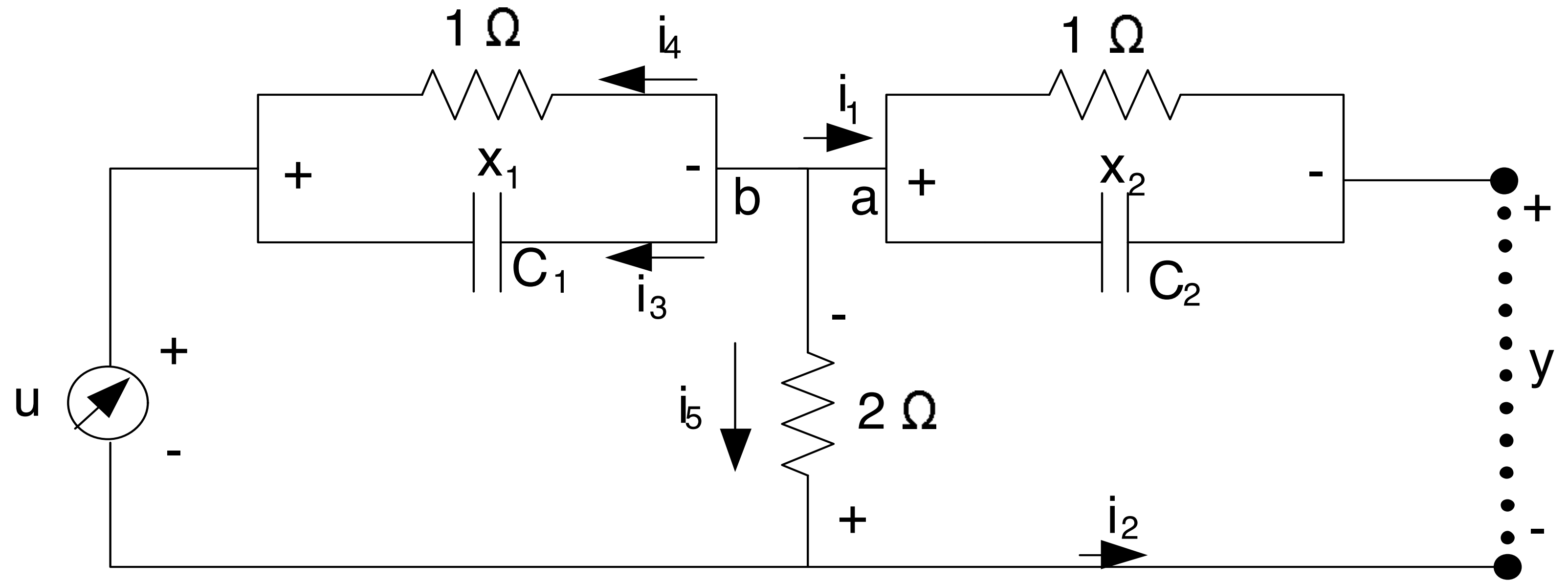
Controllability

- Definitions
- Theorems
- Proofs
- Examples

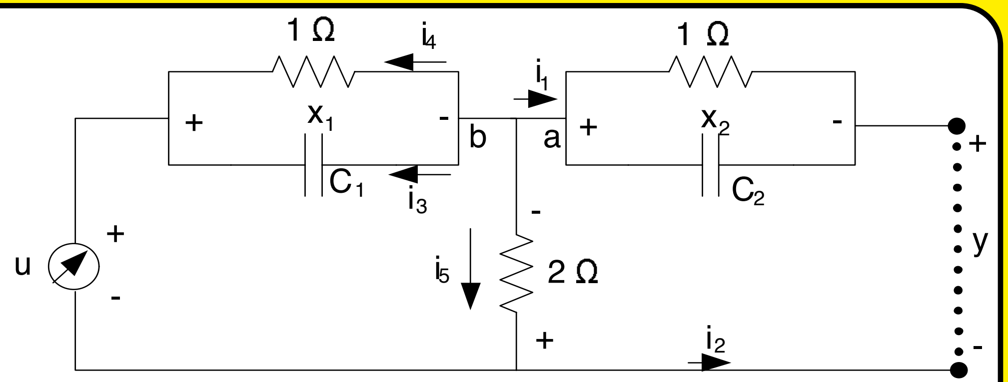
What are Controllability / Observability

- Example:



- Input: *u (current)*
- State variables: x_1, x_2 (voltages)
- Output: y (voltage)

Are the system states observable (through y)?
Are the system states controllable (through u)?



- $i_1 = i_2 = 0$ (Kirchhoff current law)
- Current law for junc. a:

$$\dot{x}_2 C_2 = -\frac{x_2}{1} = -x_2$$

$$\dot{x}_2 = -\frac{x_2}{C_2}$$

$v = iR \qquad i = C\, dv/dt$

x_2 is not affected by u (uncontrolled)

- Current law for junc. b:

$$i_5 = u = x_1 + C_1 \dot{x}_1$$

$$\dot{x}_1 = -\frac{x_1}{C_1} + \frac{u}{C_1}$$

- Voltage law for left (open circuit): $y = x_2 + 2u$

y is not affected by x_1 (since x_2 is not), so x_1 is unobserved.

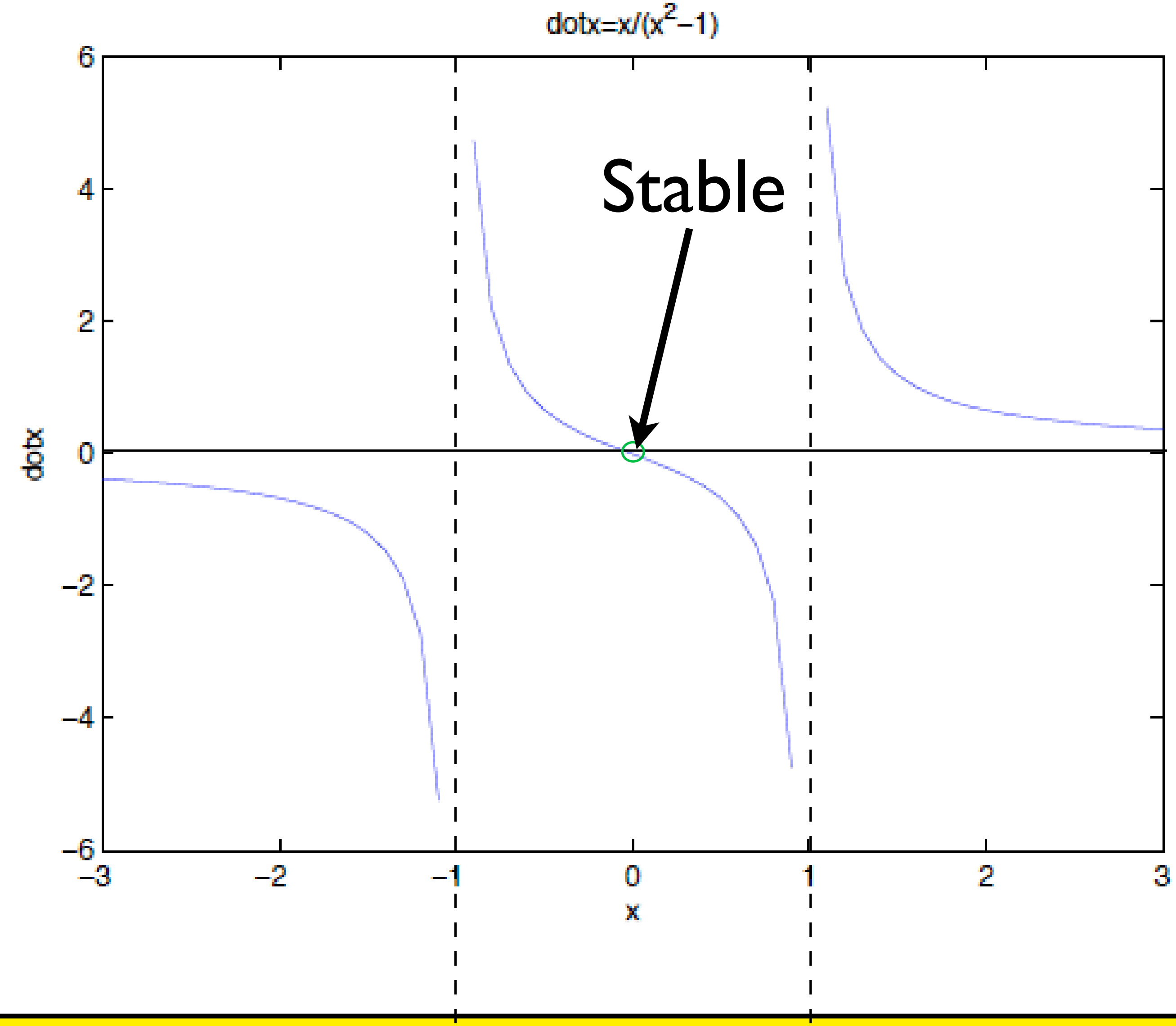
- This system is uncontrolled and unobserved (at least in part).

Definitions

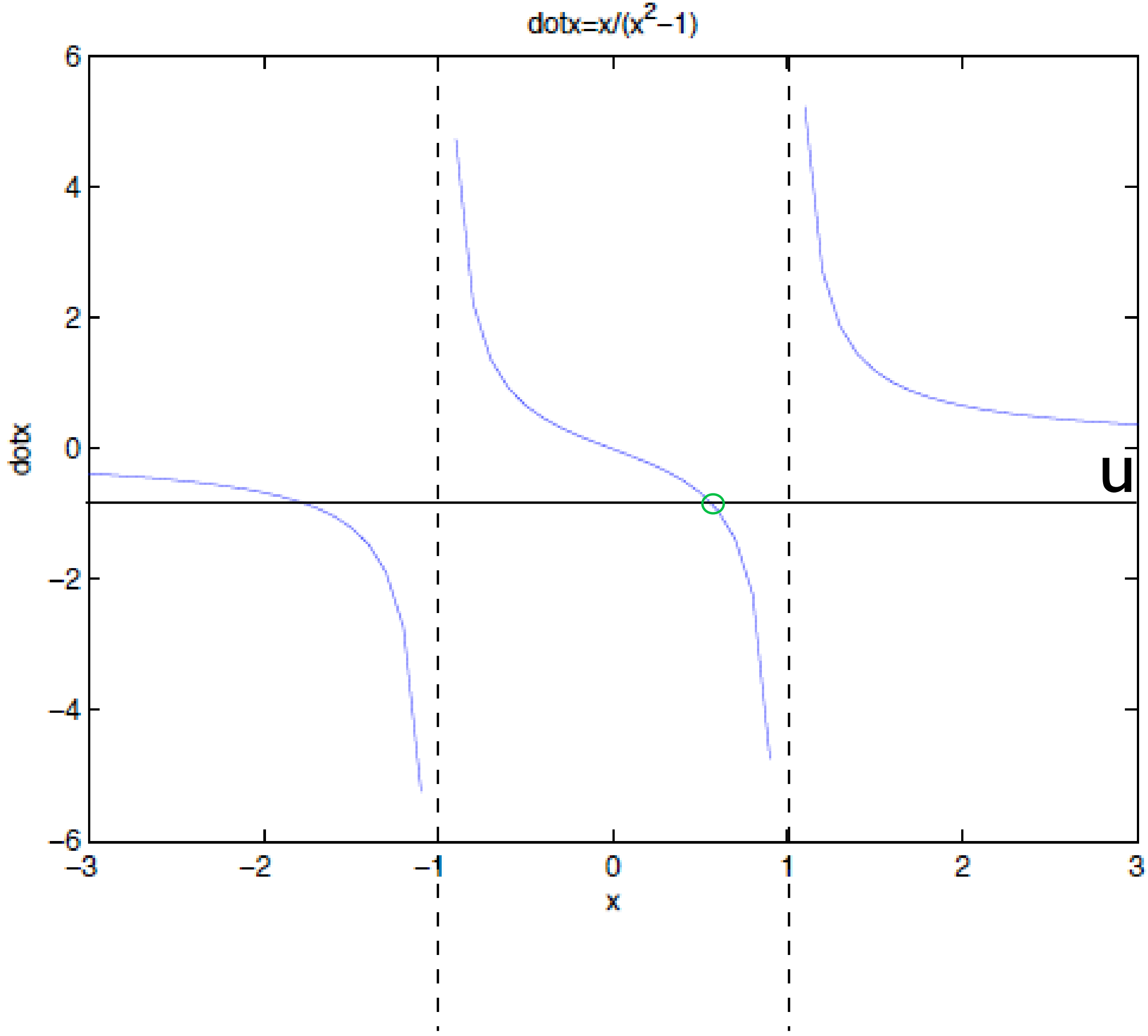
- System is ***controllable*** if for every, $\mathbf{x}_0$, $\mathbf{x}_f$ and t_f there is a control signal $\mathbf{u}$ that brings the system from $\mathbf{x}(0) = \mathbf{x}_0$ to $\mathbf{x}(t_f) = \mathbf{x}_f$ at time t_f.
- Sometimes definition requires $\mathbf{x}_f = \mathbf{0}$ (makes no difference if system is linear)
- The ***reachable region*** at time t_f are the possible values of $\mathbf{x}(t_f)$ if $\mathbf{x}(0) = \mathbf{0}$ (and all possible $\mathbf{u}$s). Similarly, the ***controllable region*** are the possible values of $\mathbf{x}(0)$ so that there exists a control signal that takes the system to $\mathbf{x}(t_f) = \mathbf{0}$.
- So, if system is controllable, the controllable region is the whole state space. Otherwise the controllable region may still be non-empty.
- In LTI systems:
    - The controllable region is always a linear subspace of the state space.
    - Reachable region = controllable region

Nonlinear System Example

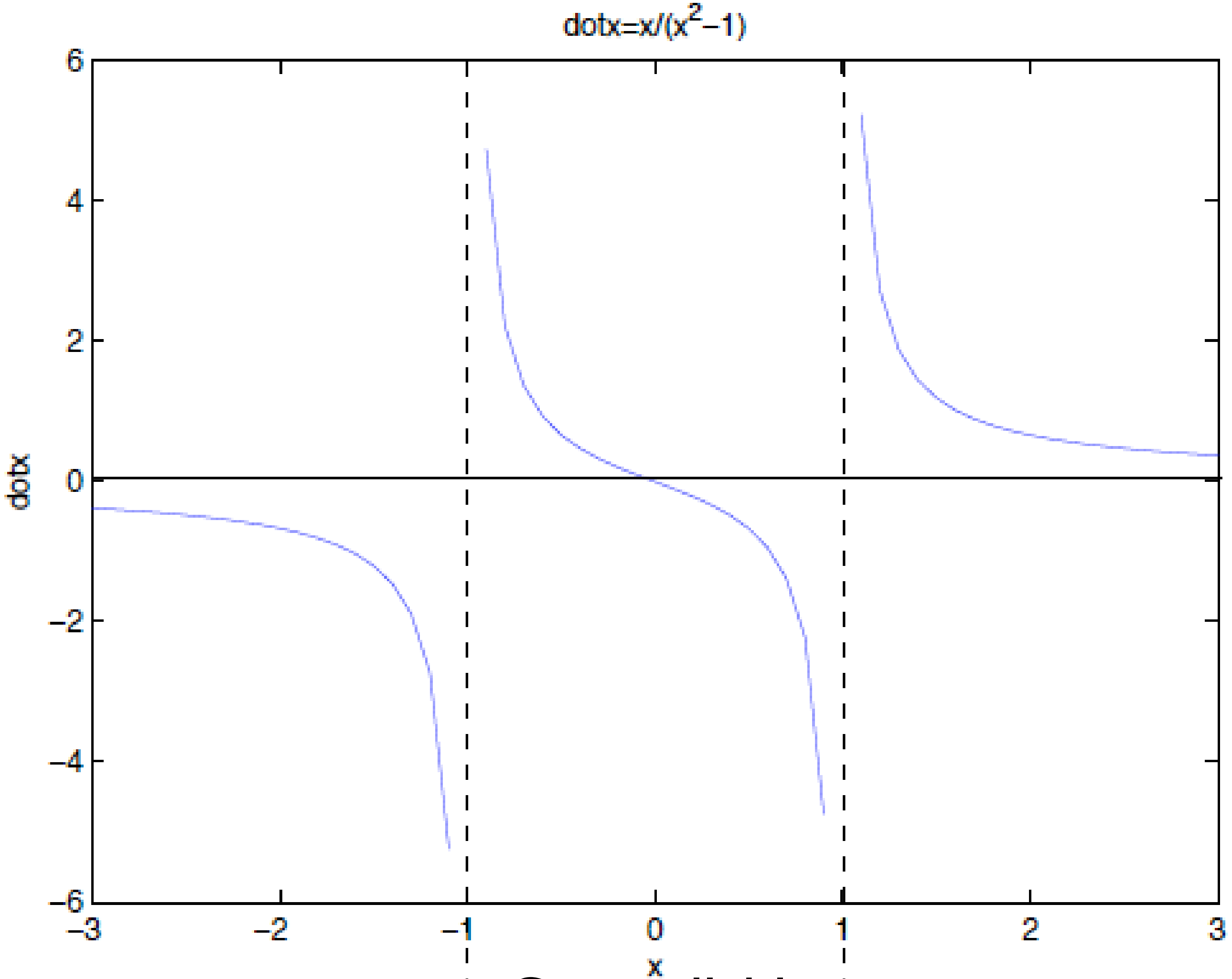
- $\dot{x} = \frac{x}{(x^2-1)} + u$
- For u = 0



- For $u \neq 0$



Controllable?
Reachable?



Controllable
Reachable

Reachability

consider state transfer from $x(0) = 0$ to $x(t)$

we say $x(t)$ is *reachable* (in t seconds or epochs)

we define $\mathcal{R}_t \subseteq \mathbf{R}^n$ as the set of points reachable in t seconds or epochs

for CT system $\dot{x} = Ax + Bu$,

$$\mathcal{R}_t = \left\{ \int_0^t e^{(t-\tau)A} Bu(\tau)\, d\tau \;\middle|\; u : [0,t] \to \mathbf{R}^m \right\}$$

and for DT system $x(t+1) = Ax(t) + Bu(t)$,

$$\mathcal{R}_t = \left\{ \sum_{\tau=0}^{t-1} A^{t-1-\tau} Bu(\tau) \;\middle|\; u(t) \in \mathbf{R}^m \right\}$$

- $\mathcal{R}_t$ is a subspace of $\mathbf{R}^n$
- $\mathcal{R}_t \subseteq \mathcal{R}_s$ if $t \leq s$

  ($i.e.$, can reach more points given more time)

we define the *reachable set* $\mathcal{R}$ as the set of points reachable for some t:

$$\mathcal{R} = \bigcup_{t \geq 0} \mathcal{R}_t$$

Theorem (controllability):

Given a system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$, $\mathbf{x} \in \mathrm{R}^n$ the following are equivalent:

1. The (**A**,**B**) pair is controllable.

2. The *controllability grammian* matrix,

$$\mathbf{W}_c(t) = \int_0^t e^{A\tau} B B' e^{A'\tau} d\tau = \int_0^t e^{A(t-\tau)} B B' e^{A'(t-\tau)} d\tau$$

   is non-singular (invertible) for all *t*.

3. The $n \times np$ *controllability matrix*, C=[$\mathbf{B}$ $\mathbf{AB}$ $\mathbf{A}^2\mathbf{B}$ **...** $\mathbf{A}^{n-1}\mathbf{B}$] has full row rank (*n*).

4. The matrix [$\mathbf{A}-\mathbf{I}\lambda_i$ $\mathbf{B}$] has full row rank for every eigenvalue λ_i of **A**.

Reachability for discrete-time LDS

DT system $x(t+1) = Ax(t) + Bu(t)$, $x(t) \in \mathbf{R}^n$

$$x(t) = \mathcal{C}_t \begin{bmatrix} u(t-1) \\ \vdots \\ u(0) \end{bmatrix}$$

where $\mathcal{C}_t = \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix}$

so reachable set at t is $\mathcal{R}_t = \mathrm{range}(\mathcal{C}_t)$

by C-H theorem, we can express each A^k for $k \geq n$ as linear combination of $A^0, \ldots, A^{n-1}$

hence for $t \geq n$, $\mathrm{range}(\mathcal{C}_t) = \mathrm{range}(\mathcal{C}_n)$

thus we have

$$\mathcal{R}_t = \begin{cases} \operatorname{range}(\mathcal{C}_t) & t < n \\ \operatorname{range}(\mathcal{C}) & t \geq n \end{cases}$$

where $\mathcal{C} = \mathcal{C}_n$ is called the *controllability matrix*

- any state that can be reached can be reached by $t = n$
- the reachable set is $\mathcal{R} = \operatorname{range}(\mathcal{C})$

Controllable system

system is called *reachable* or *controllable* if all states are reachable ($i.e.$, $\mathcal{R} = \mathbf{R}^n$)

system is reachable if and only if $\mathbf{Rank}(\mathcal{C}) = n$

example: $x(t+1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t)$

controllability matrix is $\mathcal{C} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

hence system is not controllable; reachable set is

$$\mathcal{R} = \mathrm{range}(\mathcal{C}) = \{ \, x \mid x_1 = x_2 \, \}$$

General state transfer

with $t_f > t_i$,

$$x(t_f) = A^{t_f - t_i} x(t_i) + \mathcal{C}_{t_f - t_i} \begin{bmatrix} u(t_f - 1) \\ \vdots \\ u(t_i) \end{bmatrix}$$

hence can transfer $x(t_i)$ to $x(t_f) = x_{\mathrm{des}}$

$$\Leftrightarrow \qquad x_{\mathrm{des}} - A^{t_f - t_i} x(t_i) \in \mathcal{R}_{t_f - t_i}$$

- general state transfer reduces to reachability problem
- if system is controllable any state transfer can be achieved in $\leq n$ steps
- important special case: driving state to zero (sometimes called regulating or controlling state)

Least-norm input for reachability

assume system is reachable, $\mathbf{Rank}(\mathcal{C}_t) = n$

to steer $x(0) = 0$ to $x(t) = x_{\rm des}$, inputs $u(0), \ldots, u(t-1)$ must satisfy

$$x_{\rm des} = \mathcal{C}_t \begin{bmatrix} u(t-1) \\ \vdots \\ u(0) \end{bmatrix}$$

among all u that steer $x(0) = 0$ to $x(t) = x_{\rm des}$, the one that minimizes

$$\sum_{\tau=0}^{t-1} \|u(\tau)\|^2$$

is given by

$$\begin{bmatrix} u_{\rm ln}(t-1) \\ \vdots \\ u_{\rm ln}(0) \end{bmatrix} = \mathcal{C}_t^T(\mathcal{C}_t\mathcal{C}_t^T)^{-1}x_{\rm des}$$

$u_{\rm ln}$ is called *least-norm* or *minimum energy* input that effects state transfer

can express as

$$u_{\rm ln}(\tau) = B^T(A^T)^{(t-1-\tau)}\left(\sum_{s=0}^{t-1} A^sBB^T(A^T)^s\right)^{-1} x_{\rm des},$$

for $\tau = 0, \ldots, t-1$

$\mathcal{E}_{\min}$, the minimum value of $\sum_{\tau=0}^{t-1} \|u(\tau)\|^2$ required to reach $x(t) = x_{\mathrm{des}}$, is sometimes called *minimum energy* required to reach $x(t) = x_{\mathrm{des}}$

$$
\begin{aligned}
\mathcal{E}_{\min} &= \sum_{\tau=0}^{t-1} \|u_{\mathrm{ln}}(\tau)\|^2 \\
&= \left(\mathcal{C}_t^T (\mathcal{C}_t \mathcal{C}_t^T)^{-1} x_{\mathrm{des}}\right)^T \mathcal{C}_t^T (\mathcal{C}_t \mathcal{C}_t^T)^{-1} x_{\mathrm{des}} \\
&= x_{\mathrm{des}}^T (\mathcal{C}_t \mathcal{C}_t^T)^{-1} x_{\mathrm{des}} \\
&= x_{\mathrm{des}}^T \left(\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau \right)^{-1} x_{\mathrm{des}}
\end{aligned}
$$

- $\mathcal{E}_{\min}(x_{\rm des}, t)$ gives measure of how hard it is to reach $x(t) = x_{\rm des}$ from $x(0) = 0$ (*i.e.*, how large a u is required)

- $\mathcal{E}_{\min}(x_{\rm des}, t)$ gives practical measure of controllability/reachability (as function of $x_{\rm des}$, t)

- ellipsoid $\{ z \mid \mathcal{E}_{\min}(z, t) \le 1 \}$ shows points in state space reachable at t with one unit of energy

  (shows directions that can be reached with small inputs, and directions that can be reached only with large inputs)

$\mathcal{E}_{\min}$ as function of t:

if $t \geq s$ then

$$\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau \geq \sum_{\tau=0}^{s-1} A^\tau B B^T (A^T)^\tau$$

hence

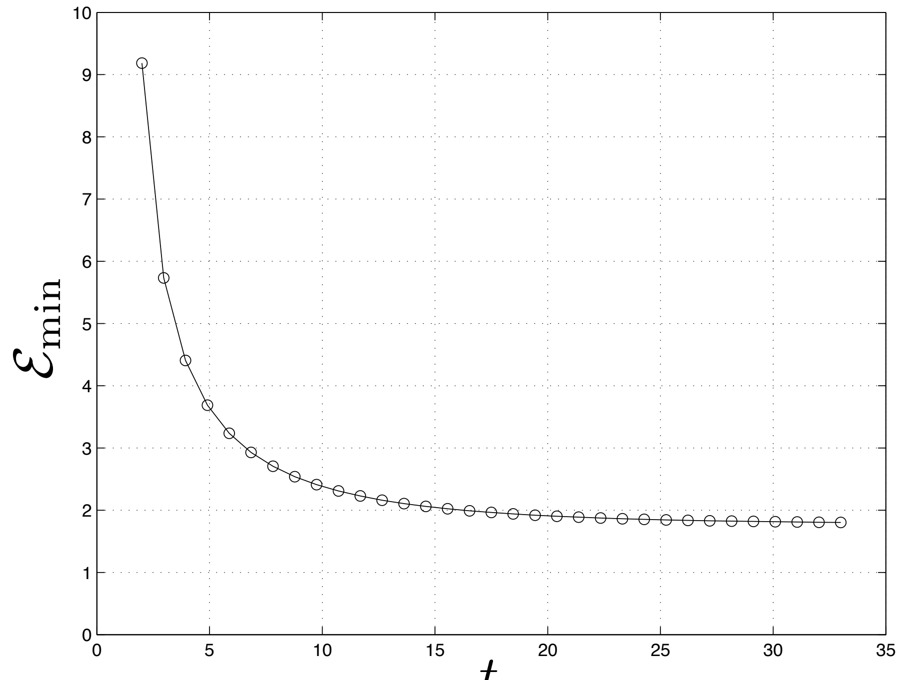
$$\left(\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau \right)^{-1} \leq \left(\sum_{\tau=0}^{s-1} A^\tau B B^T (A^T)^\tau \right)^{-1}$$

so $\mathcal{E}_{\min}(x_{\rm des}, t) \leq \mathcal{E}_{\min}(x_{\rm des}, s)$

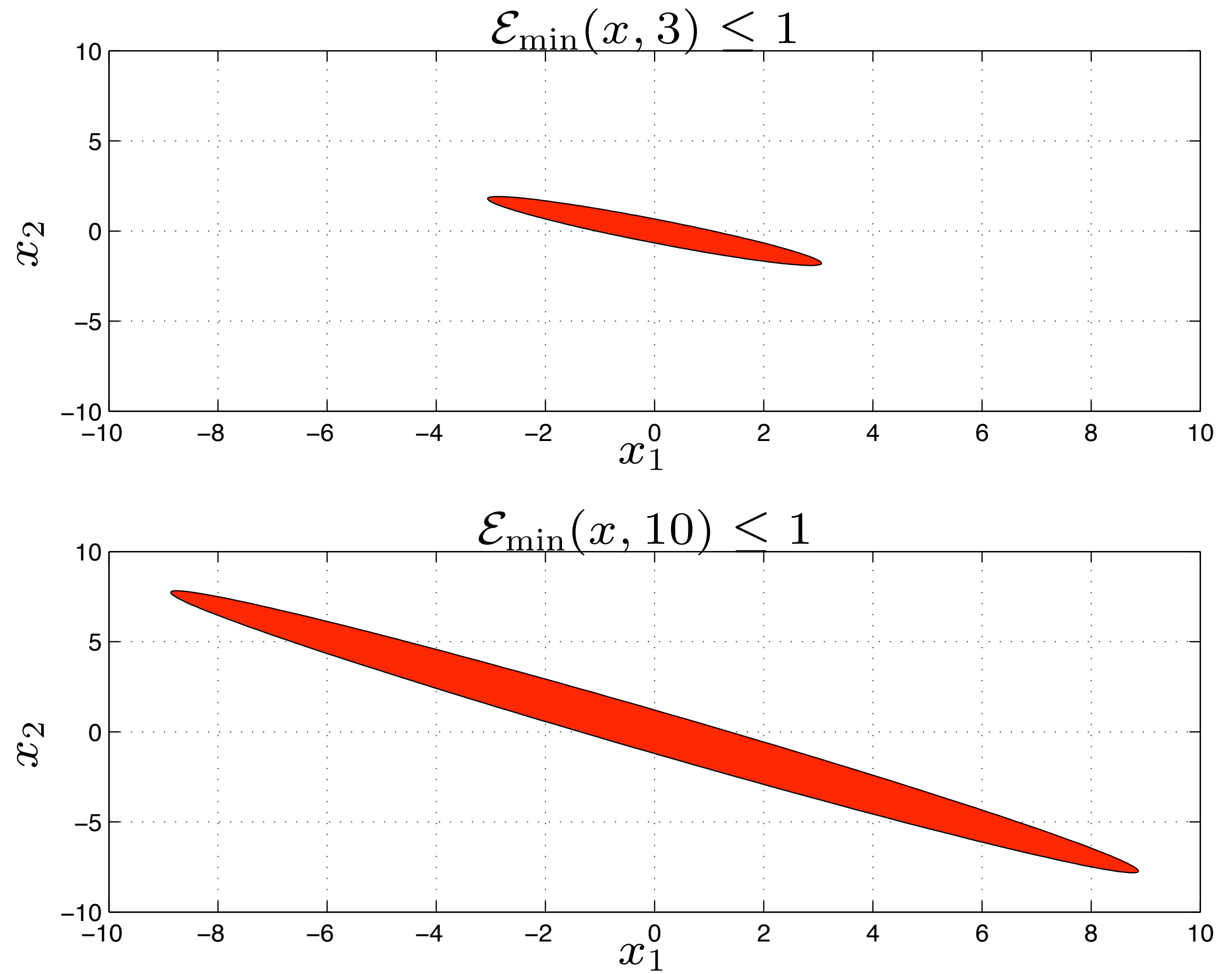
i.e.: takes less energy to get somewhere more leisurely

example: $x(t+1) = \begin{bmatrix} 1.75 & 0.8 \\ -0.95 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$

$\mathcal{E}_{\min}(z,t)$ for $z = [1\ 1]^T$:



ellipsoids $\mathcal{E}_{\min} \leq 1$ for $t = 3$ and $t = 10$:



Minimum energy over infinite horizon

the matrix

$$P = \lim_{t\to\infty} \left(\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau \right)^{-1}$$

always exists, and gives the minimum energy required to reach a point $x_{\rm des}$ (with no limit on t):

$$\min \left\{ \sum_{\tau=0}^{t-1} \|u(\tau)\|^2 \;\middle|\; x(0) = 0,\ x(t) = x_{\rm des} \right\} = x_{\rm des}^T P x_{\rm des}$$

if A is stable, $P > 0$ ($i.e.$, can't get anywhere for free)

if A is not stable, then P can have nonzero nullspace

- $Pz = 0$, $z \neq 0$ means can get to z using u's with energy as small as you like

  (u just gives a little kick to the state; the instability carries it out to z efficiently)

- basis of highly maneuverable, unstable aircraft

Continuous-time reachability

consider now $\dot{x} = Ax + Bu$ with $x(t) \in \mathbf{R}^n$

reachable set at time t is

$$\mathcal{R}_t = \left\{ \int_0^t e^{(t-\tau)A} Bu(\tau)\, d\tau \;\middle|\; u : [0,t] \to \mathbf{R}^m \right\}$$

fact: for $t > 0$, $\mathcal{R}_t = \mathcal{R} = \mathrm{range}(\mathcal{C})$, where

$$\mathcal{C} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$$

is the controllability matrix of (A, B)

- same $\mathcal{R}$ as discrete-time system
- for continuous-time system, any reachable point can be reached as fast as you like (with large enough u)

first let's show for *any* u (and $x(0)=0$) we have $x(t) \in \mathrm{range}(\mathcal{C})$

write e^{tA} as power series:

$$e^{tA} = I + \frac{t}{1!}A + \frac{t^2}{2!}A^2 + \cdots$$

by C-H, express $A^n, A^{n+1}, \ldots$ in terms of $A^0, \ldots, A^{n-1}$ and collect powers of A:

$$e^{tA} = \alpha_0(t)I + \alpha_1(t)A + \cdots + \alpha_{n-1}(t)A^{n-1}$$

therefore

$$\begin{aligned} x(t) &= \int_0^t e^{\tau A} B u(t-\tau)\, d\tau \\ &= \int_0^t \left(\sum_{i=0}^{n-1} \alpha_i(\tau) A^i \right) B u(t-\tau)\, d\tau \end{aligned}$$

$$= \sum_{i=0}^{n-1} A^i B \int_0^t \alpha_i(\tau) u(t-\tau)\, d\tau$$
$$= \mathcal{C} z$$

where $z_i = \displaystyle\int_0^t \alpha_i(\tau) u(t-\tau)\, d\tau$

hence, $x(t)$ is always in $\mathrm{range}(\mathcal{C})$

need to show converse: every point in $\mathrm{range}(\mathcal{C})$ can be reached

Impulsive inputs

suppose $x(0_-) = 0$ and we apply input $u(t) = \delta^{(k)}(t) f$, where $\delta^{(k)}$ denotes kth derivative of δ and $f \in \mathbf{R}^m$

then $U(s) = s^k f$, so

$$\begin{aligned} X(s) &= (sI - A)^{-1} B s^k f \\ &= \left(s^{-1}I + s^{-2}A + \cdots\right) B s^k f \\ &= (\underbrace{s^{k-1} + \cdots + sA^{k-2} + A^{k-1}}_{\text{impulsive terms}} + s^{-1}A^k + \cdots) Bf \end{aligned}$$

hence

$$x(t) = \text{ impulsive terms } + A^k Bf + A^{k+1} Bf \frac{t}{1!} + A^{k+2} Bf \frac{t^2}{2!} + \cdots$$

in particular, $x(0_+) = A^k Bf$

thus, input $u = \delta^{(k)} f$ transfers state from $x(0_-) = 0$ to $x(0_+) = A^k B f$

now consider input of form

$$u(t) = \delta(t) f_0 + \cdots + \delta^{(n-1)}(t) f_{n-1}$$

where $f_i \in \mathbf{R}^m$

by linearity we have

$$x(0_+) = B f_0 + \cdots + A^{n-1} B f_{n-1} = \mathcal{C} \begin{bmatrix} f_0 \\ \vdots \\ f_{n-1} \end{bmatrix}$$

hence we can reach any point in $\mathrm{range}(\mathcal{C})$

(at least, using impulse inputs)

can also be shown that any point in $\mathrm{range}(\mathcal{C})$ can be reached for any $t > 0$ using *nonimpulsive* inputs

fact: if $x(0) \in \mathcal{R}$, then $x(t) \in \mathcal{R}$ for all t (no matter what u is)

to show this, need to show $e^{tA}x(0) \in \mathcal{R}$ if $x(0) \in \mathcal{R}$ . . .

Continuous time system is controllable if the *controllability grammian*

$$\mathbf{W}_c(t) = \int_0^t e^{A\tau} B B' e^{A'\tau} d\tau = \int_0^t e^{A(t-\tau)} B B' e^{A'(t-\tau)} d\tau$$

is non-singular (invertible) for all t.

i.e. for every $\mathbf{x}_0$, $\mathbf{x}_\mathrm{f}$ and t_f we construct the appropriate **u**

- Choose $\mathbf{u}(t) = -\mathbf{B}' e^{\mathbf{A}'(t_f - t)} \mathbf{W}_c^{-1}(t_f) \left[e^{\mathbf{A} t_f} \mathbf{x}_0 - \mathbf{x}_f \right]$
- Plug into dynamic system solution:

$$\mathbf{x}(t_f) \quad = \quad e^{\mathbf{A} t_f} \mathbf{x}_0 + \int_0^{t_f} e^{\mathbf{A}(t_f - \tau)} \mathbf{B} \mathbf{u}(\tau) d\tau$$

and get

$$\mathbf{x}(t_f) \quad = \quad e^{\mathbf{A} t_f} \mathbf{x}_0 - \int_0^{t_f} e^{\mathbf{A}(t_f - \tau)} \mathbf{B} \mathbf{B}' e^{\mathbf{A}'(t_f - \tau)} d\tau \, \mathbf{W}_c^{-1}(t_f) \left[e^{\mathbf{A} t_f} \mathbf{x}_0 - \mathbf{x}_f \right]$$

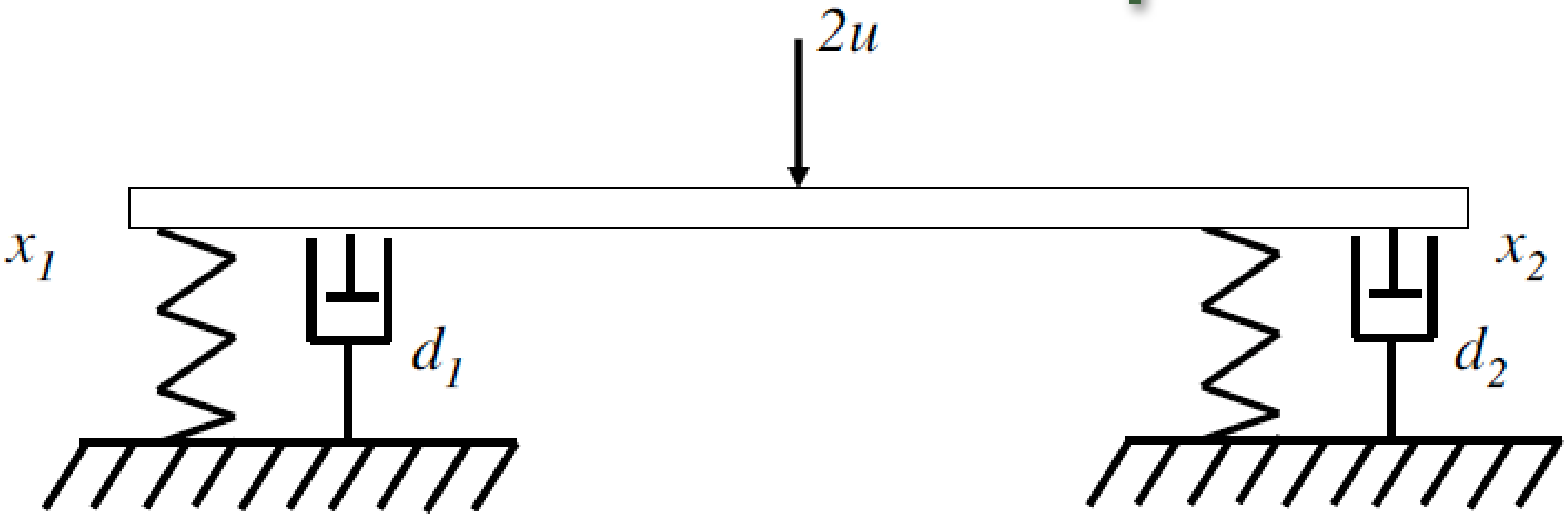
$\mathrm{W}_\mathrm{c}(t_f)$ ←——→ u(t)

$$= \quad e^{\mathbf{A} t_f} \mathbf{x}_0 - \mathbf{W}_c \mathbf{W}_c^{-1}(t_f) \left[e^{\mathbf{A} t_f} \mathbf{x}_0 - \mathbf{x}_f \right]$$

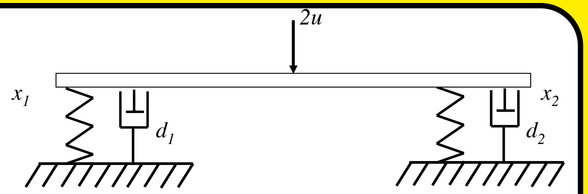
$$= \quad \mathbf{x}_f$$

Will see: above **u** is minimal energy solution

Control Example



- One control signal, divided evenly to two shock absorbers.
- Spring constant, $k = 1$
- Damper constants are d_1, d_2
- Shock absorbers' heights, x_1 and x_2 are the states
- Can we bring the states to zero from any starting heights at finite time using the same force on both absorbers? (i.e. is system controllable?)



- Force equations: $u = x_i + d_i \dot{x}_i$
(no mass...)

In standard form:

$$\dot{\mathbf{x}} = \begin{bmatrix} -\frac{1}{d_1} & 0 \\ 0 & -\frac{1}{d_2} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \frac{1}{d_1} \\ \frac{1}{d_2} \end{bmatrix} u$$

- Easy to see that system must be stable (negative eigenvalues) but will take forever to reach (0,0) with no control..

- Suppose $d1 = 2$ and $d_2 = 1$ we get system:

$$\mathbf{A} = \begin{bmatrix} -1/2 & 0 \\ 0 & -1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}.$$

- Controllability matrix is

$$\mathcal{C} = [\mathbf{b}\,\mathbf{A}\mathbf{b}] = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ 1 & -1 \end{bmatrix}$$

Obviously of rank 2 ➟ controllable

- Lets find $\mathbf{W}_c(2)$, i.e. at t = 2

$$\begin{aligned}
\mathbf{W}_c(2) &= \int_0^2 \left(\begin{bmatrix} e^{-\frac{1}{2}\tau} & 0 \\ 0 & e^{-\tau} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} e^{-\frac{1}{2}\tau} & 0 \\ 0 & e^{-\tau} \end{bmatrix} \right) d\tau \\
&= \int_0^2 \left(\begin{bmatrix} \frac{1}{4}e^{-\tau} & \frac{1}{2}e^{-\frac{3}{2}\tau} \\ \frac{1}{2}e^{-\frac{3}{2}\tau} & e^{-3\tau} \end{bmatrix} \right) d\tau \\
&= \begin{bmatrix} 0.2162 & 0.3167 \\ 0.3167 & 0.4908 \end{bmatrix}
\end{aligned}$$

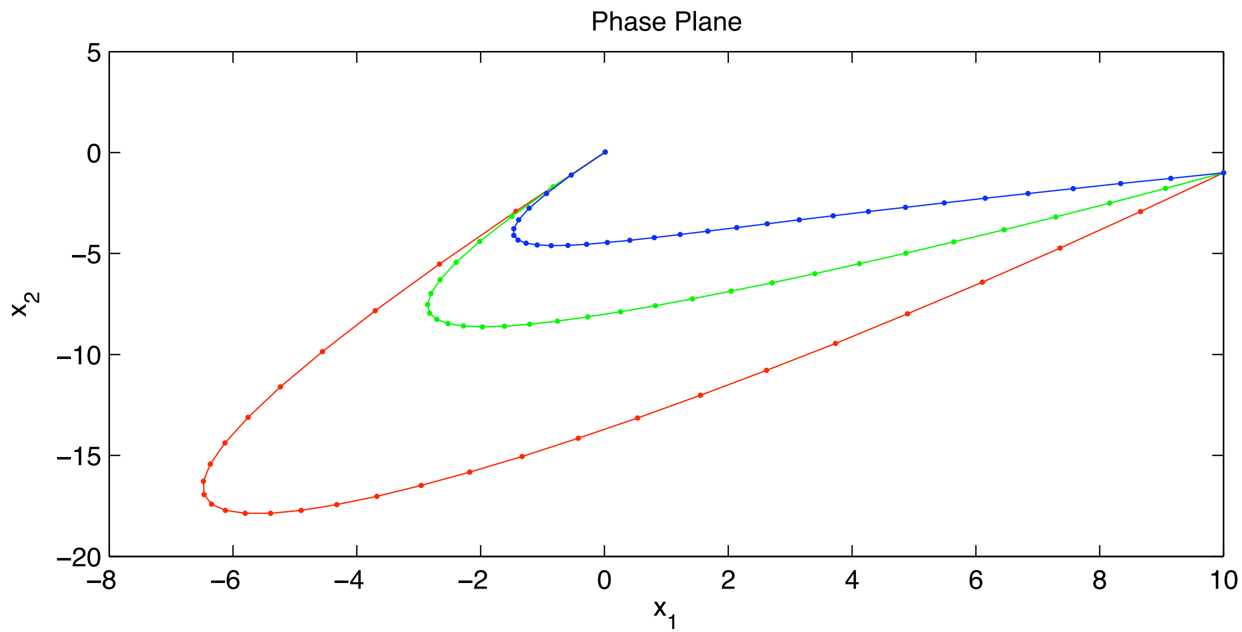
- Now, its inverse:

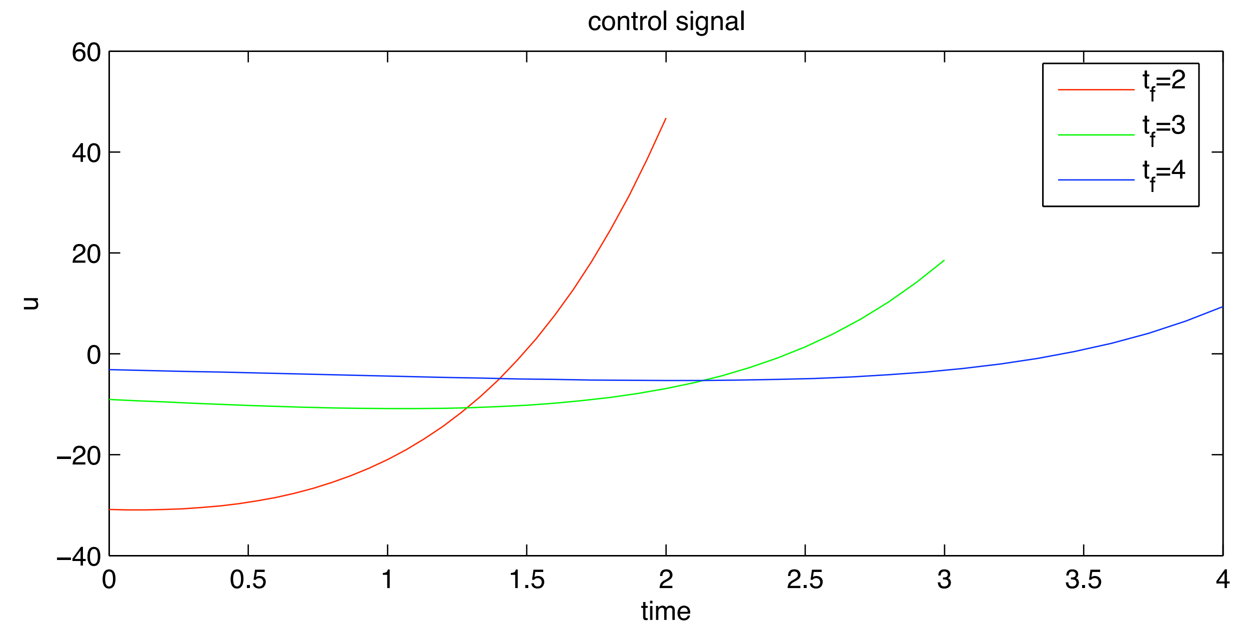
$$\mathbf{W}_c^{-1}(2) = \begin{bmatrix} 84.91 & -54.79 \\ -54.79 & 37.39 \end{bmatrix}$$

- Suppose starting position is $\mathbf{x}(0) = \begin{bmatrix} 10 \\ -1 \end{bmatrix}$

$$\begin{aligned}
\mathbf{u}(t) &= -\mathbf{B}'e^{\mathbf{A}'(t_f-t)}\mathbf{W}_c^{-1}(t_f)\left[e^{\mathbf{A}t_f}\mathbf{x}_0 - \mathbf{x}_f\right] \\
&= -\begin{bmatrix} \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} e^{-\frac{1}{2}(2-t)} & 0 \\ 0 & e^{-(2-t)} \end{bmatrix} \begin{bmatrix} 84.91 & -54.79 \\ -54.79 & 37.39 \end{bmatrix} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^{-2} \end{bmatrix} \begin{bmatrix} 10 \\ -1 \end{bmatrix} \\
&= -58.82e^{\frac{1}{2}t} + 27.96e^{t}
\end{aligned}$$

- As t gets short the control signal and state path become more extreme:



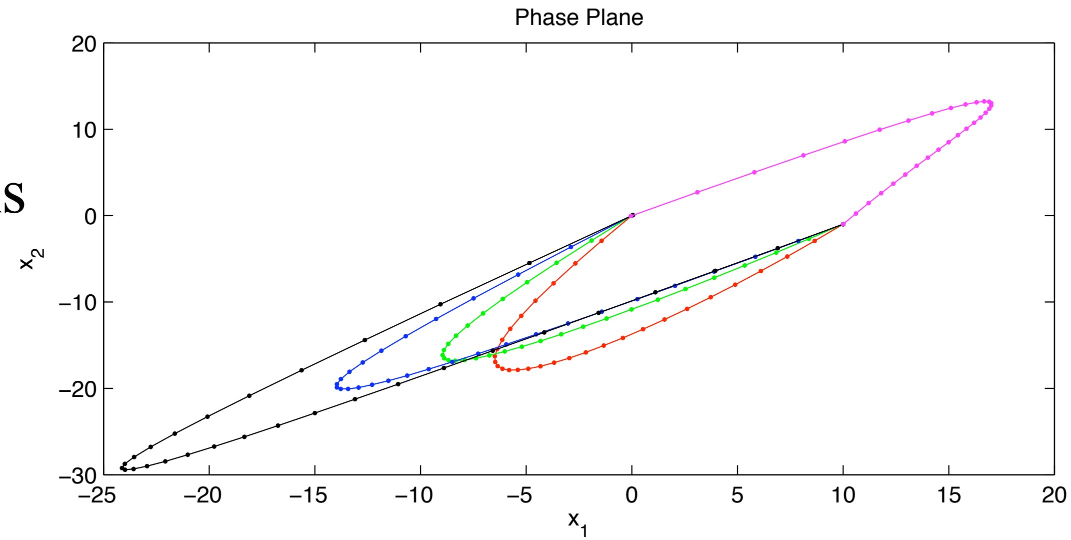


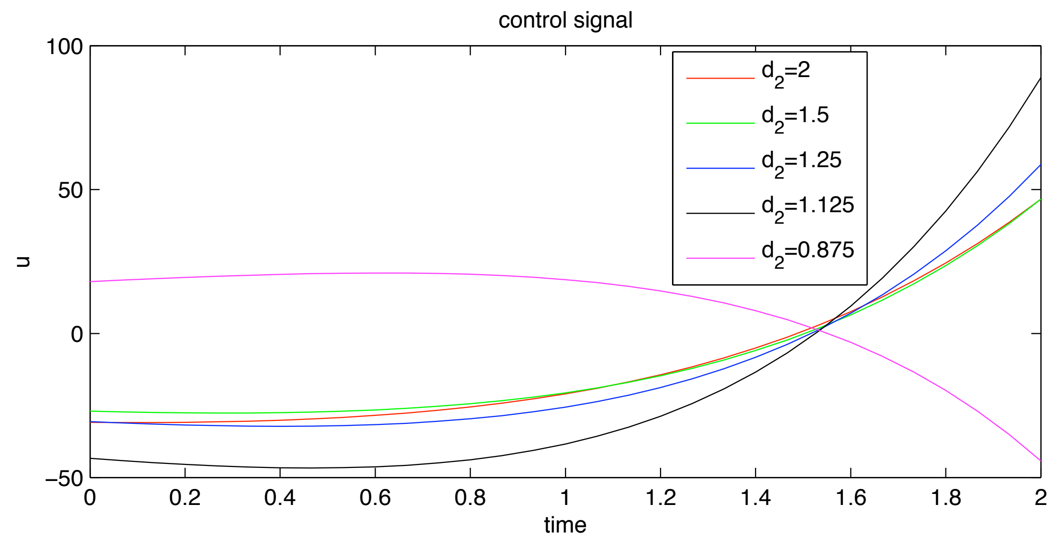
- As $d_1 \rightarrow d_2$ the controllability matrix becomes more singular: When $d_1 = \{1.25, 1.1, 1\}$ the controllability matrices are:

$$\left\{ \begin{bmatrix} 1.25 & -1.237 \\ 1.00 & -1.00 \end{bmatrix}, \begin{bmatrix} 1.10 & -1.089 \\ 1.00 & 1.00 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \right\}$$

So rank goes from 2 to 1.

- For same t_f, the control paths (of state and control signal) become more extreme:





Linear Time Varying (LTV)

- We define controllability in terms of time interval [t_0,t_f]
- Let $\mathbf{\Phi}(t_f, t_0)$ be the system's transfer matrix,
(for LTI system $\mathbf{\Phi}(t_f, t_0)$ = expm(A(t_f - t_0)))
then the system's state is
$$\mathbf{x}(t_f) = \mathbf{\Phi}(t_f, t_0)\mathbf{x}_0 + \int_{t_0}^{t_f} \mathbf{\Phi}(t_f, \tau)\mathbf{B}(\tau)\mathbf{u}(t)d\tau$$
- Define the controllability grammian:
$$\mathbf{W}_c(t_0, t_f) = \int_{t_0}^{t_f} \mathbf{\Phi}(t_f, \tau)\mathbf{B}(\tau)\mathbf{B}'(\tau)\mathbf{\Phi}'(t_f, \tau)d\tau$$
- As for LTI systems, the following control will work:
$$\mathbf{u}(t) = -\mathbf{B}'(t)\mathbf{\Phi}'(t_f, t)\mathbf{W}_c^{-1}(t_0, t_f)[\mathbf{\Phi}'(t_f, t_0)\mathbf{x}_0 - \mathbf{x}_f]$$