

Summary so far:

I/O Linearization:	• suitable for tracking • output y is an intrinsic physical variable
Full state linearization:	• set point stabilization • output is not intrinsic, selected to enable a linearizing change of variables.

Remaining question:

- When is a system feedback linearizable, *i.e.*, how do we know whether a relative degree $r = n$ output exists?

Basic Definitions from Differential Geometry

Definition: The Lie bracket of two vector fields f and g is a new vector field defined as:

$$[f, g](x) = \frac{\partial g}{\partial x} f(x) - \frac{\partial f}{\partial x} g(x).$$

Note:

1. $[f, g] = -[g, f]$,
2. $[f, f] = 0$,
3. If f, g are constant then $[f, g] = 0$.

Notation for repeated applications:

$$[f, [f, g]] = \mathrm{ad}_f^2 g, \quad [f, [f, [f, g]]] = \mathrm{ad}_f^3 g, \quad \cdots$$

$$\mathrm{ad}_f^0 g(x) \triangleq g(x), \quad \mathrm{ad}_f^k g \triangleq [f, \mathrm{ad}_f^{k-1} g] \quad k = 1, 2, 3, \ldots$$

Definition: Given vector fields $f_1, \ldots, f_k$, a distribution Δ is defined as $\Delta(x) = \mathrm{span}\{f_1(x), \ldots, f_k(x)\}$.

$f \in \Delta$ means that there exist scalar functions $\alpha_i(x)$ such that

$$f(x) = \alpha_1(x) f_1(x) + \cdots + \alpha_k(x) f_k(x).$$

Definition: Δ is said to be nonsingular if $f_1(x), \ldots, f_k(x)$ are linearly independent for all x.

Definition: Δ is said to be involutive if

$$g_1 \in \Delta, g_2 \in \Delta \implies [g_1, g_2] \in \Delta$$

that is, Δ is closed under the Lie bracket operation.

Proposition: $\Delta = \text{span}\{f_1, \ldots, f_k\}$ is involutive if and only if

$$[f_i, f_j] \in \Delta \quad 1 \leq i, j \leq k.$$

Example 1: $\Delta = \text{span}\{f_1, \ldots, f_k\}$ where $f_1, \ldots, f_k$ are constant vectors

Example 2: a single vector field $f(x)$ is involutive since $[f, f] = 0 \in \Delta$

Definition: A nonsingular k-dimensional distribution

$$\Delta(x) = \text{span}\{f_1(x), \ldots, f_k(x)\} \quad x \in \mathbb{R}^n$$

is said to be completely integrable if there exist $n-k$ functions

$$\phi_1(x), \ldots, \phi_{n-k}(x)$$

such that

$$\frac{\partial \phi_i}{\partial x} f_j(x) = 0 \quad i = 1, \ldots, n-k, \quad j = 1, \ldots, k$$

and $d\Phi_i(x) := \dfrac{\partial \phi_i}{\partial x}$, $i = 1, \ldots, n-k$, are linearly independent.

Example 3: If $f_1, \ldots, f_k$ are linearly independent constant vectors, then we can find $n-k$ independent row vectors $T_1, \ldots, T_{n-k}$ s.t.

$$T_i[f_1 \ldots f_k] = 0.$$

Therefore, $\Delta = \text{span}\{f_1, \ldots, f_k\}$ is completely integrable and

$$\phi_i(x) = T_i x, \quad i = 1, \ldots, n-k.$$

Frobenius Theorem: A nonsingular distribution is completely integrable if and only if it is involutive.

Example 3 above is a special case since Δ is involutive by Example 1.

Back to (Full State) Feedback Linearization

Recall: $\dot{x} = f(x) + g(x)u$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}$ is feedback linearizable if we can find an output $y = h(x)$ such that relative degree $r = n$.

How do we determine if a relative degree $r = n$ output exists?

$$L_g h(x) = L_g L_f h(x) = \cdots = L_g L_f^{n-2} h(x) = 0 \text{ in a nbhd of } x_0 \tag{1}$$

$$L_g L_f^{n-1} h(x_0) \neq 0. \tag{2}$$

Proposition:[2] (1)-(2) are equivalent to:

[2] follows from (5) below with $j = 0$

$$L_g h(x) = L_{\mathrm{ad}_f g} h(x) = \cdots = L_{\mathrm{ad}_f^{n-2} g} h(x) = 0 \text{ in a nbhd of } x_0 \tag{3}$$

$$L_{\mathrm{ad}_f^{n-1} g} h(x_0) \neq 0. \tag{4}$$

The advantage of (3) over (1) is that it has the form:

$$\frac{\partial h}{\partial x}[g(x) \quad \mathrm{ad}_f\, g(x) \quad \ldots \quad \mathrm{ad}_f^{n-2} g(x)] = 0$$

which is amenable to the Frobenius Theorem.

Theorem: $\dot{x} = f(x) + g(x)u$ is feedback linearizable around x_0 if and only if the following two conditions hold:

C1) $[g(x_0) \quad \mathrm{ad}_f\, g(x_0) \quad \ldots \quad \mathrm{ad}_f^{n-1} g(x_0)]$ has rank n

C2) $\Delta(x) = \mathrm{span}\{g(x), \mathrm{ad}_f\, g(x), \ldots, \mathrm{ad}_f^{n-2} g(x)\}$ is involutive in a neighborhood of x_0.

Proof: (if) Given C1 and C2 show that there exists $h(x)$ satisfying (3)-(4).

$\Delta(x)$ is nonsingular by C1 and involutive by C2. Thus, by the Frobenius Theorem, there exists $h(x)$ satisfying (3) and $dh(x) \neq 0$.

To prove (4) suppose, to the contrary, $L_{\mathrm{ad}_f^{n-1}} h(x_0) = 0$. This implies

$$dh(x_0)\underbrace{[g(x_0) \quad \mathrm{ad}_f\, g(x_0) \quad \ldots \quad \mathrm{ad}_f^{n-1} g(x_0)]}_{\text{nonsingular by C1}} = 0.$$

Thus $dh(x_0) = 0$, a contradiction.

(only if) Given that $y = h(x)$ with $r = n$ exists, that is (3)-(4) hold, show that C1 and C2 are true.

We will use the following fact[3] which holds when $r = n$:

[3] see, *e.g.*, Khalil, Lemma C.8

$$L_{\mathrm{ad}_f^i g} L_f^j h(x) = \begin{cases} 0 & \text{if } i + j \leq n-2 \\ (-1)^{n-1-j} L_g L_f^{n-1} h(x) \neq 0 & \text{if } i + j = n-1. \end{cases} \tag{5}$$

Define the matrix

$$M = \begin{bmatrix} dh \\ dL_f h \\ \vdots \\ dL_f^{n-1} h \end{bmatrix} \begin{bmatrix} g & -\mathrm{ad}_f\, g & \mathrm{ad}_f^2\, g & \ldots & (-1)^{n-1}\mathrm{ad}_f^{n-1} g \end{bmatrix} \tag{6}$$

and note that the (k,ℓ) entry is:

$$\begin{aligned} M_{k\ell} &= dL_f^{k-1}h(x)(-1)^{\ell-1}\,\mathrm{ad}_f^{\ell-1}\,g(x) \\ &= (-1)^{\ell-1}L_{\mathrm{ad}_f^{\ell-1}g}L_f^{k-1}h(x). \end{aligned}$$

$$\begin{bmatrix} 0 & 0 & \cdots & \star \\ 0 & & / & \vdots \\ \vdots & \star & & \vdots \\ \star & \cdots & \cdots & \star \end{bmatrix}$$

Then, from (5):

$$M_{k\ell} = \begin{cases} 0 & \ell + k \leq n \\ \neq 0 & \ell + k = n+1. \end{cases}$$

Since the diagonal entries are nonzero, M has rank n and thus the factor

$$\begin{bmatrix} g & -\mathrm{ad}_f\,g & \mathrm{ad}_f^2\,g & \ldots & (-1)^{n-1}\,\mathrm{ad}_f^{n-1}g \end{bmatrix}$$

in (6) must have rank n as well. Thus C1 follows.

This also implies $\Delta(x)$ is nonsingular; thus, by the Frobenius Thm,

$$\text{complete integrability} \;\equiv\; \text{involutivity}.$$

$\Delta(x)$ is completely integrable since $h(x)$ satisfying (3) exists by assumption; thus, we conclude involutivity (C2). $\square$

Example:

$$\begin{aligned} \dot{x}_1 &= x_2 + 2x_1^2 \\ \dot{x}_2 &= x_3 + u \\ \dot{x}_3 &= x_1 - x_3 \end{aligned}$$

Feedback linearizability was shown on page 1 by inspection: $y = x_3$ gives relative degree $= 3$. Verify with the theorem above:

$$f(x) = \begin{bmatrix} x_2 + 2x_1^2 \\ x_3 \\ x_1 - x_3 \end{bmatrix} \qquad g(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$[f,g](x) = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \qquad [f,[f,g]](x) = \begin{bmatrix} 4x_1 \\ 0 \\ 1 \end{bmatrix}$$

Conditions of the theorem:

1. $\begin{bmatrix} 0 & -1 & 4x_1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ full rank

2. $\Delta = \mathrm{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \right\}$ involutive

   $\dfrac{\partial h}{\partial x}\begin{bmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$ satisfied by $h(x) = x_3$.

EE C222/ME C237 - Spring'18 - Lecture 21 Notes[1]

Murat Arcak

April 18 2018

Feedback Linearization Continued

Recall "strict feedback systems" discussed in Lecture 13:

$$\begin{aligned}
\dot{x}_1 &= f_1(x_1) + g_1(x_1)x_2 \\
\dot{x}_2 &= f_2(x_1, x_2) + g_2(x_1, x_2)x_3 \\
\dot{x}_3 &= f_3(x_1, x_2, x_3) + g_3(x_1, x_2, x_3)x_4 \\
&\vdots \\
\dot{x}_n &= f_n(x) + g_n(x)u.
\end{aligned} \tag{1}$$

Such systems are feedback linearizable when $g_i(x_1, \ldots, x_i) \neq 0$ near the origin, $i = 1, 2, \cdots, n$, because the relative degree is n with the choice of output $y = h(x) = x_1$:

$$y^{(n)} = L_f^n h(x) + \underbrace{g_1(x_1)g_2(x_1, x_2) \cdots g_n(x)}_{L_g L_f^{n-1} h(x) \neq 0} u.$$

Feedback linearizability is lost when $g_i(0) = 0$ for some i; however, backstepping may be applicable as the following example illustrates:

<u>Example 1:</u>

$$\begin{aligned}
\dot{x}_1 &= x_1^2 x_2 \\
\dot{x}_2 &= u.
\end{aligned}$$

Treat x_2 as virtual control and let $\alpha_1(x_1) = -x_1$ which stabilizes the x_1-subsystem, as verified with Lyapunov function $V_1(x_1) = \frac{1}{2}x_1^2$. Then $z_2 := x_2 - \alpha_1(x_1)$ satisfies $\dot{z}_2 = u - \dot{\alpha}_1$, and

$$u = \dot{\alpha}_1 - \frac{\partial V_1}{\partial x_1} x_1^2 - k_2 z_2 = -x_1^2 x_2 - x_1^3 - k_2(x_2 + x_1)$$

achieves global asymptotic stability:

$$V = \frac{1}{2}x_1^2 + \frac{1}{2}z_2^2 \quad \Rightarrow \quad \dot{V} = -x_1{}^4 - k_2 z_2^2.$$

In contrast the system is not feedback linearizable, because condition (C1) in the theorem for feedback linearizability (Lecture 20, p.4) fails. To see this note that

$$f(x) = \begin{bmatrix} x_1^2 x_2 \\ 0 \end{bmatrix}, \quad g(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathrm{ad}_f\, g(x) = [f, g](x) = \begin{bmatrix} -x_1^2 \\ 0 \end{bmatrix},$$

thus, with $n = 2$ and $x_0 = 0$,

$$[g(x_0) \quad \mathrm{ad}_f\, g(x_0) \quad \ldots \quad \mathrm{ad}_f^{n-1} g(x_0)] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

which is rank deficient.

Multi-Input Multi-Output Systems

Consider now a MIMO system with m inputs and m outputs:

$$\begin{aligned} \dot{x} &= f(x) + \sum_{i=1}^{m} g_i(x) u_i \\ y_i &= h_i(x), \quad i = 1, \cdots, m. \end{aligned} \tag{2}$$

Let r_i denote the number of times we need to differentiate y_i to hit at least one input. Then,

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_m^{(r_m)} \end{bmatrix} = \underbrace{\begin{bmatrix} L_f^{r_1} h_1(x) \\ \vdots \\ L_f^{r_m} h_m(x) \end{bmatrix}}_{=:\, B(x)} + \underbrace{\begin{bmatrix} L_{g_1} L_f^{r_1-1} h_1(x) & \cdots & L_{g_m} L_f^{r_1-1} h_1(x) \\ \vdots & & \vdots \\ L_{g_1} L_f^{r_m-1} h_m(x) & \cdots & L_{g_m} L_f^{r_m-1} h_m(x) \end{bmatrix}}_{=:\, A(x)} \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}.$$

If $A(x)$ is nonsingular, then the feedback law

$$u = A(x)^{-1}(-B(x) + v)$$

input/output linearizes the system, creating m decoupled chains of integrators:

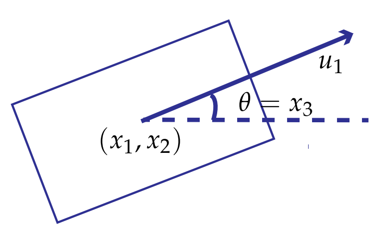
$$y_i^{(r_i)} = v_i, \quad i = 1, \ldots, m.$$

We say that the system has *vector relative degree* $\{r_1, \cdots, r_m\}$ if the matrix $A(x)$ defined above is nonsingular.

Example 2: The kinematic model of a unicycle, depicted below, is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} \cos x_3 \\ \sin x_3 \\ 0 \end{bmatrix} u_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u_2,$$

where u_1 is the speed and u_2 is the angular velocity.



Let $y_1 = x_1$ and $y_2 = x_2$, and note that

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \cos x_3 & 0 \\ \sin x_3 & 0 \end{bmatrix}}_{=:\, A(x)} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Since $A(x)$ is singular, the system does not have a well-defined vector relative degree. □

The notion of zero dynamics and the normal form can be extended to MIMO systems[2]. If the system has vector relative degree $\{r_1, \cdots, r_m\}$, then $r := r_1 + \cdots + r_m \leq n$ and

[2] see, *e.g.*, Sastry, Section 9.3

$$\zeta := [h_1(x)\ L_f h_1(x) \cdots L_f^{r_1-1} h_1(x)\ \cdots\ h_m(x)\ L_f h_m(x) \cdots L_f^{r_m-1} h_m(x)]^T$$

defines a partial set of coordinates. As in normal form discussed in Lecture 19, one can find $n - r$ additional functions $z_1(x), \cdots, z_{n-r}(x)$ so that $x \mapsto (z, \zeta)$ is a complete coordinate transformation.

Full-state feedback linearization amounts to finding m output functions $h_1, \cdots, h_m$ such that the system has vector relative degree $\{r_1, \cdots, r_m\}$ with $r_1 + \cdots + r_m = n$. Necessary and sufficient conditions for the existence of such functions, analogous to those in Lecture 20 for SISO systems, are available[3].

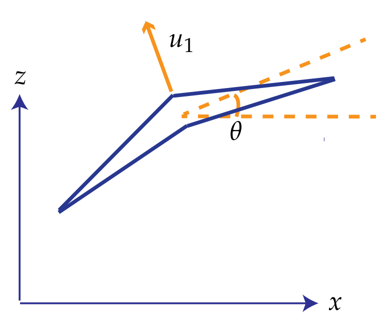
[3] see, *e.g.*, Sastry, Proposition 9.16

Example 3: Consider the following model of a *planar vertical take-off and landing* (PVTOL) aircraft[4]

[4] Sastry, Section 10.4.2

$$\begin{aligned} \ddot{x} &= -\sin(\theta) u_1 + \mu \cos(\theta) u_2 \\ \ddot{z} &= \cos(\theta) u_1 + \mu \sin(\theta) u_2 - 1 \\ \ddot{\theta} &= u_2, \end{aligned}$$

where μ is a constant that accounts for the coupling between the rolling moment and translational acceleration, and -1 in the second equation is the gravitational acceleration, normalized to unity by appropriately scaling the variables.



If we take x and z as the two outputs we get

$$\begin{bmatrix}\ddot{x}\\ \ddot{z}\end{bmatrix} = \begin{bmatrix}0\\ -1\end{bmatrix} + \underbrace{\begin{bmatrix}-\sin\theta & \mu\cos\theta\\ \cos\theta & \mu\sin\theta\end{bmatrix}}_{A(\theta)}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$

where $A(\theta)$ is invertible when $\mu \neq 0$:

$$A^{-1}(\theta) = \begin{bmatrix}-\sin\theta & \cos\theta\\ \frac{1}{\mu}\cos\theta & \frac{1}{\mu}\sin\theta\end{bmatrix}.$$

Thus the systems has vector relative degree $\{2,2\}$ when $\mu \neq 0$, and the input/output linearizing controller is

$$\begin{bmatrix}u_1\\ u_2\end{bmatrix} = \begin{bmatrix}-\sin\theta & \cos\theta\\ \frac{1}{\mu}\cos\theta & \frac{1}{\mu}\sin\theta\end{bmatrix}\left(\begin{bmatrix}0\\ 1\end{bmatrix} + \begin{bmatrix}v_1\\ v_2\end{bmatrix}\right).$$

The zero dynamics is obtained by substituting $u_2^* = \frac{1}{\mu}\sin\theta$, needed to maintain z at a constant value and $\dot{z}$ at zero, in the dynamical equation for θ:

$$\ddot{\theta} = \frac{1}{\mu}\sin\theta.$$

The system is nonminimum phase for $\mu > 0$, since $\theta = 0$ is unstable.

Drift-Free Systems

Suppose $f(x) = 0$ for all x in (2). Such system are called *drift-free* and encompass linear systems of the form

$$\dot{x} = Bu, \quad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m.$$

Assuming the columns of the $n \times m$ matrix B are linearly independent, we can find $n - m$ row vectors T_i, $i = 1, \cdots, n-m$, such that

$$T_i B = 0.$$

This means that $\phi_i(x) := T_i x$ satisfies

$$\frac{d}{dt}\phi_i(x(t)) = 0 \quad \Rightarrow \quad \phi_i(x(t)) = \phi_i(x(0)) \tag{3}$$

regardless of the control inputs. Since there are $n - m$ such constraints, controllability is not possible in drift-free linear systems with fewer control inputs than the state dimension ($m < n$).

The Frobenius Theorem (Lecture 20) implies that constraints of the form (3), called *holonomic constraints*, also exist for nonlinear drift-free systems

$$\dot{x} = \sum_{i=1}^{m} g_i(x)u_i \tag{4}$$

when the distribution $\Delta = \text{span}\{g_1, \cdots, g_m\}$ is nonsingular and involutive.

When Δ is non-involutive, however, controllability may be possible with $m < n$ – another essentially nonlinear phenomenon.

Indeed, *Chow's Theorem* states that (4) is controllable if the *involutive closure*[5] of $\Delta = \text{span}\{g_1, \cdots, g_m\}$ has dimension n. This condition means that the Lie brackets of $g_1, \cdots, g_m$ span new dimensions that are not already spanned by these basis vector fields. Drift-free systems satisfying Chow's Theorem are called *nonholonomic*.

[5] the smallest involutive distribution that containts Δ

Example 4: Recall the unicycle model discussed in Example 2, where

$$g_1(x) = \begin{bmatrix} \cos x_3 \\ \sin x_3 \\ 0 \end{bmatrix}, \quad g_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad [g_1, g_2](x) = \begin{bmatrix} -\sin x_3 \\ \cos x_3 \\ 0 \end{bmatrix}.$$

$\Delta = \text{span}\{g_1, g_2\}$ is non-involutive, as $[g_1, g_2]$ generates a new direction. Taken together, g_1, g_2, and $[g_1, g_2]$ span the entire three-dimensional space at each point x; therefore, the system is controllable by Chow's Theorem. This conclusion sheds light on how parallel parking is possible despite lack of sideways actuation. □

To present an interpretation of the Lie bracket $[g_1, g_2]$, we let $\Phi_t^{g_i}(x_0)$ denote the solution of the system $\dot{x} = g_i(x)$ at time t from initial condition x_0. Then it can be shown that

$$\Phi_t^{-g_2}(\Phi_t^{-g_1}(\Phi_t^{g_2}(\Phi_t^{g_1}(x_0)))) = t^2[g_1, g_2](x_0) + \mathcal{O}(t^3),$$

which suggests that motion in the direction of the Lie bracket $[g_1, g_2]$ can be generated by alternating actuation of the two inputs u_1 and u_2 with positive and negative signs, as one does in parallel parking.

EE C222/ME C237 - Spring'18 - Lecture 22 Notes[1]

Murat Arcak

April 23 2018

Finite Time Convergence

Systems with Lipschitz continuous dynamics converge to equilibrium points no faster than exponentially (Homework 10, Problem 1). Finite-time convergence is thus possible only with non-Lipschitz or discontinuous dynamics, as illustrated in the following examples.

Example 1: Consider the system $\dot{x} = -x^{1/3}$, where the right-hand side is not Lipschitz. We rearrange the differential equation as $x^{-1/3}\dot{x} = \frac{3}{2}\frac{d}{dt}x^{2/3} = -1$ for $x \neq 0$, and obtain the solution

$$x(t)^{2/3} = x(0)^{2/3} - \frac{2t}{3}$$

which holds until $x(t)$ reaches 0 at $t = \frac{3}{2}x(0)^{2/3}$.

Example 2: Consider the system $\dot{x} = -\mathrm{sgn}(x)$ where

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

The solution is $x(t) = x(0) - t$ when $x(0) > 0$ and $x(t) = x(0) + t$ when $x(0) < 0$, until $x(t)$ reaches zero at $t = |x(0)|$ in each case.

The following proposition allows us to conclude finite time convergence from a Lyapunov function.

Proposition: Consider the system $\dot{x} = f(t,x)$, $f(t,0) = 0\ \forall t$. If there exists a positive definite, continuously differentiable and radially unbounded function $V : \mathbb{R}^n \mapsto \mathbb{R}$, and constants $c > 0$ and $\alpha \in (0,1)$ such that, for all t and x,

$$\dot{V}(x) := \nabla V(x)^T f(t,x) \leq -cV(x)^\alpha$$

then all trajectories converge to the origin in finite time.

The proof follows by defining $w(t) := V(x(t))$, which satisfies the differential inequality $\dot{w}(t) \leq -cw(t)^\alpha$. Finite time convergence of $w(t)$ and, thus of $x(t)$, can then be argued by rearranging and solving the differential equation $\dot{\bar{w}} = -c\bar{w}^\alpha$, $\bar{w}(0) = w(0)$, as in Example 1 above, and noting that $w(t) \leq \bar{w}(t)$.

As an illustration, in Example 2 above $V = \frac{1}{2}x^2$ yields

$$\dot{V} = -x\,\mathrm{sgn}(x) = -|x| = -\sqrt{2V},$$

which satisfies the proposition above with $c = \sqrt{2}$ and $\alpha = 1/2$.

Example 3: Consider the control system

$$\dot{x} = u + \delta(x), \quad x \in \mathbb{R}, u \in \mathbb{R},$$

where $\delta(x)$ is unknown, but an upper bound $\rho(x)$ is available:

$$|\delta(x)| \leq \rho(x).$$

To stabilize the origin despite the unknown $\delta(x)$ we can apply

$$u = -(\rho(x) + \rho_0)\mathrm{sgn}(x),$$

where $\rho_0 > 0$ is a constant. Then $V = \frac{1}{2}x^2$ gives

$$\begin{aligned}
\dot{V} &= -(\rho(x) + \rho_0)|x| + x\delta(x) \\
&\leq -(\rho(x) + \rho_0)|x| + |x||\delta(x)| \\
&= -\rho_0|x| - (\rho(x) - |\delta(x)|)|x| \\
&\leq -\rho_0|x| = -\rho_0\sqrt{2V}.
\end{aligned}$$

This implies that, in addition to dominating the uncertain term $\delta(x)$, we achieve finite time stability of $x = 0$.

Sliding Mode Control

Example 3 demonstrated the ability of a discontinuous controller to dominate uncertain terms. Sliding mode control extends this idea to higher order systems, as illustrated in the following example.

Example 4: Consider the second order system

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= h(x) + g(x)u
\end{aligned} \tag{1}$$

where $g(x) \geq g_0 > 0$ $\forall x$. If we can drive the trajectories to the surface

$$s := x_2 + a_1 x_1 = 0,$$

where $a_1 > 0$ is a design parameter, then x_1 is governed by $\dot{x}_1 = -a_1 x_1$ on this surface and converges to zero along with $x_2 = -a_1 x_1$. To ensure $s(t) \to 0$ note that

$$\dot{s} = a_1 x_2 + h(x) + g(x)u, \tag{2}$$

and let ρ be a function such that

$$\frac{|a_1 x_2 + h(x)|}{g(x)} \leq \rho(x). \tag{3}$$

Then apply the controller

$$u = -(\rho(x) + \rho_0)\operatorname{sgn}(s), \ \rho_0 > 0, \tag{4}$$

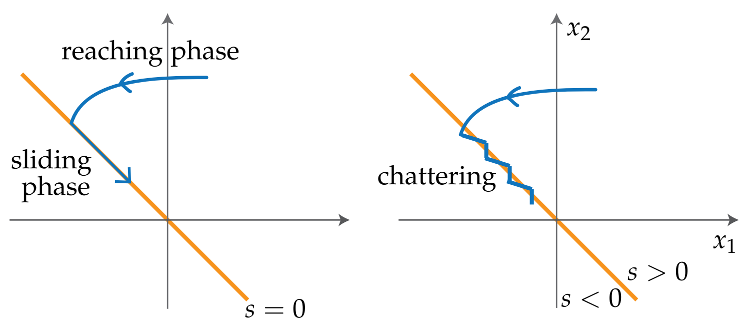
and note that $V = \frac{1}{2}s^2$ satisfies

$$\begin{aligned} \dot{V} &= s[a_1 x_2 + h(x) - g(x)(\rho(x) + \rho_0)\operatorname{sgn}(s)] \\ &\leq |a_1 x_2 + h(x)||s| - g(x)(\rho(x) + \rho_0)|s| \\ &\leq \underbrace{(|a_1 x_2 + h(x)| - g(x)\rho(x))}_{\leq 0 \text{ by } (3)}|s| - g_0\rho_0|s|. \end{aligned}$$

Thus, $\dot{V} \leq -g_0\rho_0|s| = -\sqrt{2}g_0\rho_0 V^{1/2}$, and the proposition on page 1 implies $s(t) \to 0$ in finite time.

An advantage of the controller (4) is that it does not require exact knowledge of h and g; it relies only on the upper bound (3).

The closed-loop system evolves in two phases. In the *reaching* phase the controller forces the trajectories to the surface $s = 0$ in finite time. In the *sliding* phase the trajectories slide on this surface to the origin.



In practice delays in switching lead to "chattering" around the sliding surface, as illustrated in the figure above (right).

To mitigate chattering one idea is to divide the control into a continuous part for the nominal dynamical model and a discontinuous part for the remaining uncertain terms. With this approach the magnitude of the discontinuity is reduced and, thus, chattering is less severe.

Example 4 revisited: Let $\hat{h}$ and $\hat{g} > 0$ denote nominal models for h and g. Define

$$\delta(x) := h(x) - \hat{h}(x).$$

and rewrite (2) as

$$\dot{s} = a_1 x_2 + \hat{h}(x) + \delta(x) + g(x)u.$$

Then we can attempt to cancel the first two, known terms with

$$u = -\frac{a_1 x_2 + \hat{h}(x)}{\hat{g}(x)} + v, \tag{5}$$

where v is left to be designed. Because the cancelation is inexact when $\hat{g} \neq g$, this results in

$$\dot{s} = \underbrace{\left(1 - \frac{g(x)}{\hat{g}(x)}\right)(a_1 x_2 + \hat{h}(x)) + \delta(x)}_{=:\,\Delta(x)} + g(x)v,$$

and the task for v is to dominate the combined uncertain term $\Delta(x)$. This is accomplished with the choice

$$v = -(r(x) + r_0)\mathrm{sgn}(s), \quad r_0 > 0, \tag{6}$$

where r is a function satisfying

$$\frac{|\Delta(x)|}{g(x)} \leq r(x). \tag{7}$$

The finite time convergence of $s(t)$ to zero follows from a Lyapunov analysis similar to that in Example 4 above.

The advantage of the control (5)-(6) is that the continuous part (5) accounts for the nominal term $\hat{h}$, save for the inexact cancelation when $\hat{g} \neq g$. Thus, the magnitude of r in the discontinuous term (6) can be significantly smaller than ρ in (4), leading to reduced chattering.

Example 5: For a specific illustration of the control design (5)-(6), consider the model

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \theta x_1^2 + u, \end{aligned}$$

where θ is an uncertain parameter in the interval $[0.9, 1.1]$. This model is of the form (1) with $h(x) = \theta x_1^2$ and $g(x) = 1$. We let $\hat{h}(x) = x_1^2$, and $\hat{g} = g = 1$, since the latter is perfectly known. Thus

$$\delta(x) := h(x) - \hat{h}(x) = (\theta - 1)x_1^2$$

where $|\theta - 1| \leq 0.1$, and we can take $r(x) = 0.1x_1^2$ to satisfy (7). The controller (5)-(6) is then

$$u = -a_1 x_2 - x_1^2 - (0.1x_1^2 + r_0)\mathrm{sgn}(a_1 x_1 + x_2), \quad a_1 > 0,\ r_0 > 0.$$

Note that the magnitude of the sgn function is diminished relative to the purely discontinuous controller (4) where ρ must satisfy (3), *e.g.*,

$$\rho(x) = a_1|x_2| + 1.1x_1^2.$$

EE C222/ME C237 - Spring'18 - Lecture 23 Notes[1]

Murat Arcak

April 25 2018

Sliding Mode Control Continued

We generalize the sliding mode control examples of the last lecture to the class of systems

$$\begin{aligned} \dot{\eta} &= f_a(\eta, \omega) \\ \dot{\omega} &= f_b(\eta, \omega) + \delta(\eta, \omega) + G(\eta, \omega)u, \end{aligned} \tag{1}$$

where $\omega \in \mathbb{R}^p$, $u \in \mathbb{R}^p$, $\eta \in \mathbb{R}^{n-p}$. The uncertain terms are $\delta(\eta, \omega)$ and the $p \times p$ matrix $G(\eta, \omega)$, assumed to be diagonal with entries

$$g_i(\eta, \omega) \geq g_0 > 0, \quad i = 1, \cdots, p.$$

Let $\phi(\eta)$ be a virtual control law for ω that stabilizes the origin of the η-subsystem, $\dot{\eta} = f_a(\eta, \phi(\eta))$. To drive the trajectories to the sliding surface $\omega = \phi(\eta)$, we note that $s := \omega - \phi(\eta) \in \mathbb{R}^p$ satisfies

$$\dot{s} = f_b(\eta, \omega) - \frac{\partial \phi(\eta)}{\partial \eta} f_a(\eta, \omega) + \delta(\eta, \omega) + G(\eta, \omega)u,$$

and let

$$u = -\hat{G}^{-1}(\eta, \omega) \left[f_b(\eta, \omega) - \frac{\partial \phi(\eta)}{\partial \eta} f_a(\eta, \omega) \right] + v$$

where $\hat{G}(\eta, \omega)$ is a nominal model for $G(\eta, \omega)$, and v is to be designed. Then,

$$\dot{s} = \underbrace{(I - G\hat{G}) \left[f_b(\eta, \omega) - \frac{\partial \phi(\eta)}{\partial \eta} f_a(\eta, \omega) \right] + \delta(\eta, \omega)}_{=: \Delta(\eta, \omega)} + G(\eta, \omega)v,$$

which means that the ith entry of s satisfies

$$\dot{s}_i = \Delta_i(\eta, \omega) + g_i(\eta, \omega)v_i, \quad i = 1, \cdots, p.$$

We let

$$v_i = -(\rho_i(\eta, \omega) + \rho_0)\text{sgn}(s_i), \quad \rho_0 > 0, \tag{2}$$

where $\rho_i(\eta, \omega)$ is a function such that

$$\frac{|\Delta_i(\eta, \omega)|}{g_i(\eta, \omega)} \leq \rho_i(\eta, \omega).$$

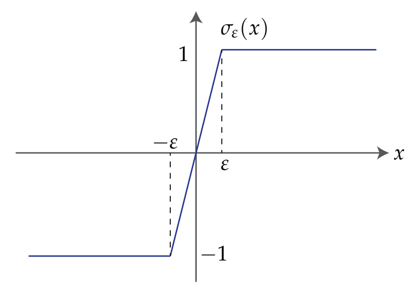
Then the Lyapunov function $V_i = \frac{1}{2}s_i^2$ satisfies $\dot{V}_i \leq -\sqrt{2}\rho_0 g_0 V^{1/2}$, which guarantees finite time convergence of s_i to 0, as discussed in Lecture 22.

Thus the trajectories reach the sliding surface $\omega = \phi(\eta)$ in finite time and, if the subsystem $\dot{\eta} = f_a(\eta, \phi(\eta) + s)$ is ISS with respect to s, then η remains bounded during the reaching phase and converges to zero asymptotically during the sliding phase.

Continuous Approximation of Sliding Mode Control

To avoid the chattering phenomenon discussed in the previous lecture, we can employ the continuous function

$$\sigma_\varepsilon(x) := \begin{cases} x/\varepsilon & \text{when } x \in [-\varepsilon, \varepsilon] \\ \mathrm{sgn}(x) & \text{otherwise,} \end{cases}$$



which approximates $\mathrm{sgn}(\cdot)$ when $\varepsilon > 0$ is a small constant.

If we implement (2) above with $\sigma_\varepsilon(s_i)$ instead of $\mathrm{sgn}(s_i)$, the Lyapunov analysis is unchanged when $|s_i| \geq \varepsilon$, where the two functions are identical. Thus, $|s_i| \geq \varepsilon$ implies $\dot{V}_i \leq -\sqrt{2}\rho_0 g_0 V^{1/2} < 0$, from which we conclude that s_i reaches the interval $[-\varepsilon, \varepsilon]$ in finite time and remains in it thereafter. Likewise, if $\dot{\eta} = f_a(\eta, \phi(\eta) + s)$ is ISS with respect to s, then η converges to a residual set around $\eta = 0$ whose size shrinks as $\varepsilon \to 0$.

Therefore, the continuous approximation eliminates chattering, but guarantees convergence to a small set around the origin rather than to the origin.

<u>Example:</u> For the system

$$\begin{aligned} \dot{x}_1 &= x_1 x_2 \\ \dot{x}_2 &= \theta x_1^2 + u, \quad |\theta| \leq 2, \end{aligned}$$

the virtual control $\phi(x_1) = -x_1^2$ and the variable $s := x_2 - \phi(x_1) = x_2 + x_1^2$ result in

$$\dot{x}_1 = -x_1^3 + x_1 s,$$

which is ISS with respect to s. To drive s to zero we note that

$$\dot{s} = 2x_1^2 x_2 + \theta x_1^2 + u$$

and apply the control

$$u = -2x_1^2 x_2 + v,$$

which guarantees global asymptotic stability of the origin $(x_1, x_2) = (0, 0)$ with the discontinuous feedback $v = -(2x_1^2 + \rho_0)\mathrm{sgn}(s)$.

If we apply the continuous approximation $v = -(2x_1^2 + \rho_0)\sigma_\varepsilon(s)$ we achieve convergence to a set which shrinks to the origin as $\varepsilon \to 0$.

Tracking Control

Consider a model represented in the normal form for input-output linearization:

$$\begin{aligned}
\dot{z} &= f_0(z,\zeta) \\
\dot{\zeta}_1 &= \zeta_2 \\
&\vdots \\
\dot{\zeta}_{r-1} &= \zeta_r \\
\dot{\zeta}_r &= b(z,\zeta) + a(z,\zeta)u \\
y &= \zeta_1,
\end{aligned}$$

where $a(z,\zeta)$ and $b(z,\zeta)$ are imperfectly known, but

$$a(z,\zeta) \geq g_0 > 0$$

with some positive constant g_0. In addition we assume the zero dynamics subsystem $\dot{z} = f_0(z,\zeta)$ is ISS with respect to ζ.

This system is of the general form (1) with $\eta = [z^T, \zeta_1, \cdots, \zeta_{r-1}]^T$ and $\omega = \zeta_r$, and we can design a virtual control

$$\zeta_r = -k_{r-1}\zeta_{r-1} - \cdots - k_1\zeta_1 \tag{3}$$

with coefficients $k_{r_1}, \cdots k_1$ such that

$$\begin{bmatrix} \dot{\zeta}_1 \\ \vdots \\ \vdots \\ \dot{\zeta}_{r-1} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ & & \ddots & 1 \\ -k_1 & \cdots & \cdots & -k_{r-1} \end{bmatrix}}_{=:\, A_0} \begin{bmatrix} \zeta_1 \\ \vdots \\ \vdots \\ \zeta_{r-1} \end{bmatrix} \tag{4}$$

is asymptotically stable.

The dynamics restricted to the sliding surface (3) consist of the subsystem (4) driving the ISS zero dynamics; therefore the trajectories converge to the origin. Finite time convergence to the surface is achieved with the standard design approach discussed on page 1.

When the goal is to ensure that the output ζ_1 tracks the desired trajectory $y_d(t)$, we define the tracking error variables

$$e_1 := \zeta_1 - y_d(t),\ e_2 := \zeta_2 - \dot{y}_d(t),\ \cdots, e_r := \zeta_r - y_d^{(r-1)}(t)$$

and rewrite the system equations as

$$\begin{aligned} \dot{z} &= f_0(z, e_1 + y_d(t), \ldots, e_r + y_d^{(r-1)}(t)) \\ \dot{e}_1 &= e_2 \\ &\vdots \\ \dot{e}_{r-1} &= e_r \\ \dot{e}_r &= b(z,\zeta) - y_d^{(r)}(t) + a(z,\zeta)u. \end{aligned}$$

As the sliding surface we select

$$s := e_r + k_{r-1}e_{r-1} + \cdots + k_1 e_1 = 0, \tag{5}$$

where $k_{r-1}, \cdots k_1$ are such that the matrix A_0 defined in (4) has all eigenvalues with negative real parts. Thus, $e(t) \to 0$ on the sliding surface and $z(t)$ remains bounded by the ISS assumption, and by the boundedness of $y_d(t)$ and its derivatives.

For the reaching phase we note that

$$\begin{aligned} \dot{s} &= \dot{e}_r + k_{r-1}\dot{e}_{r-1} + \cdots + k_1\dot{e}_1 \\ &= b(z,\zeta) - y_d^{(r)}(t) + k_{r-1}e_r + \cdots + k_1 e_2 + a(z,\zeta)u, \end{aligned}$$

and select

$$u = -\frac{1}{\hat{a}(z,\zeta)}\left[\hat{b}(z,\zeta) - y_d^{(r)}(t) + k_{r-1}e_r + \cdots + k_1 e_2\right] + v. \tag{6}$$

This yields

$$\dot{s} = \underbrace{\left(1 - \frac{a(z,\zeta)}{\hat{a}(z,\zeta)}\right)[\,\cdots\,] + (b(z,\zeta) - \hat{b}(z,\zeta))}_{=:\,\Delta(z,\zeta,t)} + a(z,\zeta)v$$

where $[\,\cdots\,]$ is the square bracketed term in (6), and $\Delta(z,\zeta,t)$ depends on t due to the derivatives of $y_d(t)$ occuring in this expression.

We then choose $\rho(z,\zeta,t)$ such that

$$\frac{|\Delta(z,\zeta,t)|}{a(z,\zeta)} \le \rho(z,\zeta,t)$$

and complete the design (6) with

$$v = -(\rho(z,\zeta,t) + \rho_0)\mathrm{sgn}(s), \;\; \rho_0 > 0. \tag{7}$$

Note that, if we set $y_d(t) \equiv 0$, the tracking controller (6)-(7) reduces to a stabilizing controller for the origin $(z,\zeta) = 0$.

Example: Consider the system

$$\begin{aligned} \dot{x}_1 &= x_2 + \sin x_1 \\ \dot{x}_2 &= \theta_1 x_1^2 + (1+\theta_2)u \quad |\theta_1| \le 2, \; |\theta_2| \le 0.5, \\ y &= x_1. \end{aligned} \tag{8}$$

To design a tracking controller we first bring the system to the normal form with the new variables $\zeta_1 = x_1$ and $\zeta_2 = x_2 + \sin x_1$:

$$\begin{aligned}\dot{\zeta}_1 &= \zeta_2 \\ \dot{\zeta}_2 &= (x_2 + \sin x_1)\cos x_1 + \theta_1 x_1^2 + (1+\theta_2)u.\end{aligned} \tag{9}$$

Then we define the error variables $e_1 = \zeta_1 - y_d(t)$ and $e_2 = \zeta_2 - \dot{y}_d(t)$, which are governed by

$$\begin{aligned}\dot{e}_1 &= e_2 \\ \dot{e}_2 &= (x_2 + \sin x_1)\cos x_1 - \ddot{y}_d(t) + \theta_1 x_1^2 + (1+\theta_2)u,\end{aligned}$$

and select the sliding surface

$$s := e_2 + k_1 e_1 = 0, \quad k_1 > 0.$$

Thus,

$$\dot{s} = (x_2 + \sin x_1)\cos x_1 - \ddot{y}_d(t) + k_1 e_2 + \theta_1 x_1^2 + (1+\theta_2)u$$

and the feedback

$$\begin{aligned}u &= -(x_2 + \sin x_1)\cos x_1 + \ddot{y}_d(t) - k_1 e_2 + v \\ v &= -(\rho(x,t) + \rho_0) \quad \rho_0 > 0\end{aligned}$$

results in

$$\dot{s} = \underbrace{\theta_2(-(x_2 + \sin x_1)\cos x_1 + \ddot{y}_d(t) - k_1 e_2) + \theta_1 x_1^2}_{=:\,\Delta(x_1, x_2, t)} + (1+\theta_2)v.$$

Using the bounds $|\theta_1| \leq 2$, $|\theta_2| \leq 0.5$ we get

$$\begin{aligned}\frac{|\Delta(x_1, x_2, t)|}{1+\theta_2} &\leq \frac{0.5|(x_2 + \sin x_1)\cos x_1 - \ddot{y}_d(t) + k_1 e_2| + 2x_1^2}{0.5} \\ &= |(x_2 + \sin x_1)\cos x_1 - \ddot{y}_d(t) + k_1 e_2| + 4x_1^2\end{aligned}$$

and, substituting $e_2 = \zeta_2 - \dot{y}_d(t) = x_2 + \sin x_1 - \dot{y}_d(t)$, we select

$$\rho(x,t) = |(x_2 + \sin x_1)\cos x_1 - \ddot{y}_d(t) + k_1(x_2 + \sin x_1 - \dot{y}_d(t)| + 4x_1^2.$$

It is important to note that sliding mode control can address only limited forms of uncertainty. In the example (8) the uncertain terms appear in the same equation as the control input; that is, they are "matched" to the input. The first equation in (8) contains no uncertainty, which allowed us to bring the system to the normal form (9).