EEE 598 B

SPECIAL TOPICS: ADAPTIVE CONTROL (MWF 12^{40} – 1^{30})

Instructor: KOSTAS TSAKALIS

- ERC 333
- ph # 965-1467
- off. hrs: TUE: 10^{30} – 12^{00}
  WED: 1^{45} – 3^{30}
  OR BY APPOINTMENT.

SCOPE OF THE COURSE:

→ ANALYSIS + DESIGN OF ADAPTIVE CONTROLLERS

- ANALYTICAL TOOLS.
- UNDERLYING PRINCIPLES.
- DESIGN GUIDLINES.
- WHAT CAN GO WRONG ?
- MODIFICATIONS + IMPROVEMENT OF ADAPTIVE CONTROLLERS

→ APPLICATIONS OF ADAPTIVE CONTROL

REFERENCES

(*) 1. CLASS NOTES

(*) 2. KEY JOURNAL PUBLICATIONS

} PERIODIC CIRCULATION THROUGH LIBRARY (FILE "EEE 598 B")

3. OTHER RELATED PUBLICATIONS

4. USEFUL BOOKS

a) LINEAR SYSTEMS, e.g. • KAILATH, "LINEAR SYSTEMS", PRENTICE HALL, 1980 etc.

b) NON LINEAR SYSTEMS, e.g. • VIDYASAGAR, "NONLINEAR SYSTEMS ANALYSIS", PRENTICE HALL, 1978
• R. MILLER + A. MICHELL "ORDINARY DIFFERENTIAL EQUATIONS", ACADEMIC PRESS 1982

c) ESTIMATION THEORY / PARAMETER ESTIMATION (any standard textbook)

d) ADAPTIVE SYSTEMS

- K.S. NARENDRA + A.M. ANNASWAMY, "STABLE ADAPTIVE SYSTEMS", PRENTICE HALL, 1989.
- G.C. GOODWIN + K.S. SIN, "ADAPTIVE FILTERING PREDICTION + CONTROL", PRENTICE HALL, 1984.
- Y. LANDAU, "ADAPTIVE CONTROL : THE MODEL REFERENCE APPROACH", MARCEL-DEKKER, 1979.

e) OTHER

- C. DESOER + M. VIDYASAGAR, "FEEDBACK SYSTEMS: INPUT-OUTPUT PROPERTIES", ACADEMIC PRESS, 1975
- B. FRANCIS, "A COURSE IN H_∞ CONTROL THEORY", SPRINGER VERLAG, 1987

3

- GRADING POLICY / COURSE FORMAT
  - → THEORY (Proofs, Mathematical derivations etc.)
  - → "APPLICATIONS" (Examples, Simulations)
- 6 HW SETS
  - 2 WKs EACH
  - BEST 5/6 ⟶ 70%
- FINAL
  - TAKEHOME, 1 WK ⟶ 30%

- **REMARKS ON SIMULATIONS**

— SOFTWARE PACKAGES / LIBRARY SUBROUTINES

1) REQUIRED : SOLUTION OF ORDINARY DIFFERENTIAL EQUATIONS (ODE's)

$$\underset{\circ}{\dot{x}} = f(x) \quad \text{OR} \quad \begin{aligned} \dot{x}(1) &= f_1(x(1), x(2), \dots) \\ \dot{x}(2) &= f_2(x(1), x(2), \dots) \\ &\vdots \end{aligned}$$

E.g. IMSL (RUNGE-KUTTA, GEAR etc)

2) RECOMMENDED : PLOTTING (x,y)

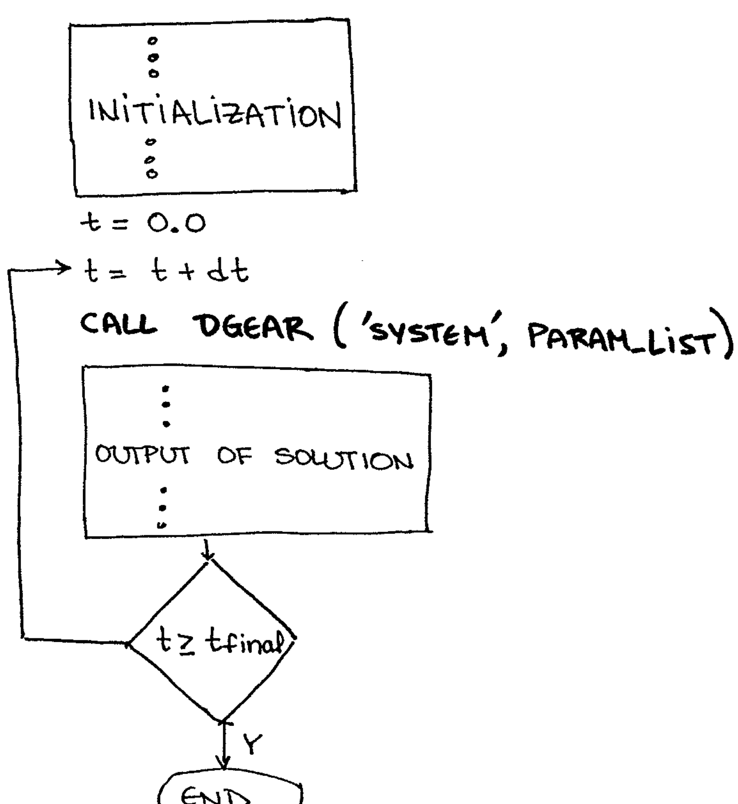
— USEFUL SOFTWARE PACKAGES FOR PLOTTING AND/OR SIMULATION

- MATLAB
- MATRIX-x
- CTRL-C

e.t.c.

— TYPICAL STRUCTURE OF A FORTRAN PROGRAM FOR THE SIMULATION OF ADAPTIVE SYSTEMS

MAIN :



1) DGEAR: NAME OF IMSL SUBROUTINE SOLVING A SYSTEM OF ODE's DEFINED BY A USER-SUPPLIED SUBROUTINE -
2) SYSTEM: NAME OF THE USER SUPPLIED SUBROUTINE DEFINING THE SYSTEM OF ODE's TO BE SOLVED
3) PARAM_LIST : LIST OF PARAMETERS

```
SUBROUTINE SYSTEM (N,XDOT, X, T)
XDOT (1) = f( X(1), X(2),... )
XDOT (2) = f ( X(1) , X(2),... )
   :
XDOT (N) = f ( X(1), X(2),... )
RETURN
END
```

$$\dot{x} = f(x) \quad \left\{ \begin{array}{l} \text{SYSTEM (PLANT)} \\ + \\ \text{CONTROLLER} \\ + \\ \text{ADAPTIVE LAW} \end{array} \right.$$

COMMENTS:

1) CONSULT MANUALS FOR THE PRECISE FORMAT
2) DOUBLE PRECISION VARIABLES (ALWAYS !)
3) LIBRARY SUBROUTINES (EFFICIENT + RELIABLE)
4) A FEW PARAMETERS MAY BE PASSED FROM MAIN TO SYSTEM THROUGH "COMMON" BLOCKS
5) AVOID "TOO GENERAL" SUBROUTINES FOR $\dot{x}=f(x)$
7 (EXTENSIVE DEBUGGING REQUIREMENTS /TIME CONSUMING)

INTRODUCTION

FEEDBACK SYSTEMS : MEANS TO COUNTERACT UNCERTAINTY

e.g. MODELING ERROR, EXTERNAL (UNMEASURED) DISTURBANCES

LINEAR FEEDBACK

$\wedge$

ADAPTIVE FEEDBACK

$\wedge$

NON LINEAR FEEDBACK

SOME DEFINITIONS + NOTATION

1. SYSTEM : An aggregation of "objects" united by some form of interaction

2. DYNAMICAL SYSTEM : One or more aspects of the system change with time

3. INPUTS : Influences originating outside the system; not directly affected by the behavior of the system.

4. OUTPUTS : Quantities of Interest, affected by the inputs.

5. CONTROL INPUTS : Inputs determined by the designer.

6. STATES : Quantities (signals) which describe the dynamical behavior of the system

---

EX. SYSTEMS DESCRIBED BY ORDINARY VECTOR DIFFERENTIAL EQUATIONS.

$$\left.\begin{aligned} \frac{dx(t)}{dt} \triangleq \dot{x}(t) &= f(x(t), u(t), \theta, t) \\ y(t) &= h(x(t), \theta, t) \end{aligned}\right\} \text{MATHEMATICAL DESCRIPTION OF A DYNAMICAL SYSTEM}$$

$t \in \mathbb{R}^+$: time

$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$: states

$u = \begin{bmatrix} u_1 \\ \vdots \\ u_p \end{bmatrix} \in \mathbb{R}^p$: inputs

$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m$: outputs

$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_r \end{bmatrix} \in \mathbb{R}^r$: "parameters"

f, h mappings

$f: \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^r \times \mathbb{R}^+ \longrightarrow \mathbb{R}^n$

$h: \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^+ \longrightarrow \mathbb{R}^m$

$$u \rightarrow \boxed{\begin{aligned} \dot{x} &= f(x,u,\theta,t) \\ y &= h(x,\theta,t) \end{aligned}} \rightarrow y \quad (\text{input } \theta)$$

OR

$$u \rightarrow \boxed{P(\theta)} \rightarrow y \qquad (*)$$

e.g. $P(\theta): \begin{cases} \dot{x} = \theta x + u \\ y = x \end{cases}$

$x, u, y, \theta \in \mathbb{R}$; θ : constant

- $\{P(\theta),\ \theta \in \mathbb{R}\}$: A FAMILY OF SYSTEMS

i.e. $\theta = 1$: $(*) \leftrightarrow u \rightarrow \boxed{\frac{1}{s-1}} \rightarrow y$

$\theta = 0$: $(*) \leftrightarrow u \rightarrow \boxed{\frac{1}{s}} \rightarrow y$ etc.

11

EXAMPLE OF AN ADAPTIVE SYSTEM

Consider the system

$$\dot{x} = a x + u$$
$$y = x$$

where a is an unknown negative constant and u, y are available for measurment.

Further, construct the system

$$\dot{w} = \theta(t)\, w + u$$

Q: Determine a function $f(y,w)$ s.t. selecting

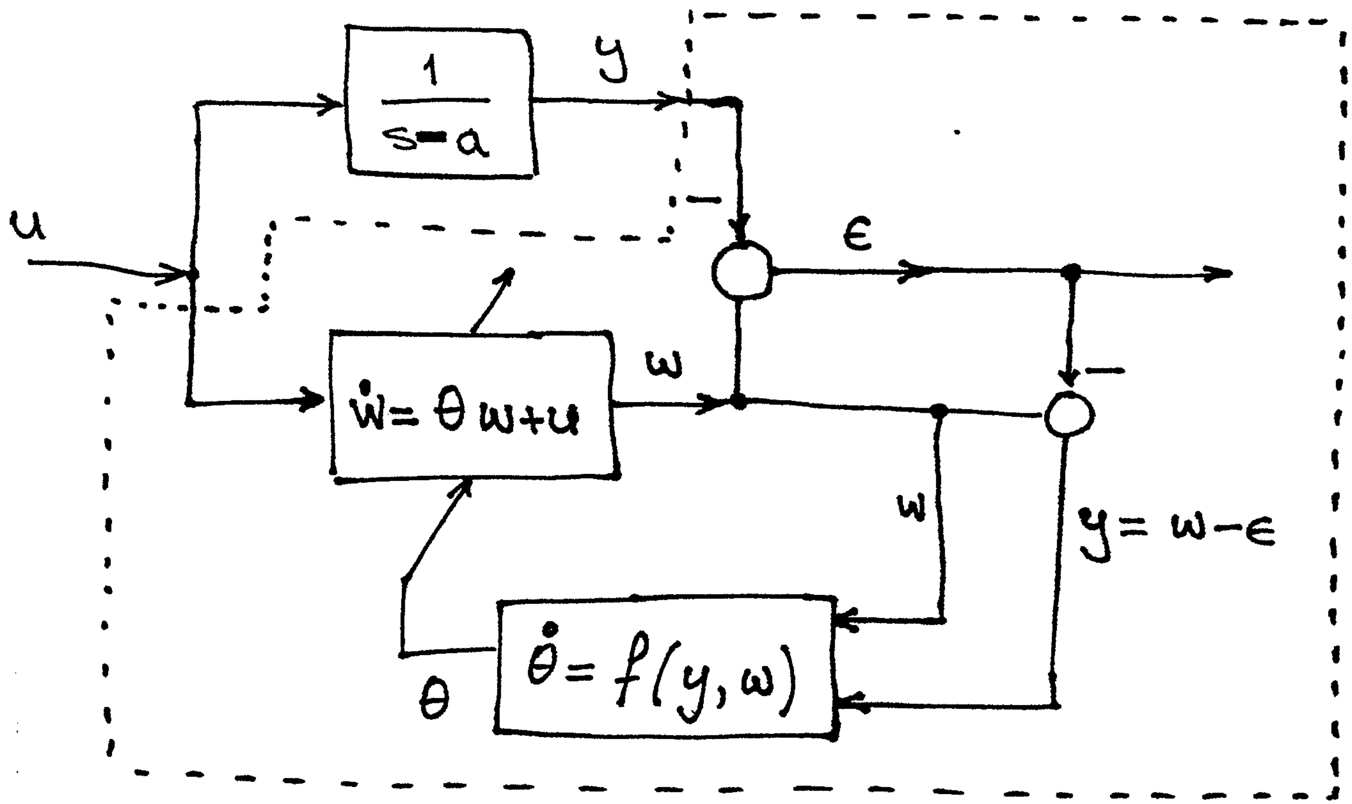
$$\dot{\theta} = f(y,w)$$

we have

$$w(t) - y(t) \longrightarrow 0 \quad \text{as } t \to \infty$$

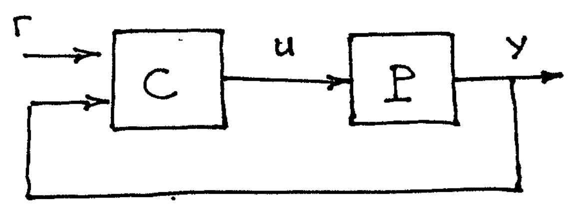
for any bounded input u and any initial conditions $x(0)$, $w(0)$, $\theta(0)$

i.e.



"DEFINITION": An adaptive system is a system which is provided with a means of continuously monitoring its own performance in relation to a given figure of merit or optimal condition and a means of modifying its own parameters or structure by a closed-loop action so as to approach this optimum.

FEEDBACK SYSTEMS



P : PLANT, SYSTEM TO BE CONTROLLED

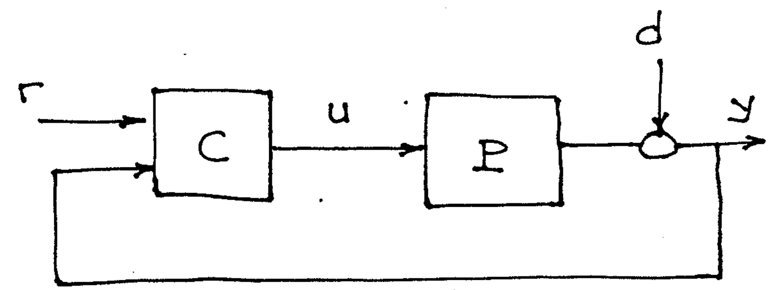
C : CONTROLLER

r : reference signal

u : control input

y : output of the plant.

Q: DESIGN C s.t. y "follows" r as CLOSELY AS POSSIBLE, MINIMIZING THE EFFECTS OF EXTERNAL DISTURBANCES + MODELING UNCERTAINTY



- d : EXTERNAL DISTURBANCE ; PARTIALLY KNOWN (e.g. d = constant)

- P : PARTIALLY KNOWN DYNAMICAL SYSTEM

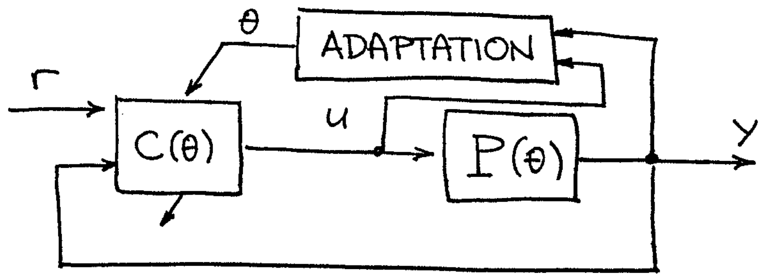
  e.g. $P = P_o + \Delta P$

  P_o : KNOWN LTI SYSTEM (NOMINAL PLANT)

  ΔP : MODELING UNCERTAINTY

  $\|\Delta P(s)\|_\infty < 1$

ADAPTIVE CONTROL



- $P(\theta) = P_0(\theta) + \Delta P$

  PLANT DESCRIPTION PARAMETRIZED BY θ. (FAMILY OF NOMINAL PLANTS)

- ▶ GIVEN θ, $P_0(\theta)$ IS KNOWN
- ▶ θ : PARTIALLY KNOWN

  e.g. $\|\theta\| < 1$

- ▶ ΔP : PARTIALLY KNOWN

  e.g. $\|\Delta P(s)\|_\infty < 1$

(NOTE: ΔP MAY DEPEND ON θ)

Q: DESIGN THE "ADAPTATION" + $C(\theta)$ s.t. THE CLOSED LOOP SYSTEM HAS CERTAIN DESIRED PROPERTIES

e.g. $y \to y_m$ as $t \to \infty$

for any bounded r & INIT. COND.

where, y_m is the output of a reference model with input r

$$y_m = M r$$

M: A KNOWN, "WELL-BEHAVED" DYNAMICAL SYSTEM.
(STABILITY, BANDWIDTH, DC GAIN, ROLL-OFF)

W.R.T. MODELING UNCERTAINTY:

1). **NON-ADAPTIVE FEEDBACK**

"UNSTRUCTURED UNCERTAINTY"

(e.g. $\|\Delta P\|_\infty < 1$)

2) **ADAPTIVE FEEDBACK**

"PARTIALLY STRUCTURED UNC/TY"

(e.g. $\|\Theta\| < 1$; $\|\Delta P\|_\infty < 1$)

W.R.T. TYPE OF FEEDBACK

1). **NON-ADAPTIVE FEEDBACK**

SIGNAL INFORMATION

(e.g. y)

2). **ADAPTIVE FEEDBACK**

SIGNAL & OPERATOR INFORMATION

(e.g. y, θ)

EXAMPLE OF AN ADAPTIVE CONTROLLER

CONSIDER THE PLANT

$$\dot{y} = a_p y + u$$

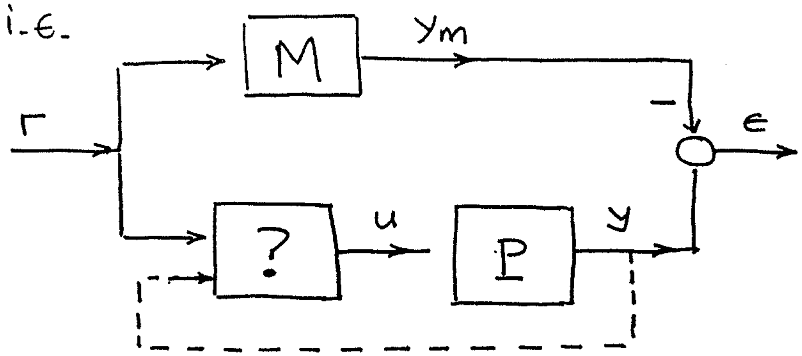
where a_p is an unknown constant.

Q: DESIGN u s.t. GIVEN THE REFERENCE MODEL

$$\dot{y}_m = a_m y_m + r \quad ; \quad a_m < 0$$

$(y - y_m) \to 0$ as $t \to \infty$, for ANY BOUNDED REFERENCE INPUT r AND ANY INITIAL CONDITIONS $y(0)$, $y_m(0)$.

i.e.



where,

- $M: r \to y_m$; $\dot{y}_m = a_m y_m + r$
- $P: u \to y$; $\dot{y} = a_p y + u$

IDEA: IF a_p WERE KNOWN, WE COULD SELECT

$$u = K^* y + r$$

$$K^* = a_m - a_p$$

THEN:

$$\dot{y} = a_p y + K^* y + r = a_m y + r$$

$$\Rightarrow \dot{(y - y_m)} = a_m (y - y_m)$$

$$\Rightarrow y - y_m = e^{a_m t} (y - y_m)(0) \to 0 \text{ as } t \to \infty$$

HOWEVER, K^* is UNKNOWN.

LET

$$u = Ky + r$$

$$\dot{K} = f(y_m, y)$$

where $f(\cdot,\cdot)$ is A FUNCTION TO BE DETERMINED (e.g. an ESTIMATOR)

THEN,

$$\dot{y} = \underbrace{(a_p + k)}_{\theta} y + r$$

$$\dot{\theta} = \dot{K} = f(y_m, y)$$

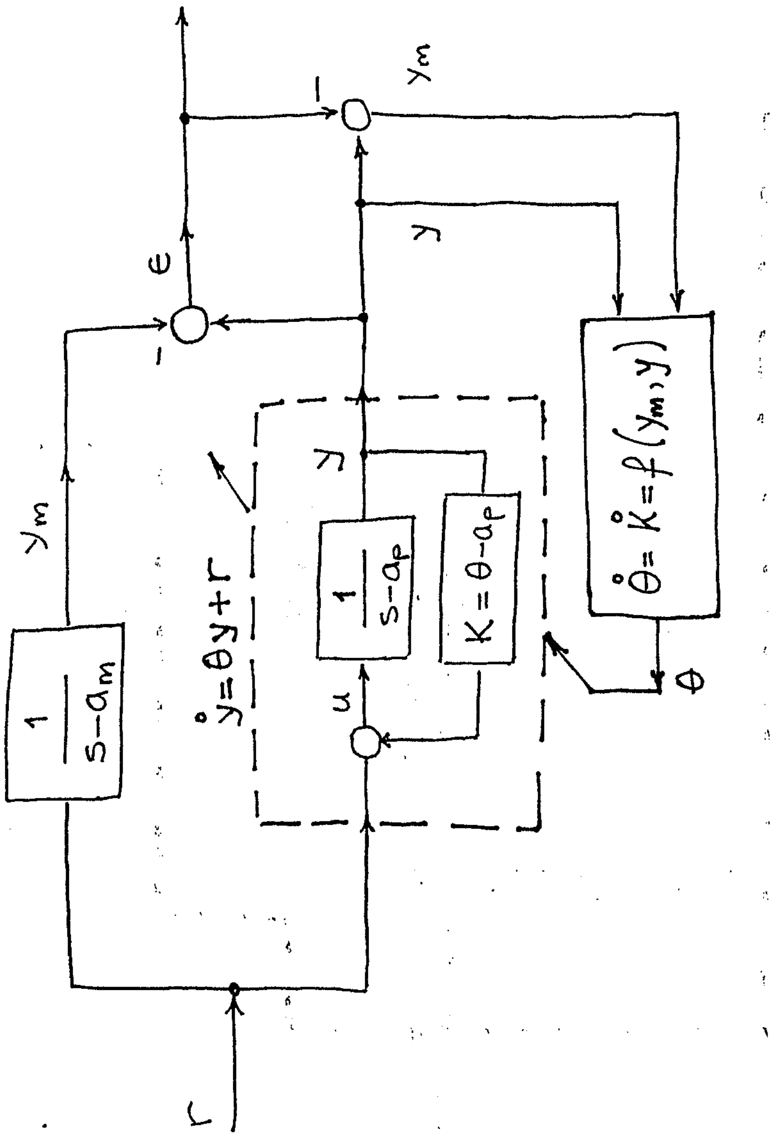
∴ SELECT $f(\cdot,\cdot)$ s.t. THE OUTPUT OF

$$\dot{y} = \theta y + r$$

"TRACKS" THE OUTPUT OF

$$\dot{y}_m = a_m y_m + r$$

I.E.



I. MATHEMATICAL PRELIMINARIES

GUIDED BY THE PREVIOUS EXAMPLE WE NEED SOME MATH BACKGROUND ON THE FOLLOWING TOPICS:

- 1). CONTROL LAW DESIGN

  e.g. $u = Ky + r$

  $K(\theta) = \theta - a_P$

— TYPICAL PROBLEM: GIVEN A FAMILY OF PLANTS $P(\theta)$, FIND $C(\theta)$ S.T. FOR ANY GIVEN θ, THE FEEDBACK SYSTEM

$$y = P(\theta)\,u \;;\; u = -C(\theta)\,y$$

IS STABLE.

(LINEAR SYSTEM THEORY, STABILIZATION POLE PLACEMENT etc.)

- 2). ADAPTIVE LAW DESIGN

  e.g. $\dot{\theta} = f(y_m, y)$

NON LINEAR SYSTEMS, STABILITY, LYAPUNOV THEORY

- 3). ADAPTIVE CONTROL SYSTEMS

  (1) + (2) + I/O OPERATORS

NONLINEAR VECTOR ODEs

$$\dot{x} = f(t, x(t), u(t)), \quad t \geq 0; \; x(0)$$

1) EXISTENCE OF SOLUTIONS
2) UNIQUENESS "
3) SOLUTION DEFINED OVER THE ENTIRE half-Line $[0, \infty)$
4) CONTINUOUS DEPENDENCE ON $x(0)$

EX.

- $\dot{x} = \dfrac{1}{2x(t)} \quad t \geq 0 \quad x(0) = 0$

  $\rightsquigarrow x_1(t) = t^{1/2} \; ; \; x_2(t) = -t^{1/2}$

- $\dot{x} = x^2 \quad t \geq 0 \quad x(0) = x_0 > 0$

  $\rightsquigarrow x(t) = \dfrac{x_0}{1 - t x_0}$

EXISTENCE AND UNIQUENES OVER $[0, \frac{1}{x_0})$

However $x(t) \to \infty$ as $t \to \frac{1}{x_0}$ AND $x(t)$ is NOT DEFINED AT $t = \frac{1}{x_0}$ (FINITE ESCAPE TIME)

NOTE THAT IN GENERAL IT MAY NOT BE POSSIBLE TO OBTAIN THE EXACT SOLUTION OF AN ODE

$\therefore$ WE NEED SOME "TOOLS" TO ESTABLISH

1) "WELL-BEHAVEDNESS"

2) "BOUNDS"

OF SOLUTIONS OF ODES WITHOUT ACTUALLY SOLVING THEM.

W.O.L.O.G. LET US CONSIDER THE ODE

$$\dot{x} = f(t,x) \qquad (1)$$

• <u>DEF</u>. THE SYSTEM (1) IS SAID TO BE <u>AUTONOMOUS</u> IF $f(t,x)$ IS INDEPENDENT OF t AND IS SAID TO BE <u>NON AUTONOMOUS</u> OTHERWISE.

• <u>DEF</u> $x_0 \in \mathbb{R}^n$ IS SAID TO BE AN <u>EQUILIBRIUM</u> OF (1) AT TIME $t_0 \in \mathbb{R}_+$ IF $f(t,x_0) = 0 \quad \forall\ t \geq t_0$

(STATIONARY, SINGULAR POINT)

<u>REM</u> IF (1) IS AUTONOMOUS THEN $x_0 \in \mathbb{R}^n$ IS AN EQUILIBRIUM POINT OF (1) AT SOME TIME <u>IFF</u> IT IS AN EQUILIBRIUM POINT OF (1) AT ALL TIMES

<u>NOTE:</u> IF x_0 IS AN EQUILIBRIUM POINT OF (1) AT $t=t_0$, THEN FOR $t_1 \geq t_0$

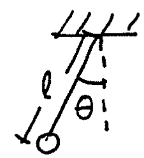
$$\dot{x} = f(t, x(t)), \quad t \geq t_1 ; \quad x(t_1) = x_0$$

HAS THE UNIQUE SOLUTION $x(t) = x_0,\ \forall t \geq t_1$

<u>EX</u>. $\dot{x} = Ax$ HAS THE EQUILIBRIUM POINTS : $x_0 \in \{ x_0 \mid Ax_0 = 0 \} = \mathcal{N}(A)$

<u>DEF</u> AN EQUILIBRIUM POINT x_0 OF (1), AT t_0, IS SAID TO BE <u>ISOLATED</u> IF $\exists\ N(x_0) \subset \mathbb{R}^n$: $N(x_0)$ CONTAINS NO EQUILIBRIUM POINTS AT t_0 OF (1) OTHER THAN x_0.

<u>EX</u>. CONSIDER THE MOTION OF A FRICTIONLESS PENDULUM



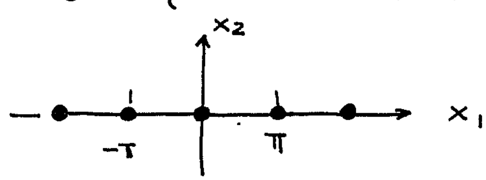
$$\ddot{\theta} + \frac{g}{\ell}\sin[\theta(t)] = 0$$

i.e. $\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \triangleq \begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} x_2 \\ -g/\ell \, \sin(x_1) \end{bmatrix}$

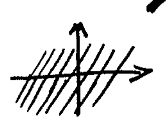
EQUILIBRIUM $x_0 = [x_{10}, x_{20}]^T$ iff

$x_{20} = 0$, $\sin(x_{10}) = 0$

i.e., $x_0 \in \{ x_0 \in \mathbb{R}^2 : x_0 = (n\pi, 0)^T, n \in \mathbb{Z} \}$



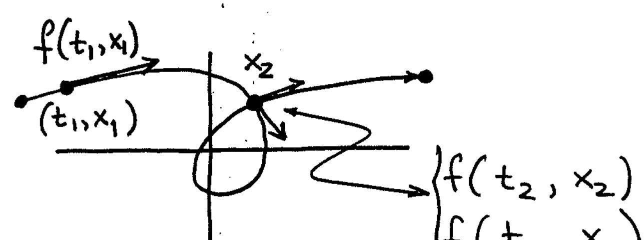
<u>EX</u>. $\dot{x} = Ax$. $x \in \mathbb{R}^2$

1. A : NONSINGULAR $\Rightarrow x_0 = 0$ 
2. A : RANK 1 $\Rightarrow x_0 \in N(A)$ 
3. $A = 0$ $x_0 \in \mathbb{R}^2$ 

<u>REM</u>. CONSIDER

$$\dot{x} = f(t, x) \; ; \; x(0) \quad (*)$$

f : CONTINUOUS, AND SUPPOSE $\mathcal{C}$ IS A SOLUTION TRAJECTORY OF (*) PASSING THROUGH (t_1, x_1). THEN THE VECTOR $f(t_1, x_1)$ IS TANGENT TO $\mathcal{C}$ AT (t_1, x_1)



f IS COMMONLY REFERRED TO AS THE "VELOCITY VECTOR FIELD" OR "VECTOR FIELD" OF (*).

▶ LINEAR VECTOR SPACES

<u>DEF</u>. A FIELD IS A SET $\mathbb{F}$ TOGETHER WITH TWO OPERATIONS $+, \cdot : \mathbb{F}\times\mathbb{F} \to \mathbb{F}$

$(\mathbb{F},+)$ ABELIAN GROUP:

1) $\alpha+(\beta+\gamma) = (\alpha+\beta)+\gamma \quad \forall \alpha,\beta,\gamma \in \mathbb{F}$
2) $\exists\, 0 \in \mathbb{F} : \ \alpha+0=\alpha \quad \forall \alpha \in \mathbb{F}$
3) $\forall \alpha\ \exists -\alpha : \ \alpha+(-\alpha)=0$
4) $\alpha+\beta = \beta+\alpha \quad \forall \alpha,\beta \in \mathbb{F}$

$(\mathbb{F},\cdot)$ ABELIAN GROUP:

1) $\alpha\cdot(\beta\cdot\gamma) = (\alpha\cdot\beta)\cdot\gamma \quad \forall \alpha,\beta,\gamma$
2) $\exists\, 1 \in \mathbb{F} : \ \alpha\cdot 1 = \alpha \quad \forall \alpha \in \mathbb{F}$
3) $\forall \alpha \in F\ \exists\, \alpha^{-1} \in \mathbb{F} : \ \alpha\alpha^{-1} = 1$
4) $\alpha\cdot\beta = \beta\cdot\alpha$

<u>EX</u>. $\mathbb{R}$, $\mathbb{C}$

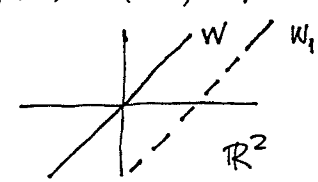
<u>DEF</u>. A VECTOR SPACE OVER A FIELD $\mathbb{F}$ IS A SET V TOGETHER WITH TWO OPERATIONS

$$+ : V\times V \to V, \quad \cdot : \mathbb{F}\times V \to V$$

s.t.
1) $(V,+)$ IS AN ABELIAN GROUP
2) $(\alpha\cdot\beta)\cdot v = \alpha\cdot(\beta\cdot v) \quad \forall \alpha,\beta \in \mathbb{F}, \forall v \in V$
3) $(\alpha+\beta)v = \alpha v + \beta v$ -1-
4) $\alpha(v+w) = \alpha v + \alpha w \quad \forall \alpha \in \mathbb{F}, \forall v,w \in V$
5) $1\cdot v = v \quad \forall v \in V$.

<u>DEF</u>. LET $(V,\mathbb{F})$ BE A VECTOR SPACE AND $W \subseteq V$, $W \neq \emptyset$. THEN $(W,\mathbb{F})$ IS SAID TO BE A SUBSPACE OF V IF $(W,\mathbb{F})$ IS A VECTOR SPACE, i.e.
1) $x+y \in W \quad \forall x,y \in W$
2) $\alpha x \in W \quad \forall x \in W, \alpha \in \mathbb{F}$.

<u>EX</u>. $(V,\mathbb{F}) = (\mathbb{R}^2, \mathbb{R})$



W_1 NOT A VECTOR SPACE $\Rightarrow$ NOT A SUBSPACE OF $\mathbb{R}^2$.

▶ <u>NORMED LINEAR SPACES</u>

DEF A NORMED LINEAR SPACE is AN ORDERED PAIR $(X, \|\cdot\|)$, WHERE X is A LINEAR VECTOR SPACE AND $\|\cdot\|$ is A REAL VALUED FUNCTION ON X ('NORM') s.t.

1) $\|x\| \geq 0 \quad \forall x \in X$; $\|x\| = 0 \Leftrightarrow x = 0_X$
2) $\|ax\| = |a| \cdot \|x\| \quad \forall x \in X, \ a \in \mathbb{F}$
3) $\|x+y\| \leq \|x\| + \|y\| \quad \forall \ x, y \in X$

I.E. $\|x\|$ is A MEASURE OF THE "SIZE" OF X OR THE DISTANCE OF X FROM 0

DEF A SEQUENCE $(x_n)_1^\infty$ IN A NORMED LINEAR SPACE $(X, \|\cdot\|)$ is SAID TO CONVERGE TO $x_0 \in X$ if $\|x_n - x_0\| \to 0$ as $n \to \infty$

EQUIVALENTLY, $\forall \varepsilon > 0 \ \exists \ N(\epsilon)$:

$$\|x_n - x_0\| < \varepsilon \quad \forall \ n \geq N(\epsilon)$$

DEF Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be two normed linear spaces AND f be a function $f: X \to Y$. We say that

- f is <u>CONTINUOUS at $x_0 \in X$</u> if $\forall \ \varepsilon > 0 \ \exists \ \delta(\varepsilon, x_0) > 0$ s.t.

$$\|f(x_0) - f(y)\|_Y < \varepsilon$$

whenever $\|x_0 - y\|_X < \delta(\varepsilon, x_0)$

- f is <u>CONTINUOUS</u> if it is continuous at every $x \in X$
- f is <u>UNIFORMLY CONTINUOUS</u> if it is continuous <u>and</u> $\forall \ \varepsilon > 0 \ \exists \ \delta(\varepsilon) > 0$ s.t. $\|f(x) - f(y)\|_Y < \varepsilon$ whenever $\|x - y\|_X < \delta(\varepsilon)$.

__DEF__. A sequence $\{x_n\}_1^\infty$ in a normed linear space $(X, \|\cdot\|)$ is said to be __Cauchy sequence__ if $\forall \varepsilon > 0$ $\exists\ N(\varepsilon) \in \mathbb{N}$ s.t.

$$\|x_n - x_m\| < \varepsilon \quad \text{whenever } n, m \geq N(\varepsilon)$$

__REM__. CONVERGENT SEQ. $\Rightarrow$ CAUCHY

__Prf__. Suppose $\{x_n\}_1^\infty$ is convergent ($x_n \in (X, \|\cdot\|)$, $x_n \to x_0 \in (X, \|\cdot\|)$)

Let $\varepsilon > 0$. Select N:

$$\|x_n - x_0\| < \frac{\varepsilon}{2} \qquad \forall n \geq N$$

Then, for $n, m \geq N$

$$\|x_n - x_m\| \leq \|x_n - x_0\| + \|x_m - x_0\| < \varepsilon$$

__DEF__ A Normed Linear Space $(X, \|\cdot\|)$ is said to be a COMPLETE NORMED LINEAR SPACE or a __BANACH__ space if every CAUCHY SEQUENCE in X converges (to an element in X).

__EX__. Consider the linear vector space $\mathbb{R}^n$ together with the function $\|\ \|_\infty$: $\mathbb{R}^n \to \mathbb{R}$ defined by

$$\|x_\infty\| = \max_{1 \leq i \leq n} |x_i|$$

Then $(\mathbb{R}^n, \|\cdot\|_\infty)$ is A NORMED LINEAR SPACE AND A BANACH SPACE.

(NOTE: $\|\cdot\|_\infty$ is a norm. Why?).

The same is true for the linear vector spaces:

▶ 1). $(\mathbb{R}^n, \|\cdot\|_1)$

where $\|\cdot\|_1 : \mathbb{R}^n \to \mathbb{R}$

$$\|x\|_1 = \sum_1^n |x_i|$$

▶ 2). $(\mathbb{R}^n, \|\cdot\|_p)$

where $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$

$$\|x\|_p = \left\{ \sum_1^n |x_i|^p \right\}^{1/p}$$

In particular, if $p=2$, $\|\cdot\|_2$ is also known as the <u>Euclidean norm</u> or l_2-norm on $\mathbb{R}^n$.

NOTE THAT $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a different entity than $(\mathbb{R}^n, \|\cdot\|_1)$ or $(\mathbb{R}^n, \|\cdot\|_2)$ even though the underlying linear vector space is the same $(\mathbb{R}^n)$

<u>SPECIAL PROPERTIES OF $\mathbb{R}^n$</u> $(+\ \mathbb{C}^n)$.

• Let $\|\cdot\|_\alpha$, $\|\cdot\|_\beta$ be any two norms on $\mathbb{R}^n$. Then there exist <u>finite</u> positive constants K_1, K_2 s.t.

$$K_1 \|x\|_\alpha \le \|x\|_\beta \le K_2 \|x\|_\alpha$$

$\forall\ x \in \mathbb{R}^n$.

(Such norms are called "equivalent norms")

<u>eg.</u> $\|x\|_\infty \le \|x\|_1 \le n \|x\|_\infty$

$\|x\|_\infty \le \|x\|_2 \le n^{1/2} \|x\|_\infty$

<u>Consequence</u>: Convergence in $\mathbb{R}^n$ is independent of the norm used.

• Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$, $\{x_n\}_1^\infty$ be a sequence in $\mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$. Then $\|x_n - x_0\| \to 0$ as $n \to \infty$

iff each component sequence $\{x_n^{(i)}\}_1^\infty$ converges to $x_0^{(i)}$ for $i = 1, \dots, n$.

- Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$, $x(\cdot)$ be a function mapping $\mathbb{R} \to \mathbb{R}^n$. Then $x(\cdot)$ is continuous (from $(\mathbb{R}, |\cdot|)$ into $(\mathbb{R}^n, \|\cdot\|)$) iff each of the component functions $x_i(\cdot)$ is a continuous function on $\mathbb{R}$.

THE NORMED LINEAR SPACE $C^n[a,b]$

Let $\|\cdot\|$ be any given norm on $\mathbb{R}^n$ and $C^n[a,b]$ denote the set of all continuous functions $[a,b] \to \mathbb{R}^n$. Define $\|\cdot\|_c : C^n[a,b] \to R$ as

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

Then, $\|\cdot\|_c$ is a norm on $C^n[a,b]$

Prf. Axioms 1 & 2 are straight forward

To test 3, Let $x(\cdot), y(\cdot) \in C^n[a,b]$

Then
$$\|x(\cdot) + y(\cdot)\| = \max_{t \in [a,b]} \|x(t) + y(t)\|$$
$$\leq \max_{t \in [a,b]} \|x(t)\| + \|y(t)\|$$

(by the triangle inequality on $\mathbb{R}^n$)

$$\leq \max_{t \in [a,b]} \|x(t)\| + \max_{t \in [a,b]} \|y(t)\|$$
$$\triangleq \|x(\cdot)\|_c + \|y(\cdot)\|_c.$$

INNER PRODUCT SPACES

Def. An Inner Product Space is a linear vector space X together with a function $\langle \cdot, \cdot \rangle : X \times X \to F$ (the associated field) s.t.

1). $\langle x,y\rangle = \overline{\langle y,x\rangle} \quad \forall\ x,y \in X$.

($\bar{\cdot}$ denotes conjugation)

2). $\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle \quad \forall\ x,y,z \in X$

3) $\langle x, \alpha y\rangle = \alpha\langle x,y\rangle \quad \forall\ x,y \in X$
$\forall\ \alpha \in \mathbb{F}$.

4). $\langle x,x\rangle \geq 0 \quad \forall\ x \in X$

$\langle x,x\rangle = 0$ iff $x = 0$.

<u>THM</u>. Given an inner product space X, define $\|\cdot\| : X \to \mathbb{R}$ by

$$\|x\| = \langle x,x\rangle^{1/2} \quad \forall\ x \in X.$$

Then $\|\cdot\|$ is a norm on X.

To prove the theorem we need the following lemma (Schwarz's inequality)

<u>LEM</u>. Let X be an inner product space. Then $\forall\ x,y \in X$

i). $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$

ii) $|\langle x,y\rangle| = \|x\|\cdot\|y\|$ iff $\alpha x + \beta y = 0$ for some $\alpha,\beta \in \mathbb{F}$ not both zero.

<u>Prf</u>. ($\mathbb{F} = \mathbb{R}$) Consider

$$f(\alpha,\beta) = \|\alpha x + \beta y\|^2 = \langle \alpha x + \beta y, \alpha x + \beta y\rangle$$

$$= \alpha^2\|x\|^2 + 2\alpha\beta\langle x,y\rangle + \beta^2\|y\|^2.$$

(i). $f(\alpha,\beta) \geq 0 \quad \forall\ \alpha,\beta \in \mathbb{R}$ iff discriminant ≤ 0 i.e. $\langle x,y\rangle^2 \leq \|x\|^2\|y\|^2$

This, together with ppty 4 of $\langle\cdot,\cdot\rangle$ proves (i).

(ii) Suppose $\alpha x + \beta y \neq 0$ whenever either α or β are nonzero. Then $f(\alpha,\beta) > 0 \iff$ discriminant $< 0 \Rightarrow$ (ii)

Prf of THM. $\|x\| = \langle x, x\rangle^{1/2}$ SATISFIES THE NORM AXIOMS:

1). $\|x\| \geq 0$, $\|x\| = 0$ iff $x = 0$ (ppty 4 of $\langle\cdot,\cdot\rangle$.

2) $\langle \alpha x, \alpha x\rangle^{1/2} = \|\alpha x\| = (\alpha^2 \langle x, x\rangle)^{1/2} = |\alpha| \cdot \|x\|$

3). $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle$

$$\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\|$$

$$= (\|x\| + \|y\|)^2 \quad \blacksquare$$

DEF An inner product space that is complete in the sense of the norm induced by the inner product, is called a HILBERT space.

Ex. $(\mathbb{R}^n, \|\cdot\|_2)$ is a Hilbert space

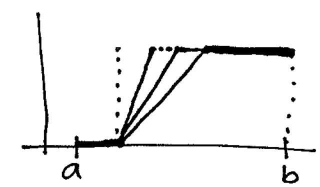
$\|x\|_2 = \left\{ \sum_1^n x_i^2 \right\}^{1/2}$ the Euclidean norm

Ex. Consider $C^n[a,b]$ and define $\langle\cdot,\cdot\rangle_c : C^n[a,b] \times C^n[a,b] \to \mathbb{R}$ as

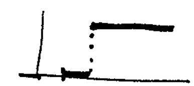
$$\langle x(\cdot), y(\cdot)\rangle_c = \int_a^b \langle x(t), y(t)\rangle_{\mathbb{R}^n}\, dt$$

Then $(C^n[a,b], \langle\cdot,\cdot\rangle_c)$ is an inner product space but not a Hilbert space.

e.g. Consider the sequence of functions:



whose limit



does not belong to $C^n[a,b]$.

The completion of $(C^n[a,b], \langle\cdot,\cdot\rangle_c)$ is a space denoted by $L_2^n[a,b]$, the space of all square-integrable, Lebesgue-measurable functions. Note however that $(C^n[a,b], \|\cdot\|_c)$ with $\|x\|_c = \max_{t\in[a,b]} \|x(t)\|$, is a Banach space.

REM Let $(X, \|\cdot\|)$ be a normed linear space. Then $\|\cdot\| : X \to \mathbb{R}$ is uniformly continuous on X.

Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. Then for each $y \in X$, $x \mapsto \langle x, y\rangle$: $X \to \mathbb{R}$ (or $\mathbb{C}$) is uniformly continuous on X.

INDUCED NORMS

The space $\mathbb{C}^{n\times n}$ $(\mathbb{R}^{n\times n})$ of all $n\times n$ matrices with complex (real) elements is a linear vector space if addition and scalar multiplication are done componentwise. Further, each $A \in \mathbb{C}^{n\times n}$ defines a linear mapping $A : C^n \to C^n$ i.e. $x \in C^n \mapsto Ax$.

Def. Let $\|\cdot\|$ be a given norm on $\mathbb{C}^n$. Then, for each $A \in \mathbb{C}^{n\times n}$ the quantity $\|A\|_i$, defined by:

$$\|A\|_i = \sup_{\substack{x\neq 0 \\ x\in\mathbb{C}^n}} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$$

is called the <u>induced</u> matrix norm of A corresponding to the vector norm $\|\cdot\|$.

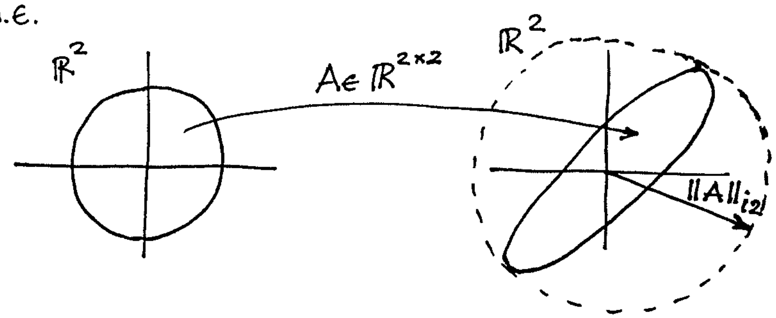
LEM. For each $\|\cdot\|$ on $\mathbb{C}^n$, $\|\cdot\|_i : \mathbb{C}^{n\times n} \to [0,\infty)$ is a norm on $\mathbb{C}^{n\times n}$.

Prf Axioms 1 & 2 by inspection. For 3 let $A, B \in \mathbb{C}^{n\times n}$. Then

$$\|A+B\|_i = \sup_{\|x\|=1} \|(A+B)x\| = \sup_{\|x\|=1} \|Ax+Bx\|$$
$$\leq \sup_{\|x\|=1} \|Ax\| + \|Bx\| \leq \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\|_i + \|B\|_i$$

REM. $\|A\|_i$ can be interpreted as the maximum "gain" of the mapping A.

i.e.



REM To each norm on $\mathbb{C}^n$ there corresponds an induced norm on $\mathbb{C}^{n\times n}$. The converse is not true in general.

LEM Let $\|\cdot\|_i$ be an induced norm on $\mathbb{C}^{n\times n}$. Then $\forall\ A, B \in \mathbb{C}^{n\times n}$

$$\|AB\|_i \le \|A\|_i \|B\|_i$$

Prf. $\|A(Bx)\| \le \|A\|_i \|Bx\| \le \|A\|_i \|B\|_i \cdot \|x\|$ $\forall\ x \in \mathbb{C}^n$.

NORM ON $\mathbb{C}^n$	INDUCED NORM ON $\mathbb{C}^{n\times n}$
$\|x\|_\infty = \max_i \|x_i\|$	$\|A\|_{i\infty} = \max_i \sum_j \|a_{ij}\|$ (row sum)
$\|x\|_1 = \sum_i \|x_i\|$	$\|A\|_{i1} = \max_j \sum_i \|a_{ij}\|$ (column sum)
$\|x\|_2 = \left(\sum_i \|x_i\|^2\right)^{1/2}$	$\|A\|_{i2} = \left[\lambda_{max}(A^*A)\right]^{1/2}$ where $\lambda_{max}(A^*A)$ = maximum eigenvalue of A^*A ; A^* = complex conjugate, transpose

NOTE: $\|A\|_{i2}$ is also known as the maximum singular value of A.

THE CONTRACTION MAPPING THEOREM

(A.K.A. BANACH FIXED POINT THEOREM).

- Very useful to derive existence + uniqueness of solutions to a class of vector ODE's
- Note: Mapping ~ function ~ operator are used interchangeably.

1. GLOBAL CONTRACTIONS

THM Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space and $T: \mathcal{X} \to \mathcal{X}$ a mapping for which there exists a fixed constant $\rho < 1$ s.t.

$$\|Tx - Ty\| \leq \rho \|x - y\| \quad \forall\, x, y \in \mathcal{X}$$

Then:

i). There exists exactly one $x^* \in \mathcal{X}$ s.t

$$Tx^* = x^*.$$

ii). For any $x \in \mathcal{X}$ the sequence $\{x_n\}_1^\infty$ in $\mathcal{X}$ defined by

$$x_{n+1} = Tx_n \quad ; \quad x_0 = x$$

converges to x^*. Moreover,

$$\|x^* - x_n\| \leq \frac{\rho^n}{1-\rho} \|x_1 - x_0\| = \frac{\rho^n}{1-\rho} \|Tx_0 - x_0\|$$

REM: Contraction: The images of any two elements are closer together than the elements are.

PRF (OUTLINE) Let $x \in \mathcal{X}$, arbitrary.

We will show that 1). The sequence $\{x_n\}_1^\infty$ is Cauchy, so it converges in the complete metric space $\mathcal{X}$. 2). The limit x^* is a fixed point of T $(Tx^* = x^*)$ 3). x^* is the unique fixed point of T.

1). $\|x_{n+1} - x_n\| \leq \rho\|x_n - x_{n-1}\| \leq \dots \leq \rho^n \|x_1 - x_0\|$

Let $m = n + r$ $r \geq 0$. Then

$$\|x_m - x_n\| \leq \|x_n - x_{n+1}\| + \|x_{n+1} - x_{n+2}\| + \dots + \|x_{m-1} - x_m\|$$
$$\leq \sum_{i=0}^{r-1} \rho^{n+i} \|x_1 - x_0\|$$
$$= \rho^n \frac{1-\rho^r}{1-\rho} \|x_1 - x_0\| \leq \frac{\rho^n}{1-\rho}\|x_1 - x_0\|$$

Hence, $\|x_m - x_n\|$ can be made arbitrarily small by choosing m, n sufficiently large, i.e., $\forall \varepsilon > 0$ $\exists N(\varepsilon)$:

$\|x_m - x_n\| < \varepsilon$ whenever $m, n > N(\varepsilon)$

$\therefore$ $\{x_n\}_1^\infty$ is Cauchy and since X is Banach $\{x_n\}_1^\infty$ converges in X.

2). Let $x^* = \lim_{n \to \infty} (x_n)$

Then, $\forall x_m$ $\|Tx^* - x^*\| \leq \|x_m - x^*\| + \|x_m - Tx\|$
$\leq \|x_m - x^*\| + \rho\|x_{m-1} - x^*\|$

Hence, $x_n \to x^* \Rightarrow \|x_m - x^*\| < \frac{\varepsilon}{2}$
$\|x_{m-1} - x^*\| < \frac{\varepsilon}{2}$

for any arbitrary $\varepsilon > 0$ and m sufficiently large. ($m \geq M(\varepsilon)$). Since x_m was arbitrary we have that $\|Tx^* - x^*\| < \varepsilon$ for any $\varepsilon > 0 \Rightarrow \|Tx^* - x^*\| = 0 \Rightarrow Tx^* = x^*$.

3). Suppose $\tilde{x}$ is another fixed point of T

Then $\|x^* - \tilde{x}\| = \|Tx^* - T\tilde{x}\| \leq \rho\|x^* - \tilde{x}\|$

Since $\rho < 1$, $\Rightarrow (1-\rho)\|x^* - \tilde{x}\| \leq 0 \Rightarrow x^* = \tilde{x}$.

<u>COMMENTS</u> : Note the repeated use of the triangle inequality $\|x+y\| \leq \|x\| + \|y\|$, in (1) + (2), and the use of $\|x\| = 0 \Leftrightarrow x = 0$ in (2) + (3). A standard technique in this kind of proofs is informally known as the "$\varepsilon/2$ technique":

Suppose we need to show that $\|x-y\|<\varepsilon$

Then, choose an "appropriate" w s.t.

$\|x-w\| < \varepsilon/2$ and $\|y-w\| < \varepsilon/2$. Using the triangle inequality

$$\|x-y\| = \|\underset{\smile}{x}+w-\underset{\smile}{w}-y\| \leq \|x-w\|+\|w-y\| < \tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{2}$$

- In analysis, a usual sufficient condition for the application of the contraction mapping theorem is that $T(\cdot)$ is continuously differentiable and $\|T'(x)\| \leq \rho < 1$.
- The condition $\|Tx-Ty\| \leq \rho\|x-y\|; \rho<1$ CANNOT be replaced by $\|Tx-Ty\|<\|x-y\|$

2. LOCAL CONTRACTIONS

A weaker version of the previous theorem holds in the case where T is a contraction only over some region M of X. (locally).

THM. Let $(X, \|\cdot\|)$ be a Banach space and M be a subset of X. Also let $T: X \to X$ and suppose there exists a constant $\rho < 1$ s.t.

$$\|Tx-Ty\| \leq \rho\|x-y\|, \quad \forall x,y \in M.$$

Further, suppose that there exists $x_0 \in X$ s.t. the ball

$$B = \left\{ x \in X : \|x - x_0\| \leq \frac{\|Tx_0 - x_0\|}{1-\rho} \right\}$$

is entirely contained within M (i.e., $B \subset M$)

Then, (i) T has exactly one fixed point in M, say x^*.

(ii) The sequence $x_{i+1} = Tx_i$, $i \geq 0$ converges to x^*. Further,

$$\|x_n - x^*\| < \frac{\rho^n}{1-\rho}\|Tx_0 - x_0\|$$

REM The Local Contraction mapping thm guarantees that, if <u>all</u> conditions are met, the sequence $\{x_0, Tx_0, T^2x_0 \dots\}$ converges to x^*. However, if y is some other element of M the sequence $\{y, Ty, T^2y, \dots\}$ may or may not converge to x^*.

Also, the theorem states that T has exactly one fixed point in M, without ruling out the possibility that T has some fixed points outside M.

An alternative version of the local contraction mapping theorem is given next. This version assumes a stronger hypothesis than before, but it will be more convenient in later applications

THM Let $(X, \|\cdot\|)$ be a Banach space and B be a closed Ball in X i.e.,

$$B = \{x : \|x - z\| \le r\}$$

for some $z \in X$ and some $r \le \infty$. Let $T : X \to X$ s.t.

(i) $Tx \in B$ whenever $x \in B$.

(ii) There exists a constant $\rho < 1$ s.t.

$$\|Tx - Ty\| \le \rho \|x - y\|, \quad \forall x, y \in B.$$

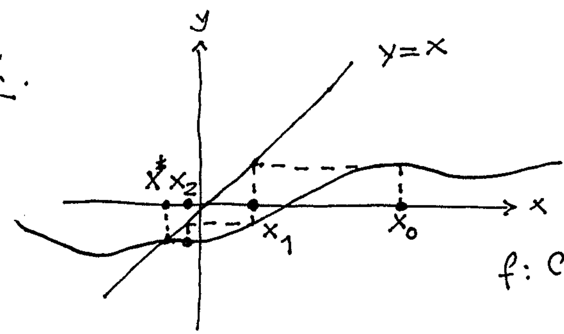
Then,

(i) T has exactly one fixed point in B (say x^*)

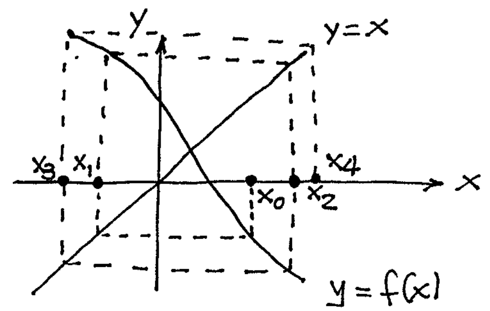
(ii) For <u>any</u> $x_0 \in B$ the sequence $\{x_n\}_1^\infty$ defined by $x_{n+1} = Tx_n$, $n \ge 0$, converges to x^*. Moreover,

$$\|x_n - x^*\| \le \frac{\rho^n}{1-\rho} \|Tx_0 - x_0\|$$

Ex.



f: Contraction



f: NOT CONTRACTION

Ex Approximate Numerical Solutions of $f(x)=0$.

1. Convert $f(x)=0$ to $x=g(x)$.

Suppose g: continuously differentiable on

$$J=[x_0-r,\ x_0+r]$$

for some x_0, r and satisfies

i) $|g'(x)| \le a < 1 \quad \forall x \in J$

ii) $|g(x_0)-x_0| < (1-a)r$

Then, $x=g(x)$ has a unique sol'n on J, the sequence

$$x_{n+1} = g(x_n) \qquad n=0,1,\ldots$$

converges to the solution x (of $x=g(x)$) and one has the error estimates:

$$|x-x_m| < a^m r$$

$$|x-x_m| \le \frac{a}{1-a}\,|x_m - x_{m-1}|$$

Ex. NEWTON'S METHOD

Let f be real valued + twice continuously differentiable on an interval $[a,b]$ and let $\hat{x}$ be a simple zero of f in (a,b). Then, the Newton's method defined by

$$x_{n+1} = g(x_n), \quad g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

is a contraction in some neighborhood of $\hat{x}$ and the iterative sequence $\{x_n\}_1^\infty$ converges to $\hat{x}$ for any x_0 sufficiently close to $\hat{x}$.

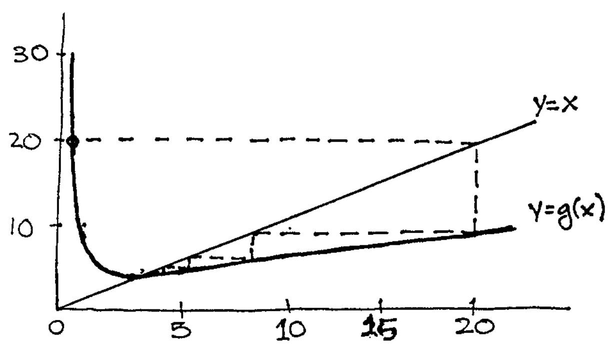
<u>Application</u>. Let c be a given positive number. Construct the iteration

$$x_{n+1} = g(x_n) = \frac{1}{2}\left(x_n + \frac{c}{x_n}\right)$$

$n = 0, 1, \ldots$. Then,

$$x_n \to \sqrt{c}$$

for some x_0. (What are the conditions on x_0 ?).



SOLUTIONS OF ODE'S

1. LOCAL EXISTENCE & UNIQUENESS

<u>THM</u>. Consider the ODE

$$\dot{x} = f(t,x), \quad t \geq 0; \quad x(0) = x_0. \qquad (*)$$

and suppose that f is continuous in t and x and satisfies the following conditions

$$\|f(t,x) - f(t,y)\| \leq \mathbf{K}\|x-y\|, \quad \forall x,y \in B, \; \forall t \in [0,\mathbf{T}]$$
(Lipschitz continuous)

$$\|f(t,x_0)\| \leq \mathbf{h} \quad \forall t \in [0,\mathbf{T}]$$

where B is a ball in $\mathbb{R}^n$ of the form

$$B = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \mathbf{r}\}$$

and K, h, r, T are some finite constants.

<u>Then</u>, $(*)$ has exactly one solution over $[0,\delta]$ whenever

$$h\delta \exp(K\delta) \leq r$$

and

$$\delta \leq \min\left(T, \frac{\rho}{K}, \frac{r}{h+Kr}\right)$$

for some constant $\rho < 1$

<u>Proof</u> (outline) Let $x_0(\cdot)$ denote the function in $C^n[0,\delta]$: $x_0(t) = x_0, \forall t \in [0,\delta]$ and let $S = \{x(\cdot) \in C^n[0,\delta] : \|x(\cdot) - x_0(\cdot)\|_c \le r\}$

Also let $P : C^n[0,\delta] \to C^n[0,\delta]$ defined by

$$(Px)(t) = x_0 + \int_0^t f(\tau, x(\tau))\, d\tau, \quad \forall t \in [0,\delta]$$

Clearly, $x(\cdot)$ is a solution of (*) over $[0,\delta]$ iff $(Px)(t) = x(t)$.

1) P is a contraction on S.

Let $x(\cdot), y(\cdot) \in S$. Then $x(t), y(t) \in B, \forall t \in [0,\delta]$

(Note that S is a set of time functions $S \subset C^n[0,t]$ while $B \subset \mathbb{R}^n$)

Then
$$\|(Px)(t) - (Py)(t)\| \le \int_0^t \|f(\tau, x(\tau)) - f(t, y(\tau))\| d\tau$$
$$\le Kt \|x(\cdot) - y(\cdot)\|_c$$
$$\le \rho \|x(\cdot) - y(\cdot)\|_c$$

Hence $\|(Px)(\cdot) - (Py)(\cdot)\| \le \rho \|x(\cdot) - y(\cdot)\|_c$

2) $P : S \to S$.

$$\|(Px)(t) - x_0\| \le \int_0^t \underbrace{\|f(\tau, x(\tau)) - f(\tau, x_0)\|}_{\le Kr \ (x \in S} d\tau + \int_0^t \underbrace{\|f(\tau, x_0)\|}_{\le h} d\tau$$
$$\le Kr\delta + h\delta \le r$$

$\therefore \|Px(\cdot) - x_0(\cdot)\|_c \le \sup_{t \in [0,\delta]} \|(Px)(t) - x_0\| \le r$

i.e. $(Px)(\cdot) \in S \overset{(1)}{\Rightarrow} P$ has only one fixed point in S

3) P has exactly one fixed point in $C^n[0,\delta]$

Suppose $x(\cdot) \in C^n[0,\delta]$ satisfies

$$x(\cdot) = x_0 + \int_0^t f(\tau, x(\tau))\, d\tau, \quad \forall t \in [0,\delta]$$

Then,
$$\|x(t) - x_0\| \le \int_0^t K \|x(\tau) - x_0\| d\tau + h\delta$$

Using the "Bellman-Gronwall lemma"

$$\|x(t) - x_0\| \le h\delta \exp(Kt) \le h\delta \exp(K\delta) \le r \quad \forall t \in [0,\delta]$$

$\therefore x(\cdot) \in S$. Hence P has exactly one fixed point in $C^n[0,\delta]$, which in fact is in S

$\therefore$ (*) Has exactly one solution over $[0,\delta]$.

COR: If $f(\cdot,\cdot)$ has continuous partial derivatives w.r.t. its second argument and continuous one sided partial derivatives w.r.t. its first argument in some neighborhood of $[0, x_0]$, then $(*)$ has a unique solution over $[0,\delta]$ for sufficiently small δ.

* THE BELLMAN-GRONWALL LEMMA *

Suppose $c \geq 0$, $r(\cdot), k(\cdot) \geq 0$ & continuous

and
$$r(t) \leq c + \int_0^t k(\tau) r(\tau)\, d\tau \quad , \quad \forall t \in [0,T]$$

Then,
$$r(t) \leq c \exp\left[\int_0^t k(\tau)\, d\tau\right], \quad \forall t \in [0,T]$$

This lemma allows the derivation of explicit upper bounds for the solutions of a class of ODE's and is particularly useful in Adaptive Control.

GLOBAL EXISTENCE & UNIQUENESS

THM: Suppose that for each $T \in [0,\infty)$ there exists finite constants k_T, h_T s.t.

1) $\| f(t,x) - f(t,y) \| \leq k_T \| x - y \|$, $\forall x, y \in \mathbb{R}^n$, $\forall t \in [0,T]$

2) $\| f(t, x_0) \| \leq h_T$, $\forall t \in [0,T]$

Then $(*)$ has exactly one solution over $[0,T]$, $\forall T \in [0,\infty)$

The proof can be obtained by applying the local existence & uniqueness thm. on an interval $[0,\delta]$ and then again on $[\delta, 2\delta]$ with initial conditions $x(\delta)$ etc.

An alternative proof can be obtained by showing that the sequence $(P^m x_0)(\cdot)$ is Cauchy in $C^n[0,T]$, and use the fact that $C^n[0,T]$ is a Banach space.

DEPENDENCE ON INITIAL CONDITIONS

THM Let f satisfy the hypotheses of the global existence & Uniqueness thm. Then, for each $z \in \mathbb{R}^n$ and each $T \in [0,\infty)$ there exists exactly one element $z_0 \in \mathbb{R}^n$ s.t. the unique sol'n over $[0,T]$ of the ODE

$$\dot{x} = f(t, x(t)) \; ; \quad x(0) = z_0$$

satisfies $x(T) = z$.

THM Let f as in the previous thm and let $T \in [0,\infty)$ be specified and suppose $x(\cdot), y(\cdot) \in C^n[0,T]$ satisfying

$$\dot{x} = f(t, x(t)) \; ; \quad x(0) = x_0$$
$$\dot{y} = f(t, y(t)) \; ; \quad y(0) = y_0$$

Then for each $\varepsilon > 0$ there exists $\delta(\varepsilon, T) > 0$ s.t.

$$\| x(\cdot) - y(\cdot) \|_C < \varepsilon \quad \text{whenever} \quad \| x_0 - y_0 \| < \delta(\varepsilon, T)$$

EX Consider the linear ODE

$$\dot{x} = A(t)\, x(t) \; ; \quad x(0) = x_0 \qquad (\#)$$

where $A(\cdot)$ is piecewise continuous. Then for every finite T, there exists a finite constant K_T s.t. $\| A(t) \|_i \le K_T, \; \forall t \in [0,T]$

Hence, $\| A(t) x - A(t) y \| \le K_T \| x - y \|$, $\forall\, x, y \in \mathbb{R}^n$; $\forall\, t \in [0,T]$ and

$$\| A(t) x_0 \| \le K_T \| x_0 \| , \quad \forall\, t \in [0,T]$$

Therefore $(\#)$ has a unique sol'n over each finite $[0,T]$ corresponding to each x_0. Moreover, this sol'n depends continuously on x_0.

EX Consider the ODE (scalar)

$$\dot{x} = -x^2 \; ; \quad x(0) = 1$$

Then $-x^2$ is only Locally Lipschitz

$\therefore$ this ODE has a unique sol'n over $[0, \delta]$

for sufficiently small δ. Note, however, that this ODE has a unique sol'n over $[0, \infty)$ namely $x(t) = \frac{1}{t+1}$, even though x^2 is not globally Lipschitz-continuous. (The previous theorems give only sufficient conditions for the existence & uniqueness of solutions.)

On the other hand, a violation of the conditions for existence & uniqueness can serve as an "indicator" that the ODE may not have a solution for some times e.g., $\dot{x} = x^2$; $x(0)=x_0$ whose sol'n $x = \frac{x_0}{1-tx_0}$ is not defined at $t = 1/x_0$.

STABILITY IN THE SENSE OF LYAPUNOV

Consider the ODE

$$\dot{x} = f(t, x), \quad t \geq 0 \qquad (*)$$

$$x \in \mathbb{R}^n$$
$$f : \mathbb{R}^+ \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

and assume that $(*)$ has a unique sol'n over $[0, \infty)$ corresponding to each initial condition $x(0)$ and that this sol'n depends continuously on $x(0)$. Also, let x_e be an equilibrium point of $(*)$ i.e.

$$f(t, x_e) = 0, \quad \forall t \geq t_0$$

Note that w.o.l.og. we can take $x_e = 0$. If this is not the case we can consider the system $\dot{z} = f_1(t,z)$ where $z = x - x_e$ and $f_1(t,z) = f(t, \{z + x_e\})$

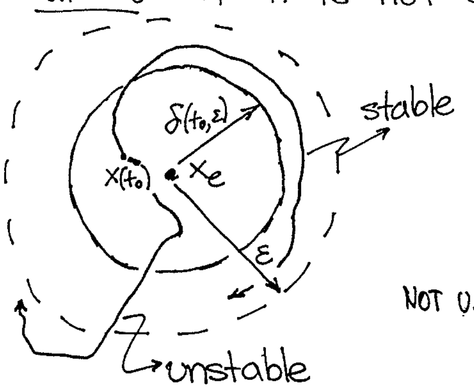
DEF: The equilibrium point x_e at time t_0 of (*) is said to be stable at t_0 if $\forall \varepsilon > 0$ $\exists\ \delta(t_0, \varepsilon) > 0$ s.t.

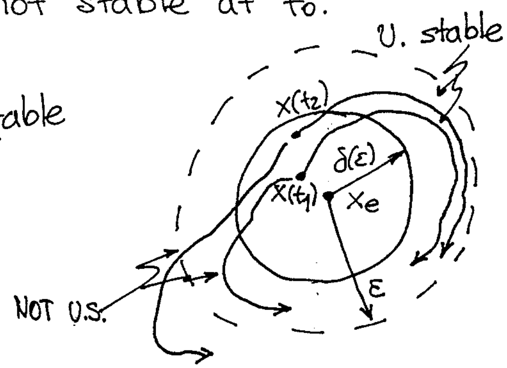
$$\| x(t_0) - x_e \| < \delta(t_0, \varepsilon) \Rightarrow \| x(t) - x_e \| < \varepsilon \quad \forall\ t \geq t_0.$$

Further, it is said to be uniformly stable over $[t_0, \infty)$ if $\forall\ \varepsilon > 0$ $\exists\ \delta(\varepsilon) > 0$ s.t.

$$\| x(t_1) - x_e \| < \delta(\varepsilon),\ t_1 \geq t_0 \Rightarrow \| x(t) - x_e \| < \varepsilon \quad \forall\ t \geq t_1.$$

The equilibrium point is said to be unstable at t_0 if it is not stable at t_0.



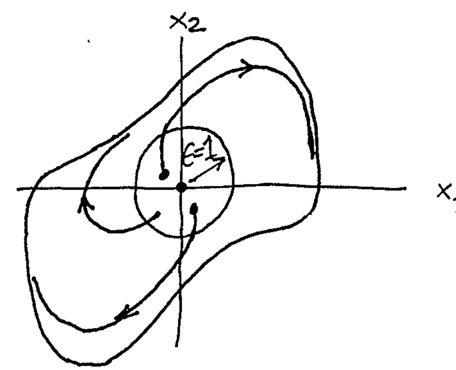


$t_1, t_2 \geq t_0$

EX. Van der Pol oscillator

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

$x_1 = x_2 = 0$ is an equilibrium point. However, solution trajectories starting from any nonzero initial state approach a limit cycle



Note that the sol'n trajectories remain uniformly bounded. The equilibrium (0,0) however is unstable

DEF The equilibrium point x_e is <u>asymptotically stable at time t_0</u> if it is <u>stable</u> at t_0 and there exists a $\delta_1(t_0) > 0$ s.t.

$$\|x(t_0) - x_e\| < \delta_1(t_0) \Longrightarrow \|x(t) - x_e\| \to 0 \text{ as } t \to \infty.$$

Further, it is <u>uniformly asymptotically stable</u> over $[t_0, \infty)$ if it is <u>uniformly stable</u> and $\exists\ \delta_1 > 0$ s.t.

$$\|x(t_1) - x_e\| < \delta_1 ,\ t_1 \geq t_0 \Longrightarrow \|x(t) - x_e\| \to 0 \text{ as } t \to \infty.$$

REM The ball

$$B_{\delta_1(t_0)} = \{ x \in \mathbb{R}^n : \|x - x_e\| < \delta_1(t_0) \}$$

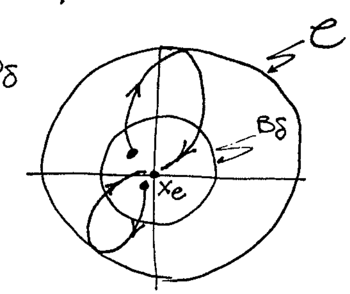
is usually called "ball (or region) of attraction". Notice that a.s. does <u>not</u>

imply that all trajectories starting in $B_{\delta_1(t_0)}$ will be confined in it. It is possible that trajectories start within $B_{\delta_1(t_0)}$ but leave $B_{\delta_1(t_0)}$ at some later time.

A.S. implies that: i) any such trajectories will ultimately return to $B_{\delta_1(t_0)}$ in finite time and $\|x(t) - x_e\| \to 0$

ii) The maximum "excursion" of $x(t)$ can be made arbitrarily small by starting closer to x_e. (see stability definition)

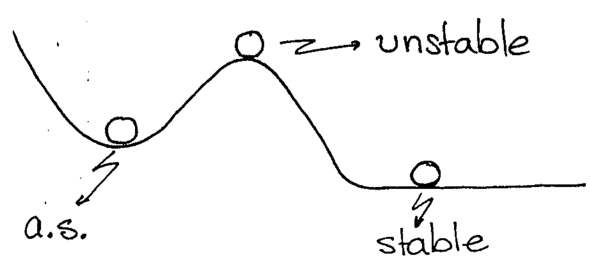
* Note that $\|x(t) - x_e\| \to 0$ alone does not imply a.s. e.g., Consider a system whose trajectories, starting inside B_δ will first touch a curve $\mathcal{C}$ before converging to the x_e



Def The equilibrium point x_e is <u>globally</u> stable / a.s. / u.s. / u.a.s. (or stable/a.s./u.s./u.a.s. in the large) if it is stable/a.s./u.s./u.a.s. regardless of what $x(t_0)$ is.

Rem: A globally asymptotically stable equilibrium is the only equilibrium of the system.

Ex



Ex Distinction between stability and uniform stability: Consider,

$$\dot{x} = \left(6t\sin(t) - 2t\right)x, \quad x(t_0) = x_0$$

whose solution is:

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

Equilibrium $x_e = 0$.

$x=0$ is a stable equilibrium at any time $t_0 \geq 0$ but is <u>not</u> u.s. over $[0,\infty)$.

Prf. Let. $t_0 \geq 0$ be any fixed initial time. Then, consider the ratio $\frac{x(t)}{x(t_0)}$: if $t - t_0 > 6$

$$\left|\frac{x(t)}{x(t_0)}\right| \leq \exp\left[12 + (t-t_0)\left[6 - (t-t_0)\right]\right]$$

and since continuous, it is bounded over $[t_0, t_0+6]$ $\therefore$

$$C(t_0) = \sup_{t \geq t_0}\left|\frac{x(t)}{x(t_0)}\right| < M(t_0)$$

where $M(t_0)$ is a finite number for any fixed t_0. Thus, given $\varepsilon > 0$, Choose $\delta(\varepsilon, t_0) = \varepsilon / C(t_0) \Rightarrow x=0$ is a stable equilibrium for all times t_0.

On the other hand, when $t_0 = 2n\pi$,

$$x[(2n+1)\pi] = x(2n\pi)\exp\{(4n+1)(6-\pi)\pi\}$$

or,

$$c(2n\pi) \geq \exp\{(4n+1)(6-\pi)\pi\}$$

which is unbounded as a function of t_0 (i.e. n). Thus, given any $\varepsilon > 0$ it is not possible to choose $\delta(\varepsilon)$ — independent of t_0 — s.t. $\|x(t_1)\| < \delta(\varepsilon)$, $t_1 \geq t_0 \Rightarrow \|x(t)\| < \varepsilon \quad \forall\, t \geq t_1$

$\therefore$ $x = 0$ is not uniformly stable over $[0, \infty)$

REM

- For autonomous systems

  stability $\Longleftrightarrow$ Uniform stability
  a.s. $\Longleftrightarrow$ U.a.s.

- For non-autonomous systems

  U.S. $\Longrightarrow$ stability
  U.a.s $\Longrightarrow$ a.s.

LEM. Suppose that the equilibrium point x_e at t_0 of $(*)$ is stable at some time $t_1 > t_0$. Then x_e is also a stable equilibrium point at all times $\tau \in [t_0, t_1]$

THM Consider $(*)$ and suppose that f satisfies

$$f(t,x) = f(t+T, x), \quad \forall\, x \in \mathbb{R}^n, \ \forall\, t \geq 0$$

for some positive number T. U.t.c. the following statements are equivalent

(i) The equilibrium x_e of $(*)$ is stable at some $t_0 \geq 0$

(ii) The equilibrium x_e of $(*)$ is U.S. over the interval $[0, \infty)$

Thm. Consider (*) and suppose that f satisfies

$$f(t,x) = f(t+T,x), \quad \forall x \in \mathbb{R}^n, \forall t \geq 0$$

for some positive number T. U.t.c. the following statements are equivalent.

(i) The equilibrium x_e of (*) is a.s. at some time $t_0 \geq 0$

(ii) The equilibrium x_e of (*) is u.a.s. over $[0, \infty)$.

DEF The equilibrium x_e of (*) is exponentially stable if $\exists\, \alpha > 0$ and $\forall \varepsilon > 0 \;\exists\, \delta(\varepsilon) > 0$ s.t.:

$$\| x(t; x_0, t_0) - x_e \| \leq \varepsilon e^{-\alpha(t-t_0)}, \quad \forall t \geq t_0, \quad \forall x_0 : \|x_0 - x_e\| < \delta(\varepsilon)$$

globally exponentially stable if $\exists\, a > 0$ and $\forall \beta > 0 \;\exists\, K(\beta) > 0$ s.t.

$$\|x(t; x_0, t_0) - x_e\| \leq K(\beta) \|x_0 - x_e\| e^{-a(t-t_0)}, \quad \begin{cases} \forall t \geq t_0 \\ \forall \|x_0 - x_e\| < \beta \end{cases}$$

STABILITY OF LINEAR EQUATIONS

$$\dot{x} = A(t)x \qquad (*)$$

Let $\Phi(t,\tau)$ be the state transition matrix (STM) of (*) (i.e. $x(t) = \Phi(t,t_0)x_0$, $\frac{d}{dt}\Phi(t,t_0) = A(t)\Phi(t,t_0)$; $\Phi(t_0,t_0) = I$)

THM : The equilibrium 0 of (*)

1). STABLE AT t_0 iff $\exists\, m(t_0)$ s.t.

$$\|\Phi(t,t_0)\| \leq m(t_0) < \infty \quad \forall t \geq t_0$$

2) U.S. over $[0,\infty)$ iff $\exists\, m_0$ s.t.

$$\|\Phi(t,t_0)\| \leq m_0 \quad \forall t \geq t_0, \forall t_0 \geq 0$$

(or, $\sup_{t_0 \geq 0} m(t_0) = \sup_{t_0 \geq 0} \sup_{t \geq t_0} \|\Phi(t,t_0)\| = m_0 < \infty$)

3) A.S. iff.

$$\|\Phi(t,t_0)\| \to 0 \quad \text{as} \quad t \to \infty$$

Note: $A(t)$: pw cont. $\Rightarrow \Phi(t,t_0)$ cont. $\|\Phi\| \to 0 \Rightarrow \|\Phi\| < C$ $\therefore$ stability.

4). U.A.S. iff $\exists\, K, \alpha > 0$ s.t.

$$\|\Phi(t,\tau)\| \leq K e^{-\alpha(t-\tau)} \quad \forall t_0 \leq \tau \leq t$$

REM • For (*) — linear systems —

U.A.S. $\Leftrightarrow$ Exponential Stability.

• In the special case of linear autonomous systems ($\dot{x} = Ax$);

1) U.S $\Leftrightarrow$ Stability $\Leftrightarrow$ $\mathrm{Re}(\lambda(A)) \leq 0$

$\lambda(A)$: eigenvalues of A — and if $\mathrm{Re}(\lambda_i(A)) = 0$ $\lambda_i(A)$ is a simple zero of the minimal polynomial of A.

2) A.S $\Leftrightarrow$ U.A.S. $\Leftrightarrow$ $\mathrm{Re}(\lambda(A)) < 0$.

• For linear systems

LOCAL STABILITY $\Leftrightarrow$ GLOBAL STABILITY.

LEM For (*)

Equil. 0 is stable at $t_0 \Leftrightarrow$ 0 is stable $\forall t_1 \geq t_0$

(NOT U.S.)

LYAPUNOV THEOREMS

I. DEFINITE & LOCALLY DEFINITE FUNCTIONS

DEF: A continuous function $V: \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ is said to be a locally positive definite function (lpdf) if there exists a continuous nondecreasing function $\alpha: \mathbb{R} \to \mathbb{R}$

s.t. i) $\alpha(0) = 0$, $\alpha(p) > 0 \quad \forall p > 0$

ii) $V(t, 0) = 0 \quad \forall t \geq 0$

iii) $V(t, x) \geq \alpha(\|x\|) \quad \forall t \geq 0$ and

$\forall x \in B_r = \{x : \|x\| \leq r\} \quad r > 0$.

Further, if (iii) holds $\forall x \in \mathbb{R}^n$ then $V(t,x)$ is said to be a

positive definite function (pdf).

(Note: Some authors define pdf's with the additional condition $\alpha(p) \to \infty$ as $p \to \infty$)

DEF: A continuous function $V: \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ is said to be <u>decrescent</u> if there exists a continuous, nondecreasing function $\beta: \mathbb{R} \to \mathbb{R}$ s.t.

(i) $\beta(0) = 0$, $\beta(p) > 0 \ \forall p > 0$

(ii) $V(t,x) \leq \beta(\|x\|) \ \forall t \geq 0, \ \forall x \in B_r$.

<u>Examples</u>:

- $W_1(x_1, x_2) = x_1^2 + x_2^2$ is a pdf and decrescent
- $V_1(t, x_1, x_2) = (t+1)(x_1^2 + x_2^2)$ is a pdf but not decrescent.
- $V_2(t, x_1, x_2) = e^{-t}(x_1^2 + x_2^2)$ is not a pdf. V_2 is decrescent.
- $W_2(x_1, x_2) = x_1^2 + \sin^2 x_2$ is an lpdf though it is not a pdf.

II. Derivative of a function $V(t,x)$ along the trajectories of $\dot{x} = f(t,x)$

Consider the system

$$\dot{x} = f(t,x) \qquad (*)$$

and a function $V: \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ s.t. V is continuously differentiable w.r.t. all its arguments. Also let ∇V denote the gradient of $V(t,x)$ w.r.t. x. Then $\dot{V}: \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\dot{V}(t,x) = \frac{\partial V}{\partial t}(t,x) + \nabla V(t,x) f(t,x)$$

and is called the <u>derivative of V along</u> the trajectories of $(*)$.

III LYAPUNOV's DIRECT METHOD.

- Consider the system

$$\dot{x} = f(t,x), \quad t \geq 0 \qquad (*)$$
$$f(t,0) = 0, \quad t \geq t_0$$

<u>**THM**</u> The equilibrium point 0 at t_0 is stable if there exists a continuously differentiable l.pdf V s.t:

$$\dot{V}(t,x) \leq 0 \quad \forall t \geq t_0, \forall x \in B_r$$

for some ball $B_r \subseteq \mathbb{R}^n$

Further, if V is also decrescent, 0 is u.s. over $[t_0, \infty)$.

<u>REMARKS</u>: This is the basic stability theorem of Lyapunov's direct method. It has a natural interpretation in terms of the "total energy" stored in the system. That is, V can be

thought of as an appropriate energy function which is 0 at the origin (equilibrium point) and positive everywhere else. Under the assumptions of the thm., V does not increase with time, hence the energy level of the system never increases beyond its initial value.

It is important to note that:

1). Only the local behavior of V around the equilibrium is considered ($\rightsquigarrow$ local stability)

2). Not any V, continuously differentiable lpdf will do. For this reason, a test function V (cont. diff. lpdf) is usually termed as a "Lyapunov function candidate". Only after $\dot{V}$ has been shown to satisfy the conditions of the thm., V can be called a "Lyapunov functi

3). The thm. gives a sufficient condition for the stability of the equilibrium point 0 of (*). If such a V can be found, we can conclude stability. (One can actually prove the converse theorem i.e. that if 0 is stable there exists a Lyapunov function, but the result is mostly of theoretical value).

EXAMPLES

Consider the system

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -f(x_2) - g(x_1)$$

where i) f, g continuous

ii) $\forall \sigma \in [-\sigma_0, \sigma_0]$ and some σ_0

$$\sigma f(\sigma) \geq 0$$
$$\sigma g(\sigma) > 0 \quad (\sigma \neq 0).$$

This example describes a typical mass-and-spring system with nonlinear, in general, characteristics. ($f(\cdot)$ represents the friction and $g(\cdot)$ represents the restoring force of the spring). Note that if we select $f(\cdot) \equiv 0$, $g(\sigma) = \sin(\sigma)$ this example is the classical description of an unforced, frictionless pendulum.

The energy stored in the system is the sum of kinetic + potential energy i.e. let

$$V(x_1, x_2) = \frac{1}{2} x_2^2 + \int_0^{x_1} g(\sigma)\, d\sigma.$$

which is a cont. diff. lpdf.

Then,

$$\dot{V}(x_1, x_2) = x_2 \dot{x}_2 + g(x_1)\dot{x}_1$$
$$= x_2[-f(x_2) - g(x_1)] + g(x_1) x_2$$

And finally $\mathring{V} = -x_2 f(x_2)$

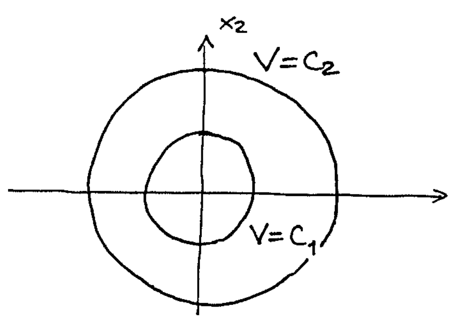
$$\mathring{V}(x_1, x_2) \leq 0 \text{ whenever } |x_2| \leq \sigma_0$$

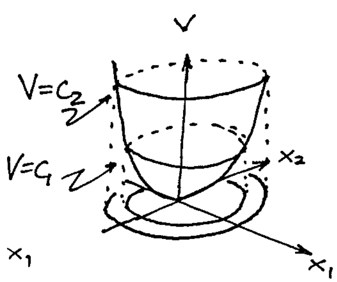
Hence, by the previous thm, 0 is a uniformly stable equilibrium.

- <u>GEOMETRICAL INTERPRETATION OF LYAPUNOV'S THEOREM</u>

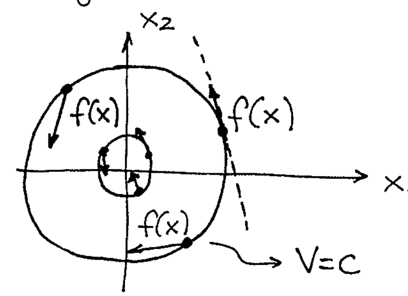
Let $x \in \mathbb{R}^2$, $V : \mathbb{R}^2 \longrightarrow \mathbb{R}$

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





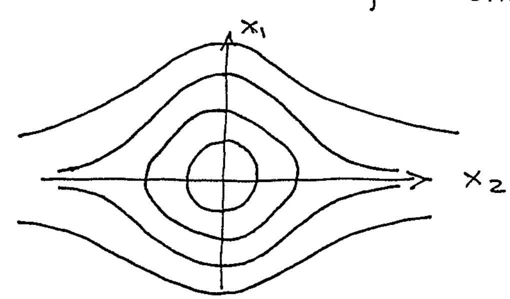
Then $\mathring{V} \leq 0$ (V non-increasing) means that at the boundary of each "surface" $V=C$ the vector field points towards the interior or is tangential to the surface $\forall c \in [0, c_0)$



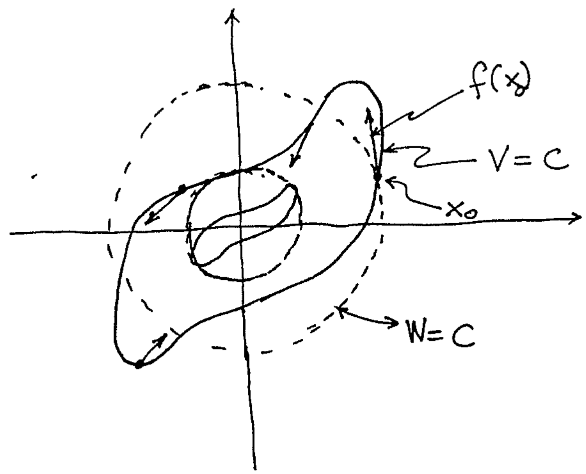
However, depending on the vector field f, the appropriate function V may look quite "strange"

e.g. $V = x_1^2 + \frac{x_2^2}{1+x_2^2} = C$

which is a closed surface only for $C \leq 1$



The meaning and the importance of selecting an appropriate V can be visualized as follows.



At x_0, the vector field points towards the interior of $V=C$ but towards the exterior of $W=C$

Note that for LTI systems ($\dot{x}=Ax$) the "appropriate" V functions are in general ellipsoids i.e. $V = x^T P x$; P : positive definite matrix, yielding a very general and quite elegant stability theory using Lyapunov functions

IV MORE LYAPUNOV THEOREMS

1) DEFINITIONS ...

... to simplify the subsequent statements.

(Note however that there are slight variations among authors)

- **Class K, KR functions** Let

$\varphi : [0,r] \to \mathbb{R}^+$ (or $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$) be a continuous function s.t.

1). $\varphi(0) = 0$; $\varphi(p) > 0$ whenever $p > 0$

2) $\varphi(\cdot)$ is **non-decreasing** on $[0,r]$ (or on $\mathbb{R}^+$)

Then $\varphi(\cdot)$ is said to belong to class K.

If in addition $\varphi(\cdot)$ satisfies

3) $\lim_{p \to \infty} \varphi(p) = \infty$ (radially unbounded)

then $\varphi(\cdot)$ is said to belong to class KR

Let $V(t,x) : \mathbb{R}^+ \times \mathbb{R}^n \longrightarrow \mathbb{R}$ be a cont. diff. function s.t. $V(t,0) = 0$ $\forall\, t \in \mathbb{R}^+$. Then V is said to be :

- <u>LOCALLY PDF</u> if there exists $\varphi \in K$ s.t. $V(t,x) \geq \varphi(\|x\|)$, $\forall\, t \in \mathbb{R}^+$ $\forall\, x \in B_r$ for some $r > 0$.

  (Recall $B_r = \{x \in \mathbb{R}^n : \|x\| \leq r\}$)

- <u>(GLOBALLY) PDF</u> if $\exists\, \varphi \in K$ s.t.

  $$V(t,x) \geq \varphi(\|x\|), \quad \forall\, t \in \mathbb{R}^+; \forall\, x \in \mathbb{R}^n$$

- <u>(LOCALLY) NEGATIVE DEFINITE</u> if $-V$ is (L)pdf.

- <u>LOCALLY positive semidefinite</u> (Lpsdf) if $V(t,x) \geq 0$ $\forall\, t \in \mathbb{R}^+$, $\forall\, x \in B_r$ and for some r.

- <u>RADIALLY UNBOUNDED</u> if $\exists\, \varphi \in KR$ s.t. $V(t,x) \geq \varphi(\|x\|)$ $\forall\, t \in \mathbb{R}^+$, $\forall\, x \in \mathbb{R}^n$

- <u>DECRESCENT</u> if $\exists\, \varphi \in K$ s.t.

  $$|V(t,x)| \leq \varphi(\|x\|)$$

  $\forall\, t \in \mathbb{R}^+$; $\forall\, x \in B_r$, for some $r > 0$.

<u>Note</u>: Local properties $\leftrightarrow$ $B_r \subset \mathbb{R}^n$

Global " $\leftrightarrow$ $B_r = \mathbb{R}^n$.

Also, when appropriate, $\mathbb{R}^+$ may be substituted by $[t_0, \infty)$.

Next, let us consider the ODE

(*) $\dot{x} = f(t,x)$, $t \geq t_0$; $x(t_0) = x_0$

where $f(t,0) = 0$ $\forall\, t \in [t_0, \infty)$ and f is sufficiently smooth s.t. (*) possesses exactly one sol'n $\forall\, t \in [t_0, \infty)$

$\forall x_0 \in B_r$. The following theorems assess the stability properties of the equilibrium $x_e = 0$ of $(*)$.

<u>THM</u> • if $\exists\ V(t,x)$: Lpdf with $\dot{V}$: Lnsdf then the equilibrium $x_e = 0$ of $(*)$ is <u>stable.</u>

• if $\exists\ V(t,x)$: lpdf, decrescent with $\dot{V}$: Lnsdf then the equilibrium $x_e = 0$ of $(*)$ is <u>Uniformly Stable</u>

• if $\exists\ V(t,x)$: lpdf, decrescent with $\dot{V}$: Lndf then the equilibrium $x_e = 0$ of $(*)$ is <u>Uniformly asymptotically stable.</u>

• If $\exists\ \varphi_1, \varphi_2, \varphi_3 \in K$ and $V(t,x)$ s.t.

$$\varphi_2(\|x\|) \le V(t,x) \le \varphi_1(\|x\|)$$
$$\dot{V}(t,x) \le -\varphi_3(\|x\|)$$

$\forall\ t \in [t_0, \infty)$, $\forall\ x \in B_r$ and some $r > 0$ and $\exists$ constants $c_1 - c_4 > 0$ s.t.

$$c_1 \varphi_1(\|x\|) \le \varphi_2(\|x\|) \le c_2 \varphi_1(\|x\|)$$
$$c_3 \varphi_1(\|x\|) \le \varphi_3(\|x\|) \le c_4\ \varphi_1(\|x\|)$$

(i.e. $\varphi_1, \varphi_2, \varphi_3$ are of the same order of magnitude) $\forall\ x \in B_r$ then the equilibrium $x_e = 0$ is <u>exponentially stable.</u>

• if $\exists\ V(t,x)$: pdf with $\dot{V}$: nsdf then 0 is <u>globally stable</u>

• if $\exists\ V(t,x)$: pdf, decrescent $(B_r = \mathbb{R}^n)$ with $\dot{V}$: nsdf then 0 is <u>globally U.S.</u>

- If $\exists\ V(t,x)$: pdf, decrescent, radially unbounded with $\dot{V}$: ndf $(B_r = \mathbb{R}^n)$ then the equilibrium $x_e = 0$ of $(*)$ is <u>globally uniformly asymptotically stable</u>.

- If $\exists\ \varphi_1, \varphi_2, \varphi_3 \in KR$ and $V(t,x)$ s.t.

$$\varphi_2(\|x\|) \leq V(t,x) \leq \varphi_1(\|x\|)$$

$$\dot{V}(t,x) \leq -\varphi_3(\|x\|)$$

$\forall\ t \in [t_0, \infty)$ $\forall\ x \in \mathbb{R}^n$ and there exist positive constants $c_1 - c_4$ s.t.

$$c_1\varphi_1(\|x\|) \leq \varphi_2(\|x\|) \leq c_2\varphi_1(\|x\|)$$

$$c_3\varphi_1(\|x\|) \leq \varphi_3(\|x\|) \leq c_4\varphi_1(\|x\|)$$

$\forall\ x \in \mathbb{R}^n$ then the equilibrium $x_e = 0$ of $(*)$ is <u>globally exponentially stable</u>

- If $\exists\ V(t,x)$: pdf, radially unbounded and $\dot{V}(t,x) \leq c\,V(x,t)$ $\forall\ x,t$ and some constant $c > 0$ then $(*)$ has <u>no finite escape time.</u>

- If $\exists\ V(t,x)$, $t \in \mathbb{R}^+$, $\|x\| \geq r > 0$ and $\exists\ \psi_1, \psi_2 \in KR$ s.t.

$$\psi_1(\|x\|) \leq V(t,x) \leq \psi_2(\|x\|)$$

$$\dot{V}(t,x) \leq 0$$

$\forall\ \|x\| \geq r$, $\forall\ t \geq 0$ then the solutions of $(*)$ are <u>uniformly bounded.</u> i.e.
$\forall\ a > 0$ and $\forall\ t_0 \in \mathbb{R}^+$, $\exists\ \beta(a) > 0$ s.t.
if $\|x_0\| < a$ then $\|x(t; x_0, t_0)\| < \beta$
$\forall\ t \geq t_0$.

If in addition $\exists\ \psi_3 \in K$ s.t.

$$\overset{\circ}{V}(t,x) \le -\psi_3(\|x\|)$$

$\forall\ \|x\| \ge r$ and $\forall t \ge 0$ then the solutions of (*) are <u>uniformly ultimately bounded</u>

i.e. $\exists\ B>0$ such that $\forall\ a>0,\ \forall t_0 \in \mathbb{R}^+$ $\exists\ T(a)>0$ s.t. $\|x_0\| < a \Rightarrow$

$\|x(t; x_0, t_0)\| < B \quad \forall\ t \ge t_0 + T$.

<u>DEF</u> A set $M \subset \mathbb{R}^n$ is said to be an <u>invariant set</u> of (*) if whenever $y \in M$ and $t_0 \ge 0$, every solution of (*) starting from an initial point in M stays within M at all future times

i.e. $\quad x(t; y, t_0) \in M, \ \forall t \ge t_0$.

● <u>THM</u> Suppose (*) is autonomous and there exists a radially unbounded pdf $V(x)$ s.t.

$$\overset{\circ}{V}(x) \le 0 \quad \forall\ x \in \mathbb{R}^n$$

and the origin $x=0$ is the only invariant subset of the set $E = \{x \in \mathbb{R}^n : \overset{\circ}{V}(x) = 0\}$ then the equilibrium $x_e = 0$ is <u>globally asymptotically stable</u>.

● <u>THM</u> Suppose (*) is autonomous and let $V : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and suppose that for some $c>0$ the set

$$\Omega_c = \{x \in \mathbb{R}^n : V(x) \le c\}$$

is bounded and V is bounded below

over the set Ω_c and that $\dot{V}(x) \le 0$ $\forall\, x \in \Omega_c$. Let E denote the set

$$E = \{ x \in \Omega_c : \dot{V}(x) = 0 \}$$

and let M be the largest invariant set of $(*)$ contained in E. Then whenever $x_0 \in \Omega_c$ the solution $x(t; x_0, 0)$ of $(*)$ approaches M as $t \to \infty$.

THM: Suppose $(*)$ is autonomous and $\exists$ $V(x)$: lpdf over some ball B_r s.t. $\dot{V}(x) \le 0$ $\forall\, x \in B_r$. Also let $m = \sup_{\|x\| \le r} V(x)$ and define

$$S = \{ x \in \mathbb{R}^n : V(x) \le m,\ \dot{V}(x) = 0 \}$$

Suppose S contains no trajectories of $(*)$ other than the trivial one $x \equiv 0$.

Then the equilibrium 0 is asymptotically stable.

REM: S may contain points outside B_r.

- **THM** (LA SALLE) Suppose $(*)$ is periodic i.e. $f(t,x) = f(t+T, x)$, $\forall\, t$; $\forall\, x \in \mathbb{R}^n$ for some $T > 0$.

Suppose that $V(t,x)$ is pdf, radially unbounded with $V(t,x) = V(t+T, x)$ and $\dot{V}(t,x) \le 0$, $\forall\, x \in \mathbb{R}^n$, $\forall\, t \ge 0$

Define

$$S = \{ x \in \mathbb{R}^n : \dot{V}(t,x) = 0,\ \forall\, t \ge 0 \}$$

and suppose that S contains no trajectories of $(*)$ other than $x \equiv 0$. Then $x_e = 0$ is a globally asymptotically stable equilibrium.

EXAMPLES

1). One of the main applications of Lyapunov theory is to obtain stability conditions involving the design parameters of the system under study. E.g. Consider the system :

$$\begin{aligned} \mathring{x}_1 &= x_2 \\ \mathring{x}_2 &= -p(t)\,x_2 - e^{-t}x_1 \end{aligned} \qquad (\#)$$

The objective is to find conditions on $p(t)$ s.t. 0 is a stable equilibrium of $(\#)$ at $t=0$. Let $V(t,x_1,x_2) = x_1^2 + e^t x_2^2$

Note that $V(t,x_1,x_2) \geq g(\|x\|) \triangleq x_1^2+x_2^2$

Then $\mathring{V} = e^t x_2^2 + 2x_1x_2 + 2e^t x_2[-p(t)\,x_2 - e^{-t}x_1]$

$= e^t x_2^2\,[-2p(t)+1]$

$\therefore\ \mathring{V} \leq 0$ provided that $p(t) \geq \frac{1}{2}, \forall t \geq 0$

Thus, 0 is a stable equilibrium at $t_0=0$ for $p(t) \geq \frac{1}{2}$, $\forall t \geq 0$.

(Note that we have not U.S. since V is not decrescent).

It should be emphasized that using a different V, we may obtain entirely different stability conditions involving $p(\cdot)$.

2). Let

$$\begin{aligned} \mathring{x}_1 &= x_1(x_1^2+x_2^2-1) - x_2 \\ \mathring{x}_2 &= x_1 + x_2(x_1^2+x_2^2-1) \end{aligned} \qquad (\#)$$

and consider the pdf

$$V = x_1^2+x_2^2.$$

Then, $\mathring{V} = 2(x_1^2+x_2^2)(x_1^2+x_2^2-1)$ which is an Lndf over $B_1 = \{x\in\mathbb{R}^2 : \|x\|\leq 1\}$

Hence, 0 is UAS (at least locally).

3). Let

$$\mathring{x}_1 = -x_2$$
$$\mathring{x}_2 = -f(x_2) - g(x_1) \qquad (\#)$$

where f, g are continuous

$f(0) = g(0) = 0$

$\sigma f(\sigma) > 0, \ \sigma g(\sigma) > 0, \ \forall \sigma \neq 0$

$\int_0^{\sigma} g(\xi)\,d\xi \to \infty$ as $|\sigma| \to \infty$

Consider

$$V(x) = \frac{x_2^2}{2} + \int_0^{x_1} g(\xi)\,d\xi$$

which is a cont. diff. pdf and radially unbounded. Then

$$\mathring{V} = -x_2 f(x_2) \leq 0, \quad \forall x \in \mathbb{R}^2.$$

<u>Note</u>: $\mathring{V}$ is <u>nsd</u> since for $x = \binom{x_1}{0}$ $\|x\| = |x_1| > 0$, $\mathring{V} = 0$ ($\mathring{V} \nless 0$ for some $\|x\| \neq 0$).

Further, let $S = \{x \in \mathbb{R}^2 : \mathring{V}(x_1, x_2) = 0\}$

i.e. $S = \{x \in \mathbb{R}^2 : x_2 = 0\}$.

If S contains any trajectories of (#) it must have $x_2(t) \equiv 0$ hence $x_1 = \text{const.}$ – say $x_1 = x_{10}$ – and $\mathring{x}_2 \equiv 0$. Hence,

$f(x_2) + g(x_{10}) = 0$, i.e. $g(x_{10}) = 0$

and therefore, using the assumptions on $g(\cdot)$, $x_{10} = 0$. Thus, the only trajectory that lies entirely within S is the trivial trajectory $x_1 \equiv x_2 \equiv 0$. Hence, 0 is a globally asymptotically stable equilibrium point.

4). Let

$$\mathring{\epsilon} = -a_m \epsilon + \phi u \quad ; \quad a_m > 0, \ |u| < a$$

$$\mathring{\phi} = -\gamma \epsilon u \quad ; \quad \gamma > 0$$

and choose

$$V = \frac{\epsilon^2}{2} + \frac{\phi^2}{2\gamma} \quad (\text{pdf})$$

$$\therefore \ \mathring{V} = -a_m \epsilon^2 + \epsilon\phi u - \phi\epsilon u = -a_m \epsilon^2$$

$$\therefore \ \mathring{V} = -a_m \epsilon^2 \leq 0 \quad (\text{nsdf})$$

$\therefore$ (ϵ, ϕ) U.B , V : U.B.
(0,0 U.S.)

Then $\int_0^T \dot{V}\, dt = V(T) - V(0)$

$$\therefore \ \int_0^T \epsilon^2 = \frac{V(0) - V(T)}{a_m} \leq K[\epsilon(0), \phi(0)]$$

where K is a constant which may depend on I.C.

Hence ϵ is square integrable $\left(\int_0^\infty \epsilon^2 < \infty.\right)$

Further, $\frac{d}{dt}(\epsilon^2) = 2\epsilon\mathring{\epsilon} = -2a_m\epsilon^2 + 2\phi u \epsilon$

and $\left|\frac{d}{dt}(\epsilon^2)\right| \leq C_1$ since u, ϕ, ϵ are U.B.

Hence (see HW-1) $\epsilon \to 0$ as $t \to \infty$.

Furthermore V is bounded from below and is non-increasing $(\dot{V} \leq 0)$ $\therefore$ $\lim_{t\to\infty} V(t)$ exists, say $V(t) \underset{t\to\infty}{\longrightarrow} V_\infty$. Since $\epsilon \to 0 \Rightarrow \phi^2 \to$ const.

Alternatively, let $\epsilon(0), \phi(0)$ be the initial conditions and $\Omega_c = \{x \in \mathbb{R}^n : V(x) \in c\}$.

Then we can always find c: $x(0) = \begin{pmatrix}\epsilon(0)\\ \phi(0)\end{pmatrix} \in \Omega_c$ (note : $V(x)$ is radially unbounded).

and V will be bounded below by 0,

Further $\dot{V} \leq 0 \ \forall x \in \Omega_c$. Let

$$E = \{x \in \Omega_c : \dot{V}(x) = 0\}$$

$$= \{\epsilon, \phi : \epsilon = 0, |\phi|^2 \leq 2\gamma c\}$$

The largest invariant subset of E will have $\epsilon \equiv 0 \Rightarrow$ 1) $\mathring{\phi} = 0$
2) $\dot{\epsilon} = 0$

Hence, $\phi = \text{Const.}$ and $\phi u = 0$;

in other words

$$M = \left\{ \epsilon, \phi : \begin{bmatrix} |\phi|^2 \leq 2\gamma c \text{ , constant} \\ \phi u = 0 \\ e = 0 \end{bmatrix} \right\}$$

and for $x_0 \in \Omega_c$ (by construction)

$(\epsilon, \phi) \to M$ as $t \to \infty$ i.e.

$$\begin{aligned} &\lim |\epsilon| = 0 \\ &\lim |\phi| = \text{constant} \quad \leq \sqrt{2\gamma c} \\ &\lim |\phi u| = 0 \end{aligned}$$

And if $|u| \geq \varepsilon > 0$ for all $t \geq t_0$ and some $t_0 \geq 0$

we get that $\lim \phi = 0$.

V LYAPUNOV THEORY & LINEAR SYSTEMS

Consider a Linear Time Invariant (LTI) system

$$\dot{x} = Ax \quad ; \quad x(0) = x_0. \qquad (*)$$

The stability of (*) (i.e. of the equilibrium 0 of (*)) can be determined by studying the eigenvalues of A.

On the other hand, using a Lyapunov approach let

$$V(x) = x^T P x$$

where P is a symmetric positive definite matrix i.e.:

$$P \in \mathbb{R}^{n\times n}, \quad P = P^T, \quad x^T P x \geq \alpha \|x\|^2 \;\; \forall x \in \mathbb{R}^n, \;\; \alpha > 0.$$

- <u>Conditions for P.D.</u>

Let $P \in \mathbb{R}^{n\times n}$, $P = P^T$. Then the following statements are equivalent:

1) $\lambda_i(P) > 0 \quad i = 1, 2 \ldots, n$
2) $\exists$ non singular A_1 : $P = A_1^T A$.
3) Every principal minor of P is positive
4) $\exists\ \alpha > 0$: $x^T P x \geq \alpha \|x\|^2$, $\forall x \in \mathbb{R}^n$.

<u>Note:</u> A symmetric matrix P has n orthogonal eigenvectors and n real eigenvalues and can be decomposed as

$$P = U^T \Lambda U$$

where U is unitary orthogonal $(U^T U = I)$
Λ is diagonal.

Hence

$$\lambda_{min}(P)\|x\|^2 \leq x^T P x \leq \lambda_{max}(P)\|x\|^2.$$

If $\lambda_{min}(P) = 0$ then P is said to be positive semi-definite $(x^T P x \geq 0 \;\forall x \in \mathbb{R}^n)$

For $P \geq 0$ $\quad \|P\|_{i2} = \lambda_{MAX}(P)$

For $P > 0$ $\quad \|P^{-1}\|_{i2} = \dfrac{1}{\lambda_{MIN}(P)}$

Thus, taking the derivative of V along the trajectories of $(*)$ we get

$$\dot{V}(x) = \dot{x}^T P x + x^T P \dot{x} = x^T (A^T P + PA) x$$

Let

$$A^T P + PA = -Q.$$

Then if Q is PD, $\dot{V}$ is ndf $\Rightarrow$

$(*)$ is (G) A.S.

THM: Let $A \in \mathbb{R}^{n\times n}$ and $\{\lambda_i\}_1^n$ be the eigenvalues of A. Then, the equation

$$A^T P + PA = -Q$$

has a unique sol'n for P corresponding to every $Q \in \mathbb{R}^{n\times n}$ iff $\lambda_i + \bar{\lambda}_j \neq 0 \;\; \forall\, i, j$

THM: Given $A \in \mathbb{R}^{n\times n}$ the following statements are equivalent:

(i) $\mathrm{Re}[\lambda_i(A)] < 0 \quad \forall i$

(ii) There exists <u>some</u> $Q \in \mathbb{R}^{n\times n}$, positive definite s.t. $A^T P + PA = -Q$ has a unique sol'n for P and this sol'n is positive definite

(iii) $\forall$ p.d. $Q \in \mathbb{R}^{n\times n}$ $\exists\, P \in \mathbb{R}^{n\times n}$ s.t.

$$A^T P + PA = -Q$$

and this P is p.d.

LEM: Consider the "Lyapunov Equation"

$$(\#) \qquad A^T P + PA = -M \;;\quad M = M^T \in \mathbb{R}^{n\times n}$$

and suppose $\mathrm{Re}[\lambda_i(A)] < 0$. Then

$$P = \int_0^\infty e^{A^T t} M e^{At}\, dt$$

is the unique sol'n of $(\#)$.

INSTABILITY THEOREMS

Consider the ODE

$$\dot{x} = f(t,x), \quad t \geq 0 \qquad (*)$$

with $f(t,0) = 0 \quad \forall\, t \geq t_0$.

<u>THM</u>: The equilibrium point 0 at t_0 of $(*)$ is unstable if there exists a continuously differentiable decrescent function $V: \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$

s.t. (i) $\dot{V}$ is lpdf

(ii) $V(t,0) = 0$ and there exist points x arbitrarily close to 0 s.t. $V(t_0, x) > 0$.

(V is <u>not</u> required to be lpdf. However, in the rare case both V and $\dot{V}$ are lpdf's the equilibrium is called <u>completely unstable</u> i.e. $\exists\ \varepsilon > 0$ s.t. <u>Every</u> trajectory $x(\cdot)$, other than $x \equiv 0$, satisfies $\|x(t)\| \geq \varepsilon$ for some t)

(I)

<u>THM</u> The equil. 0 of $(*)$ at t_0 is unstable if $\exists\ V: \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, cont. diff, decrescent and s.t. (i) $V(t_0, 0) = 0$ and $V(t_0, x)$ assumes positive values arbitrarily close to the origin, (ii) $\dot{V}(t,x)$ is of the form

$$\dot{V}(t,x) = \lambda V(t,x) + V_1(t,x)$$

where $\lambda > 0$ is a constant and $V_1: \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is s.t. $V_1(t,x) \geq 0 \quad \forall\, t \geq t_0, \forall\, x \in B_r$ for some Ball $B_r \subset \mathbb{R}^n$

<u>THM</u> (Cetaev) The equilibrium 0 at t_0 of $(*)$ is unstable if the following conditions hold: $\exists\ V: \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, cont. diff. and a closed set Ω containing 0 in its interior s.t.

(II)

1) $\exists$ open set $\Omega_1 \subset \Omega$ containing 0 on its boundary

2) $V(t,x) > 0$; $\forall\, t \geq t_0, \forall\, x \in \Omega_1$

$V(t,x) = 0$; $\forall\, t \geq t_0, \forall\, x \in \partial\Omega_1$ (the boundary of Ω_1 in Ω)

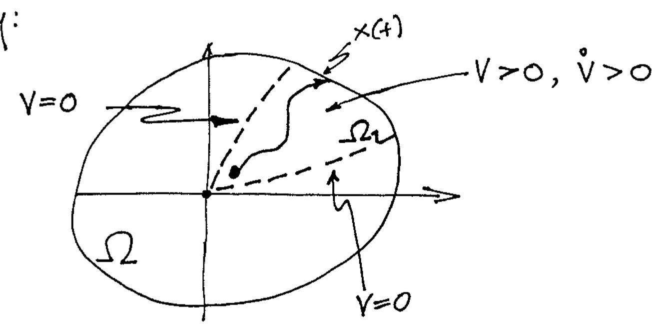
3) $V(t,x)$ is bounded above in Ω, uniformly in t

4) $\dot{V}(t,x) \geq \gamma(\|x\|)$ $\quad \forall\, t \geq t_0, \forall\, x \in \Omega_1$

where γ is a class K function

(Note $\dot{V}$ is not lpdf : 4 is required to hold in Ω_1)

Pictorially:



III

L_p SPACES & I/O STABILITY

1). A Subset S of $\mathbb{R}$ is said to be of <u>measure zero</u> if S contains either finite or countably infinite number of elements.

i.e. $S = \{s_i\}$, $i = 1, 2, \ldots$

the elements of S can be placed in 1-1 Correspondence with a subset of $\mathbb{N}$.

2) A function $f(\cdot) : \mathbb{R} \to \mathbb{R}$ is said to be measurable if it is continuous everywhere except on a set of measure zero.

DEF For all $p \in [1, \infty)$ we label as $L_p[0,\infty)$ (or simply L_p) the set of all measurable functions $f(\cdot) : [0,\infty) \to \mathbb{R}$, s.t

$$\int_0^\infty |f(t)|^p dt < \infty$$

The Label $L_\infty[0,\infty)$ denotes the set of all measurable functions $f(\cdot):[0,\infty)\to\mathbb{R}$

s.t. $\quad \underset{t\in[0,\infty)}{\text{ess. sup.}} |f(t)| < \infty$.

i.e. L_∞ is the set of all essentially bounded functions $\mathbb{R}^+\to\mathbb{R}$ (bounded except on a set of measure zero).

(i) $\forall\ p\in[1,\infty]$ L_p is a linear vector space

(ii). $\forall\ p\in[1,\infty]$, $(L_p, \|\cdot\|_p)$ is a Banach space where

$$\|f\|_p = \left[\int_0^\infty |f(t)|^p dt\right]^{1/p} \qquad p\in[1,\infty)$$

$$\|f\|_\infty = \underset{t\in[0,\infty)}{\text{ess. sup}} |f(t)|$$

(iii) for $p=2$, $(L_2, \langle\cdot,\cdot\rangle_2)$ is a Hilbert space with $\langle f,g\rangle_2 = \int_0^\infty f(t)g(t)\,dt$.

(iv) for $p\in[1,\infty]$ and $f,g\in L_p$

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

(Minkowski's inequality)

(v) for $p,q\in[1,\infty]$ s.t. $\frac{1}{p}+\frac{1}{q}=1$

let $f\in L_p$ and $g\in L_q$

then $h(t) \triangleq f(t)g(t) \in L_1$ and

$$\|h\|_1 \le \|f\|_p\|g\|_q$$

(Holder's Inequality)

DEF: Let $x(\cdot):\mathbb{R}^+\to\mathbb{R}$, measurable.

Then $\forall\ T\in\mathbb{R}^+$ the function $x_T(\cdot)$:

$$x_T(t) = \begin{cases} x(t) & 0\le t\le T \\ 0 & t>T \end{cases}$$

is called the truncation of $x(\cdot)$ to the interval $[0,T]$

Def The set of all measurable functions $f(\cdot) : [0,\infty) \to \mathbb{R}$ s.t. $f_T(\cdot) \in L_p, \forall T$ is called the <u>extended L_p</u> space and denoted by L_{pe}.

e.g. $f(t) = t \in L_{pe}, \ \forall p \in [1,\infty]$

$f(t) = t \notin L_p, \ \forall p \in [1,\infty]$

- $L_p \subset L_{pe}$; L_{pe} is a linear vector space ; L_{pe} is <u>not</u> a normed space.

Lem: For each $p \in [1,\infty]$, if $f(\cdot) \in L_{pe}$ then: (i) $\|f_T(\cdot)\|_p$ is a nondecreasing function of T.

(ii) $f(\cdot) \in L_p$ iff $\exists\, m > 0$ s.t. $\|f_T(\cdot)\|_p \le m < \infty, \ \forall T < \infty$

In this case $\|f(\cdot)\|_p = \lim_{T\to\infty} \|f_T(\cdot)\|_p$

- Note: for vector valued functions we may still define the corresponding L_p spaces as the set of $f(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ s.t. 1). $\|f(t)\|_{\mathbb{R}^n}$ is measurable

2) $\left\| \|f(t)\|_{\mathbb{R}^n} \right\|_p < \infty \quad p \in [1,\infty]$.

For simplicity, we write (2) as $\|f\|_p < \infty$.

Also, observe that, since norms in $\mathbb{R}^n$ are equivalent, any $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_+$ can be used without changing the qualitative characteristics of the analysis. The $\mathbb{R}^n$-norm selection does, however, affect the quantitative aspects of the analysis as it affects the way distance is measured.

REM: A similar development can be extended to the space of sequences $\{x_i\}_1^\infty : \mathbb{N} \to \mathbb{R}^n$

E.g. $\{x_i\}_1^\infty \in \ell_p$; $p \in [1,\infty]$ if

$$\|x\|_p \triangleq \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} \text{ exists and is finite } \quad (p \in [1,\infty))$$

$$\|x\|_\infty = \sup_i |x_i| < \infty \qquad (p = \infty)$$

CAUSALITY

Let A denote ~~the~~ mapping between the input and the output of a system i.e.

$$y = Au \qquad (\text{or } y(\cdot) = (Au)(\cdot))$$

Then a causal system is one where the value of the output at any time t depends on the values of the input up to time t. More precisely,

Def A mapping $A : L_{pe}^n \to L_{pe}^n$ is said to be causal if

$$(Au)_T = (Au_T)_T \quad \forall\, T < \infty, \ \forall\, u \in L_{pe}^n$$

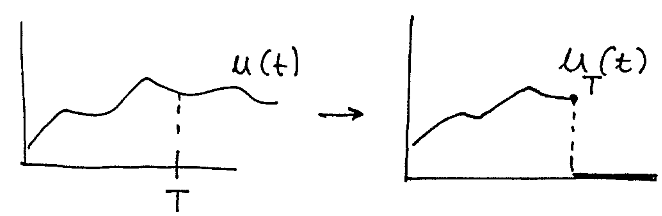
Alternatively, A is causal iff whenever $u_1, u_2 \in L_{pe}^n$ and

$$u_{1T} = u_{2T} \quad \text{for some } T < \infty$$

we have

$$(Au_1)_T = (Au_2)_T$$

(u_T denotes the truncation of u at T)



INPUT-OUTPUT STABILITY

Consider the LTI system

$\mathcal{H}: \quad \dot{x} = Ax + bu \; ; \; y = cx \qquad x(0) = 0.$

The input-output relationship of this system can be described in terms of a convolution integral which defines the mapping:

$$\mathcal{H}: u \longrightarrow \mathcal{H}u \; (= y)$$

i.e. $\quad y(t) = \int_0^t h(t-\tau)\, u(\tau)\, d\tau$

where $h(.)$ is also known as the impulse response of the system $\mathcal{H}$ $\left(h(t) = c e^{At} b\right)$

Also, assuming that the various Laplace transforms exist:

$$Y(s) = H(s)\, U(s)$$

$$H(s) = \mathcal{L}\{h(t)\}$$

Def. Let $A: L_{pe}^n \rightarrow L_{pe}^m$. We say that the mapping A (or the system represented by the mapping A) is L_p stable if

1) $Af \in L_p^m$ whenever $f \in L_p^n$
2) $\exists$ constants $k, b \; (< \infty)$:

$$\|Af\|_p \leq k \|f\|_p + b \quad \forall f \in L_p^n$$

e.g. $p = \infty$: BIBO stability

- Let $(Af)(t) = \int_0^t e^{(t-\tau)} f(\tau)\, d\tau$

  $A: L_{\infty e} \rightarrow L_{\infty e}$

  But if $f(t) \equiv 1$, $Af(t) = e^t - 1 \notin L_\infty$

  $\therefore$ A is not L_∞-stable.

- Let $(Af)(t) = \int_0^t e^{-(t-\tau)} f(\tau)\, d\tau \quad (A: L_{\infty e} \rightarrow L_{\infty e})$

$$\|(Af)(\cdot)\|_\infty \leq \sup_{t \geq 0} |f(t)| \cdot \sup_t \left[\int_0^t e^{-(t-\tau)} d\tau\right]$$

$$\leq \sup_{t \geq 0} |f(t)| = \|f\|_\infty$$

$\therefore$ A is L_∞ stable.

INDUCED NORMS OF LINEAR MAPS

Let $H: u \to Hu \triangleq h * u$ i.e.

$$Hu(t) = \int_0^t h(t-\tau)\, u(\tau)\, d\tau, \quad t \in \mathbb{R}_+.$$

Suppose that $\|h\|_1 = \int_0^\infty |h(t)|\, dt < \infty$

U.t.c.

a). $H: L_\infty \to L_\infty$

b) $\|h * u\|_\infty \leq \|h\|_1 \|u\|_\infty \quad \forall u \in L_\infty$.

and $\|h * u\|_\infty$ can be made arbitrarily close to $\|h\|_1 \|u\|_\infty$ by an appropriate choice of u.

Def: Let $|\cdot|$ be a norm on a linear space E and let A be a linear map $E \to E$.

Define $\|\cdot\|_i$: $\|A\|_i = \sup\limits_{x \neq 0} \dfrac{|Ax|}{|x|}$

$\|A\|_i$ is called the induced norm of A or the operator norm induced by the vector norm $|\cdot|$, or the gain of the operator $A : (E, |\cdot|) \to (E, |\cdot|)$.

Lem If A is a linear map $E \to E$, then the following statements are equivalent

(i) the linear function A is continuous at $0 \in E$.

(ii) the linear function A is continuous on E

(iii) The induced norm of A, is finite.

Rem $A: L_{pe} \to L_{pe}$ is L_p-stable if its induced norm on L_p is finite. Note that its induced norm will be the smallest const. K satisfying the condition given in the definition. The constant b is to cover cases of Affine maps or, in dynamical systems, initial conditions etc.

Ex: Let $H : u \to Hu \triangleq h * u$.

Then $\|H\|_{i\infty} = \|h\|_1$.

Further, assume $\|h_1\|$ is finite i.e., $h \in L_1$. Then,

(i) $H : L_2 \to L_2$

(ii) $\|H\|_{i2} = \max_{\omega \in \mathbb{R}} |\hat{h}(j\omega)|$

where $\hat{h}(s) = \mathcal{L}\{h(t)\}$.

- Assume $u \in \mathcal{L}_p$, $h \in \mathcal{L}_1$. Then for any $p \in [1, \infty]$

$$\|y\|_p \triangleq \|h * u\|_p \leq \|h\|_1 \|u\|_p .$$

The inequality is sharp for $p = 1, \infty$ only.

- Let $h(t)$ s.t. $\hat{h}(s)$ exists and is proper, rational. Then there exist A, B, C, D s.t. $\dot{x} = Ax + Bu$; $y = Cx + Du$.

has the I/O relationship $y = h * u$ (if $D \neq 0$, h contains an impulse distribution at 0).

u.t.c. $\hat{h}(s)$ is analytic in the RHP ($\mathrm{Re}[s] \geq 0$) iff $h \in \mathcal{L}_1$.

Furthermore, if $D = 0$,

1) $|h(t)| \leq \alpha_1 e^{-\alpha_0 t}$, for some $\alpha_1, \alpha_0 > 0$

2) $u \in \mathcal{L}_1 \Rightarrow y \in \mathcal{L}_1 \cap \mathcal{L}_\infty$, $\dot{y} \in \mathcal{L}_1$, y is continuous and $\lim_{t\to\infty} y = 0$

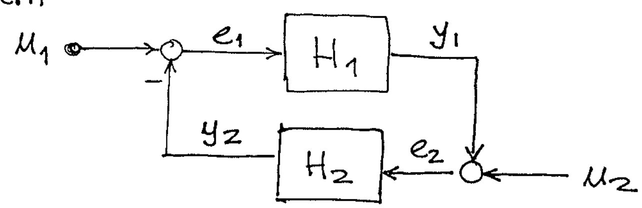
3) $u \in \mathcal{L}_2 \Rightarrow y \in \mathcal{L}_2 \cap \mathcal{L}_\infty$, $\dot{y} \in \mathcal{L}_2$, y is continuous and $\lim_{t\to\infty} y = 0$

4) For $p \in [1, \infty]$ $u \in L_p \Rightarrow y, \dot{y} \in L_p$ and y is continuous.

(see more details in Desoer + Vidyasagar)

FEEDBACK SYSTEMS

Consider the following general feedback system



$$e_1 = u_1 - H_2 e_2$$

$$e_2 = u_2 + H_1 e_1$$

where u_i, y_i and e_i $(i=1,2)$ are functions of time, usually defined for $t \geq 0$ and take values in $\mathbb{R}$ or $\mathbb{R}^n$. H_i are operators acting on its respective input e_i to produce an output y_i. The general problem under investigation is: given some assumptions on H_1, H_2, show that if u_1, u_2 belong to some

class, then e_1, e_2 and y_1, y_2 also belong to the same class.

SMALL GAIN THEOREM

The small gain theorem is a very general theorem which gives sufficient conditions under which a "bounded input" produces a "bounded output".

In our general feedback system setup, let $(L, \|\cdot\|)$ denote any $(L_p, \|\cdot\|_p)$ space and L_e be its extension (L_{pe}).

THM: Let $H_1, H_2 : L_e \to L_e$ and $e_1, e_2 \in L_e$ and define

$$u_1 = e_1 + H_2 e_2$$

$$u_2 = e_2 - H_1 e_1$$

Suppose that there exist constants

$\beta_1, \beta_2, \ \gamma_1 \geq 0, \ \gamma_2 \geq 0$ s.t.

$$\left.\begin{aligned} \|(H_1 e_1)_T\| &\leq \gamma_1 \|e_{1T}\| + \beta_1 \\ \|(H_2 e_2)_T\| &\leq \gamma_2 \|e_{2T}\| + \beta_2 \end{aligned}\right\} \forall\, T \in \mathbb{R}_+$$

Utc. if $\gamma_1 \cdot \gamma_2 < 1$ then

(1) $\|e_{1T}\| \leq (1-\gamma_1\gamma_2)^{-1}\left[\|u_{1T}\| + \gamma_2 \|u_{2T}\| + \beta_2 + \gamma_2 \beta_1\right]$

$\|e_{2T}\| \leq (1-\gamma_1\gamma_2)^{-1}\left[\|u_{2T}\| + \gamma_1 \|u_{1T}\| + \beta_1 + \gamma_1 \beta_2\right]$

(2) if, in addition, $\|u_1\|, \|u_2\| < \infty$ then

e_1, e_2, y_1, y_2. have finite norms.

<u>Rem</u>: If H_1 is causal then its gain condition can be replaced by: $\|H_1 x\| \leq \gamma_1 \|x\| + \beta_1 \quad \forall\, x \in \underline{\underline{L}}$

The interpretation of the theorem is that if the product of the gains of H_1 and H_2 is smaller than 1 then, provided that a solution exists, any bounded input (u_1, u_2)

produces a bounded output (y_1, y_2) and the map $(u_1, u_2) \rightarrow (y_1, y_2)$ has also finite gain.

Also note, that the theorem assumes the existence of e_1, e_2 from which u_1, u_2 are calculated, thus avoiding questions of existence, uniqueness and continuous dependence of solutions which must be established separately.

<u>SMALL GAIN THEOREM</u>: INCREMENTAL FORM

In the previous setup, assume that there exist $\tilde{\gamma}_1, \tilde{\gamma}_2$ s.t. $\forall\, T \in \mathbb{R}^+$ and $\forall\, \xi, \xi' \in L_e$

$$\left.\begin{aligned} \|(H_1\xi)_T - (H_1\xi')_T\| &\leq \tilde{\gamma}_1 \|\xi_T - \xi'_T\| \\ \|(H_2\xi)_T - (H_2\xi')_T\| &\leq \tilde{\gamma}_2 \|\xi_T - \xi'_T\| \end{aligned}\right. \qquad (*)$$

If $\tilde{\gamma}_1\tilde{\gamma}_2 < 1$ then

(1) $\forall\, u_1, u_2 \in \mathcal{L}_e$ $\exists$ a unique sol'n $e_1, e_2, y_1, y_2 \in \mathcal{L}_e$ which can be obtained iteratively.

(2) The map $(u_1, u_2) \to (e_1, e_2)$ is unif. cont. on $P_T \mathcal{L}_e \times P_T \mathcal{L}_e$ and on $\mathcal{L} \times \mathcal{L}$

(P_T denotes the truncation operator at T)

(3) if, in addition, the sol'n corresponding to $u_1 = u_2 = 0$ is in $\mathcal{L}$ then $u_1, u_2 \in \mathcal{L}$ $\Rightarrow e_1, e_2 \in \mathcal{L}$.

<u>REMARKS</u>: 1). If H_1 is a linear map s.t.

$$\|(H_1 \xi)_T\| \le \gamma_1 \|\xi_T\| \quad \forall\, \xi \in L_e,\ \forall\, T \in \mathbb{R}_+$$

then

$$\|(H_1(\xi - \xi'))_T\| \le \gamma_1 \|\xi_T - \xi_T'\| \quad \forall\, \xi, \xi' \in L_e,\ \forall\, T \in \mathbb{R}^+$$

2) The conditions (*) of the theorem imply that H_1, H_2 are causal.

Further more if $H : L_e \to L_e$ is a causal operator s.t. $\|(H\xi)_T - (H\xi')_T\| \le \tilde{\gamma} \|\xi_T - \xi_T'\|$ $\forall\, \xi, \xi' \in L_e$, $\forall\, T \in \mathbb{R}^+$, the smallest $\tilde{\gamma}$ which satisfies the above inequality is called the incremental gain of H.

3). Using the causality of H_1, H_2 we can write

$$e_{2T} = u_{2T} + \left\{ H_1\left[u_{1T} - (H_2 e_{2T})_T\right] \right\}_T = f(e_{2T})$$

Then, it is straightforward to show that f is a contraction on $P_T L_e$.

<u>THE LOOP TRANSFORMATION THEOREM</u>

Consider the feedback system

$$S : \begin{cases} u_1 = e_1 + H_2 e_2 \\ u_2 = e_2 - H_1 e_1 \end{cases}$$

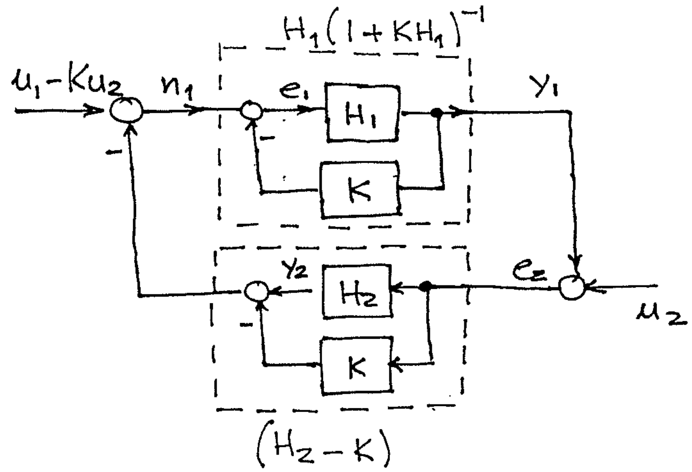
and let $K : L_e \to L_e$ and consider the system

$$S_K : \begin{cases} \bar{u}_1 = \eta_1 + (H_2 - K)e_2 \\ u_2 = e_2 - H_1(1+KH_1)^{-1}\eta_1 \end{cases}$$

where $(1+KH_1)^{-1}$ is assumed to exist : $L_e \to L_e$.

i.e. S_K can be obtained from S as follows

S_K:



<u>THM</u>: Let $H_1, H_2, K, (1+KH_1)^{-1}$ map $L_e \to L_e$. and let K be linear. U.t.c

a). If u_1, u_2, e_1, e_2 are in L_e and are solutions of S, then $(u_1 - Ku_2), u_2$, $\eta_1 = (1+KH_1)e_1$ and e_2 are in L_e and are solutions of S_K.

b) The converse of a is also true :

$(u_1 - Ku_2), u_2, \eta_1, e_2 \in L_e$ & sol'n of S_K

$\Rightarrow u_1, u_2, e_1 = (1+KH_1)^{-1}\eta_1, e_2 \in L_e$ & sol'n of S.

c) (a), (b) hold if L_e is <u>everywhere</u> replaced by L.

d) if $u_2 \equiv 0$, (a),(b),(c) hold even if K is nonlinear.

<u>Rem</u>. The loop transformation theorem is important because it allows the study of the stability of a feedback system to be performed on an "equivalent", more convenient feedback system. (see examples below).

<u>Thm</u> : Consider the equation

$$e(t) = u(t) + \int_0^t f(\tau, e(\tau), u(\tau))\, d\tau \qquad (*)$$

$u \in L_{\infty e}^n$, $u : \mathbb{R}_+ \to \mathbb{R}^n$

$f : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n$ is continuous and satisfies a global lipschitz condition, namely $\exists k$:

$$\| f(t, \xi, n) - f(t, \xi', n) \| \le k \| \xi - \xi' \|$$

$$\forall\, t \in \mathbb{R}_+ , \ \forall\, \xi, \xi', n \in \mathbb{R}^n$$

Then $(*)$ has, for each $u \in L_{\infty e}^n$, one and only one sol'n $e \in L_{\infty e}^n$

<u>Lem</u> Consider the system

$$\left.\begin{aligned} e_1 &= u_1 - y_2 \\ e_2 &= u_2 + y_1 \\ \\ y_1 &= G_1 e_1 \\ y_2 &= G_2 e_2 \end{aligned}\right\} (*)$$

and suppose that G_1, G_2 are of the form

$$(G_1 x)(t) = \int_0^t G(t,\tau)\, n_1(\tau, x(\tau))\, d\tau$$
$$(G_2 x)(t) = n_2(t, x(t))$$

where $G(\cdot,\cdot)$ is continuous , $G(t,t)$ is unif. bounded in $\mathbb{R}^+$ and n_1, n_2 satisfy

1) $n_i(t, 0) = 0 \quad \forall\, t \ge 0 \qquad i = 1, 2$

2) $\exists\, k_i \in \mathbb{R}^+$:

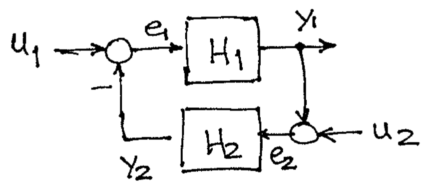
$$\| n_i(t,x) - n_i(t,y) \| \le k_i \| x - y \| \quad , \ i = 1, 2.$$
$$\forall\, t \ge 0, \ \forall\, x, y \in \mathbb{R}^n.$$

U.t.c. $G_1, G_2 : L_{pe}^n \to L_{pe}^n$. Further, given any $u_1, u_2 \in L_{pe}^n$ there exists exactly one set of $e_1, e_2, y_1, y_2 \in L_{pe}^n$ s.t. $(*)$ is satisfied.

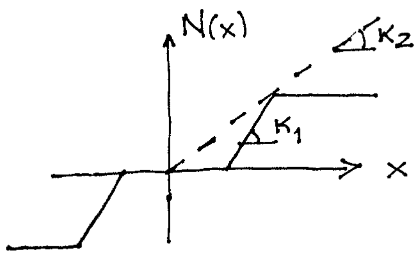
(See more details in Desoer + Vidyasagar, Vidyasagar)

SMALL GAIN THEOREM: EXAMPLES

Consider the closed loop system



Where H_1 is a (causal) LTI system with transfer function $H_1(s)$ and H_2 is the nonlinear function $N(x)$:



Assume that $e_i, y_i, u_i : \mathbb{R}^+ \to \mathbb{R}$.

and consider the space L_2 and its corresponding extended space L_{2e}.

Also assume that $H_1(s)$ is a a proper rational transfer function, analytic in the RHP.

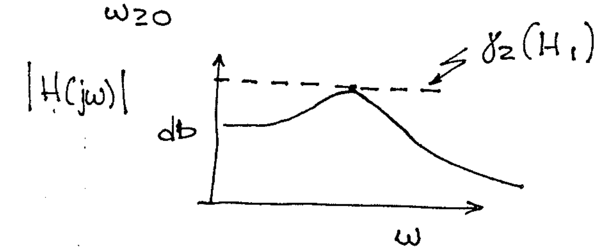
Then:

1). L_2 gain of H_1:

$$\gamma_2(H_1) = \|H_1(s)\|_\infty = \sup_{Re\, s \geq 0} |H_1(s)| =$$

$$= \max_{\omega \geq 0} |H_1(j\omega)|$$

eg.



2) Incremental L_2 gain of H_1

$$\tilde{\gamma}_2(H_1) = \gamma_2(H_1) \quad \text{(Linearity)}.$$

3). L_2 gain of H_2:

$$\gamma_2(H_2) = \sup_{\|x\| \neq 0} \left\{ \frac{\int_0^\infty N^2(x(t))\,dt}{\int_0^\infty x^2(t)\,dt} \right\}^{1/2}$$

Note : $|N(x)| \le K_2 \cdot |x|$

$\therefore \quad \gamma_2(H_2) \le \sup_{\|x\| \neq 0} \left\{ \frac{\int_0^\infty K_2^2 |x(t)|^2 dt}{\int_0^\infty |x(t)|^2 dt} \right\}^{1/2}$

$= K_2$

(= the supremum of the absolute slopes of lines drawn from the origin to points on the graph of $N(\cdot)$)

4). Incremental gain of H_2

$$\tilde{\gamma}_2(H_2) = \sup_{\|x_1 - x_2\| \neq 0} \left\{ \frac{\int_0^\infty |N(x_1(t)) - N(x_2(t))|^2 dt}{\int_0^\infty |x_1(t) - x_2(t)|^2 dt} \right\}^{1/2}$$

Note: $\| N(x_1) - N(x_2) \| \le K_1 \cdot \| x_1 - x_2 \|$

$\tilde{\gamma}_2(H_2) = K_1$

(= the Lipschitz constant of $N(\cdot)$ or the maximum absolute slope of all lines that are tangent to the graph of N)

Then, if $K_2 \cdot \| H_1(s) \|_\infty < 1$ and

. If the feedback system has solutions $e_i \in L_{2e}$ for $u_i \in L_2$, then $e_i \in L_2$ and

$$\| e_1 \|_2 \le \frac{1}{1 - K_2 \| H_1(s) \|_\infty} \left(\| u_1 \|_2 + K_2 \| u_2 \|_2 \right)$$

$$\| e_2 \|_2 \le \frac{1}{1 - K_2 \| H_1(s) \|_\infty} \left(\| u_2 \|_2 + \| H_1(s) \|_\infty \| u_1 \|_2 \right)$$

If, $K_1 \| H_1(s) \|_\infty < 1$ then $\forall u_i \in L_{2e}$ $e_i \in L_{2e}$ and are unique.

Further, the closed loop system is L_2 stable.

e.g. Let $H_1(s) = \frac{1}{s+1}$

Then $\| H_1(s) \|_\infty = 1$

Since H_1 : $y_1 = \int_0^t e^{-(t-\tau)} e_1(\tau) d\tau$

and H_2 : $y_2 = N(e_2)$ is Lipschitz
$N(0) = 0$

$u_i \in L_{2e} \Rightarrow e_i \in L_{2e}$

$\therefore$ if $K_2 < 1$

$u_i \in L_2 \Rightarrow e_i \in L_2$

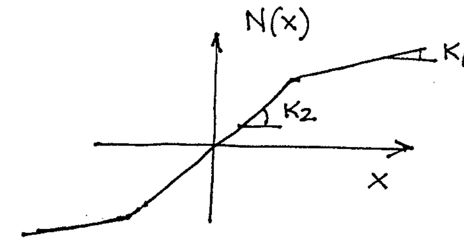
and $\|e_i\|_2 \leq \frac{1}{1-K_2} \left(\|u_1\|_2 + \|u_2\|_2 \right)$

(We have assumed throughout the example that the initial conditions of H_1 are 0. Otherwise, the constants β_1, β_2 should be included in the bounds for $\|e_i\|_2$).

Let us now consider the case where

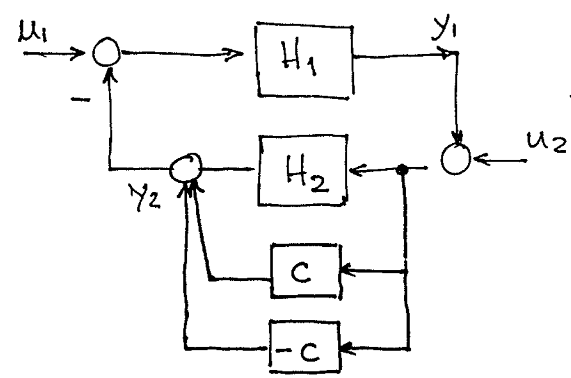
$H_1(s) = \frac{1}{s}$ which is not analytic in the [RHP] and $\|H_1(s)\|_\infty = \infty$.

Also assume that $H_2 : N_2(x)$ is of the form:

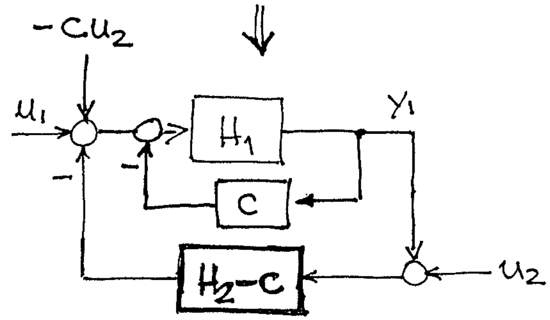


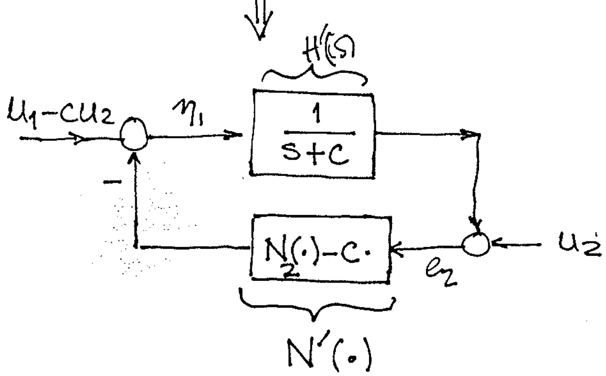
(~~saturation nonlinearity~~)

Since we cannot apply the small gain theorem let us employ the loop transformation theorem first to re-write the c.l. system as

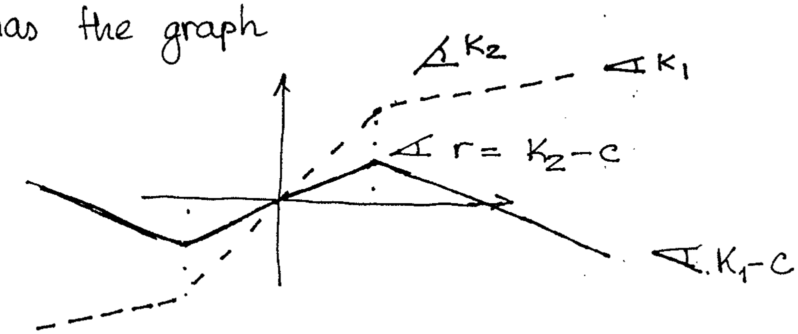


where c is a constant to be determined later





where, now, the transformed nonlinearity N' has the graph



Choose c : $|r| = |K_1 - c|$ $\therefore \begin{cases} r = \frac{K_2 - K_1}{2} \\ c = \frac{K_2 + K_1}{2} \end{cases}$

which implies that

$$|N'(x)| \leq r|x|$$

and

$$\gamma_2(H_1') = \left\| \frac{1}{s+c} \right\|_2 = \frac{2}{K_2 + K_1}$$

$$\gamma_2(H_2') = r = \frac{K_2 - K_1}{2}$$

Also note that in this case $N'(\cdot)$ has only two different slopes which, by the choice of c, are made equal in absolute value $\therefore \tilde{\gamma}_2(H_2') = r$.

$$\therefore \quad \gamma_2(H_1')\,\gamma_2(H_2') = \frac{K_2 - K_1}{K_2 + K_1} < 1$$

$\therefore$ by the small gain thm.

$$\|\eta_1\|_2 \le \frac{K_2+K_1}{2K_1}\left(\|u_1 - c u_2\|_2 + r\|u_2\|_2\right)$$

$$\|e_2\|_2 \le \frac{K_2+K_1}{2K_1}\left(\|u_2\|_2 + \frac{1}{c}\|u_1 - c u_2\|_2\right)$$

(Note that as in the previous example $\eta_1, e_2 \in L_{2e}$ for $u_i \in L_{2e}$)

and by the incremental small gain thm the transformed loop is L_2-stable.

$\therefore$ (Loop transformation thm) the original closed loop is L_2-stable.

Rem: Again, notice that in the presence of initial conditions β_1 should be included in the bounds obtained by the small gain thm.

• The constants r, c are usually referred to as the "radius" and the "center" of the cone of H_2. In general we say that $H : L_e \to L_e$ is "interior conic" if there exist constants $r \ge 0$, $c \in \mathbb{R}$ s.t.

$$\|(Hx)_T - c x_T\| \le r\|x_T\| \quad \forall x \in L_e,\ \forall T \ge 0$$

(References: Desoer + Vidyasagar,

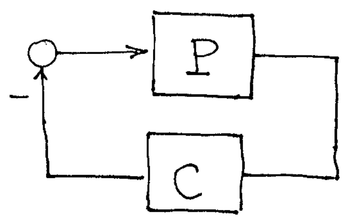
Zames: "On the Input-Output stability of Time-Varying Nonlinear Feedback Systems. Parts I, II", IEEE AC-11 April 66

And an interesting extension center + radius

Safonov: "Propagation of Conic Model Uncertainty in Hierarchical Systems", IEEE CAS-30, June 83)

Another example of the use of the small gain thm. is the following robustness problem:

Consider the closed loop system



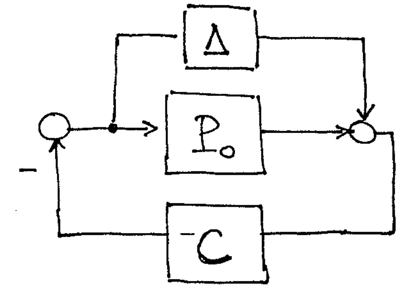
where $P, C : L_{2e} \to L_{2e}$ are LTI systems with transfer functions $P(s), C(s)$.

Suppose that $P(s)$ is given as

$$P(s) = P_0(s) + \Delta(s)$$

where $P_0(s)$ is known and $\Delta(s)$ is unknown, analytic in the RHP and such that $\|\Delta(s)\|_\infty \leq 1$.

i.e.

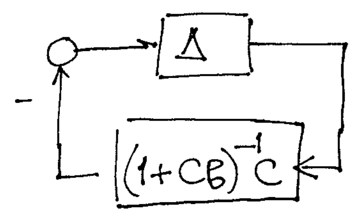


Further, suppose that C is at the disposal of the designer.

Our objective is to find some conditions on C s.t. the closed loop is L_2 stable for any $\Delta(s)$ satisfying the previous assumptions.

For simplicity let C be LTI with t.f. $C(s)$.

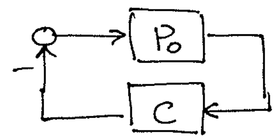
Then the closed loop can be written as



Since the closed loop should be stable $\forall \|\Delta(s)\|_\infty \leq 1$

C must be s.t.

- (1) the closed loop



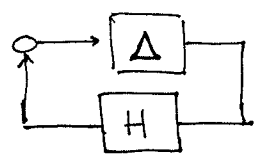
"NOMINAL CLOSED LOOP"

is L_2 stable.

(For this to be possible, certain conditions should be imposed on P_o)

(1) implies that $\|(1+CP_o)^{-1}C(s)\|_\infty < \infty$.

Furthermore from the small gain thm. we have that for any $\Delta(s)$ analytic in the RHP the closed loop



will be L_2 stable if

$$\|\Delta(s)\|_\infty \|H(s)\|_\infty < 1$$

Hence C must be s.t.

- (2) $\|(1+CP_o)^{-1}C(s)\|_\infty < 1$.

<u>REM</u> i) If the nominal c.l. is internally stable then (2) also implies that the perturbed closed loop will be internally stable.

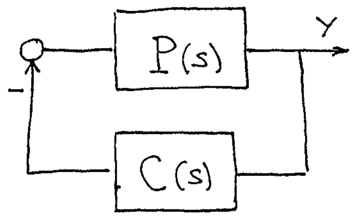
(ii) Under (1)+(2) + Linearity, the existence + uniqueness of solutions is also guaranteed.

(iii) (1)+(2) leave certain freedom in the choice of C(s). Consequently C(s) can be selected to optimize a performance objective, e.g. disturbance rejection, under the constraints (1)+(2)

(References: Francis, "A course in H_∞ Control Theory", Springer-Verlag and ref's therein).

CONTROLLER DESIGN FOR LTI SYSTEMS

Consider the closed loop system



(*)

where $P(s)$, $C(s)$ are the transfer functions of the plant and the controller respectively. It is assumed that the plant P is causal, FDLTI and completely controllable and observable. Our objective is to design $C(s)$ s.t. the closed loop is exponentially stable. Among the various design thechniques, the following three are of particular interest in Adaptive control, because they can be performed in a systematic way and yield closed-form solutions :

1). MODEL REFERENCE CONTROL (MRC)
2) POLE-PLACEMENT CONTROL (PPC)
3) LINEAR QUADRATIC CONTROL (LQ)

Controllers satisfying either 1 or 2 can be designed using algebraic methods and will be considered first.

Let $P(s) = \dfrac{N_p(s)}{D_p(s)}$ where $N_p(s)$, $D_p(s)$ are polynomials of 's' (siso case)

and $C(s) = \dfrac{N_1(s)}{N_2(s)}$ where $N_1(s)$, $N_2(s)$ are polynomials of 's' to be determined.

Then, the characteristic equation of (*) can be written as

$$1 + PC = 0 \quad \text{or,}$$

$$D_p(s)\, N_2(s) + N_p(s)\, N_1(s) = 0.$$

The PPC objective can then be stated as :

"design N_1, N_2 s.t. the characteristic equation of the closed loop has all its roots on prespecified locations in the LHP."

In other words, we want to find N_1, N_2 s.t.

$$D_p(s) N_2(s) + N_p(s) N_1(s) = A_*(s) \qquad (\#)$$

where $A_*(s)$ is the desired characteristic polynomial.

To assess the solvability of (#) we need some properties of polynomials :

Def. Two polynomials D, N are said to be coprime if there exist polynomials P, Q s.t.

$$DP + NQ = 1.$$

This definition is actually quite general and can be used to define coprimeness in more general algebraic structures.

In our case, it can be shown that two polynomials are coprime iff their only common factors are constants.

THM : If $D(s), N(s)$ are coprime and of degree n, m $(n > m)$ respectively then for any given $A_*(s)$ of degree $n_a \leq n+m$ the following equation

$$D(s) P(s) + N(s) Q(s) = A_*(s)$$

has a unique solution for $P(s), Q(s)$ with

$\deg[P] \leq m$, $\deg(Q] \leq n-1$

Rem: Equations of the form (#) are usually referred to as 'DIOPHANTINE' EQUATIONS (or 'BEZOUT' EQNS when $A_*(s)=1$)

Note that in the compensator design framework $C(s)$ must be a proper transfer function so that its implementation will be differentiator free.

i.e. $\deg[N_2] \geq \deg[N_1]$.

Cor: Let $P(s) = \frac{N_p(s)}{D_p(s)}$ where $D_p(s)$ is a monic polynomial of degree n and $N_p(s)$ is of degree $m \leq n-1$. Assume that $D_p(s), N_p(s)$ are coprime.
Then, for any $A_*(s)$ monic of degree $\leq 2n-1$, there exist polynomials $N_1(s), N_2(s)$ of degree $\leq n-1$, $N_2(s)$ monic of degree $n-1$ s.t. the characteristic equation of (*) with

$C(s) = \frac{N_1(s)}{N_2(s)}$ is $A_*(s) = 0$.

THM (Sylvester's Thm). Two polynomials $D(s), N(s)$ of degree n, m respectively, are coprime iff the matrix

$$S = \left.\underbrace{\begin{bmatrix} a_0 & & & b_0 & & & \\ a_1 & a_0 & & b_1 & b_0 & & \\ \vdots & a_1 & & \vdots & b_1 & & \\ \vdots & \vdots & a_0 & b_m & \vdots & & b_0 \\ a_n & \vdots \cdots & a_1 & 0 & b_m & \cdots & b_1 \\ 0 & a_n & & \vdots & 0 & & \vdots \\ \vdots & 0 & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_n & 0 & 0 & & b_m \end{bmatrix}}_{\leftarrow m \rightarrow \quad \leftarrow n \rightarrow}\right\} m+n$$

is nonsingular $(\det(S) \neq 0)$, where

$$D(s) = a_0 s^n + a_1 s^{n-1} + \ldots + a_n$$
$$N(s) = b_0 s^m + b_1 s^{m-1} + \ldots + b_m$$

e.g. Let $D = s^2 + \alpha s + 1$
$N = s + b$

Then

$$S = \begin{bmatrix} 1 & 1 & 0 \\ \alpha & b & 1 \\ 1 & 0 & b \end{bmatrix}$$

and $\det(S) = 1 - ab + b^2$

$\therefore$ for $\det S$ to be nonsingular $b \neq \frac{a \pm \sqrt{a^2-4}}{2}$

Note that $D(s) = (s-\alpha_1)(s-\alpha_2)$

where $\alpha_{1,2} = \dfrac{-\alpha \pm \sqrt{\alpha^2 - 4}}{2}$

and for coprimeness $-b \neq \alpha_1$ and $-b \neq \alpha_2$.

REM The Diophantine equation

$$D(s)\,P(s) + N(s)\,Q(s) = A_*(s) \qquad \begin{cases} \partial D = n \\ \partial N = m \le n-1 \\ \partial A_* = 2n-1 \end{cases}$$

can be written as a system of linear algebraic equations

$$\begin{bmatrix} L & \\ X & S \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} a \end{bmatrix}$$

where : • p, q are vectors containing the coefficients of $P(s)$, $Q(s)$ $(\partial P, \partial Q = n-1)$

• a is a vector containing the coefficients of $A_*(s)$

• L is a lower triangular matrix with diagonal elements l_{ii} = the leading coeff. of $D(s)$

• S is the Sylvester matrix of $D(s)$ and $N(s)$

• X is some matrix, generally $\neq 0$.

Controller Realization Using E.S. Filters

Consider $\quad u = C(s)\,y$

where $C(s) = \dfrac{N_1(s)}{N_2(s)}$ and let $D(s)$ be a Hurwitz polynomial (roots in LHP). Then $C(s)$ can be realized as follows:

$$u = \frac{D}{N_2} \cdot \left(\frac{N_1}{D}\, y \right)$$

$$\therefore \quad \frac{N_2}{D}\, u = \frac{N_1}{D}\, y.$$

$$\therefore \quad u + \frac{N_2 - D}{D}\, u = \frac{N_1}{D}\, y$$

$$\therefore \quad u = \frac{D - N_2}{D}\, u + \frac{N_1}{D}\, y.$$

i.e. let :

$$\dot{w}_1 = F w_1 + q u$$
$$v_1 = \theta_1 w_1$$
$$\dot{w}_2 = F w_2 + q y$$
$$v_2 = \theta_2 w_2$$

where : $\det(sI - F) = D(s)$ (Hurwitz)
F, q a completely controllable pair

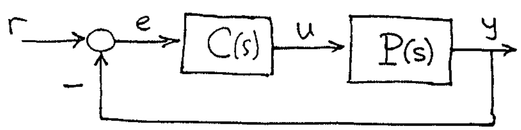
and θ_1, θ_2 are s.t.

$$\frac{D(s)-N_2(s)}{D(s)} = \theta_1 (sI-F)^{-1} q$$

$$\frac{N_1(s)}{D(s)} = \theta_2 (sI-F)^{-1} q$$

1). PPC

Consider the feedback system



Where $P(s) = \dfrac{N_p(s)}{D_p(s)}$

$$\partial D_p = n \quad \partial N_p = m \le n-1$$

Then, $C(s)$ can be selected as $C(s) = \dfrac{N_1(s)}{N_2(s)}$ and realized with ES filters (stable internal cancellations) s.t. $D_p N_2 + N_p N_1 = A_*$ is a monic Hurwitz polynomial of degree $2n-1$. In this case, the closed loop system will be internally stable and

$$y = \frac{N_p N_1}{A_*} r \; ; \; u = \frac{D_p N_1}{A_*} r$$

(Verify!)

2. PPC/IMP

The Internal model principle can be employed to design a controller s.t. the output tracks a <u>class</u> of reference inputs (or rejects a class of output disturbance)

e.g. Consider the class of reference inputs described by

$$L(s)\, r = 0$$

where $L(s)$ has distinct roots on the $j\omega$-axis e.g. $L(s) = s,\ s^2+\omega^2$, etc. Furthermore, suppose that $L(s)$ and $N_p(s)$ are coprime polynomials.

Then $C(s)$ can be designed as follows:

$$C(s) = \frac{N_1'}{N_2'} \frac{1}{L}$$

where $N_1'(s)$, $N_2'(s)$ satisfy the Diophantine eqn:

$$[D_p(s)L(s)]\,N_2'(s) + N_p(s)N_1'(s) = A_*'(s)$$

where
$$\partial A_* = \partial D_p + \partial L + \partial D_p - 1 = 2n+\ell-1$$
$$\partial N_1' = \partial D_p + \partial L - 1 = n+\ell-1$$
$$\partial N_2' = n-1$$

and A_* is a monic Hurwitz polynomial. As in the PPC case, the closed loop will be internally stable and
$$y = \frac{N_pN_1'}{A_*'}\,r.$$

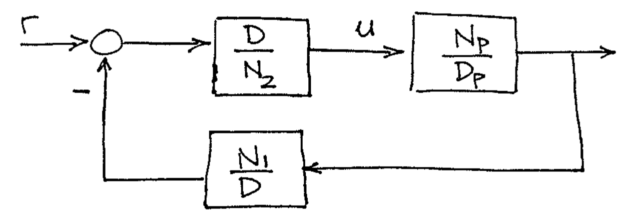
$$\therefore \quad y - r = \frac{N_pN_1' - A_*'}{A_*'}\,r = -\frac{D_pN_2'}{A_*'}[L\,r]$$

$\therefore \quad y - r \to 0$ as $t \to \infty$.

<u>3. MRC</u> Assume for the moment that N_p is a monic polynomial and consider the reference model
$$W_M(s) = \frac{N_M(s)}{D_M(s)}$$
where N_M, D_M are monic Hurwitz polynomials

s.t. $\partial D_M - \partial N_M = n-m$, $\partial D_M \le n$.

Then, if N_p is Hurwitz (min. phase assumption), a MRC can be constructed as shown below:



where N_1, N_2 satisfy
$$D_pN_2 + N_pN_1 = D_M N_p (D/N_M)$$
And D: N_M divides D.

In this case we have that the closed loop is internally stable (Note that N_p is Hurwitz)

and
$$\frac{y}{r} = \frac{N_pD}{D_M N_p (D/N_M)} = \frac{N_M}{D_M}.$$

Hence, as $t \to \infty$ $\quad y \to y_M = W_M(s)\,r$.

In the more general case where N_P, N_M are not monic we may write

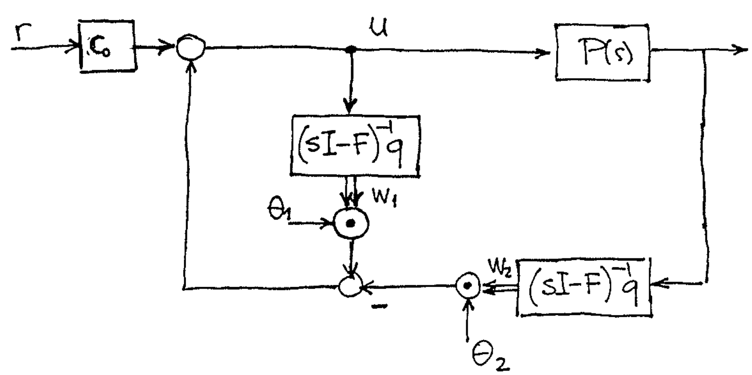
$$P(s) = K_P \frac{N_P'(s)}{D_P(s)}$$

$$W_M(s) = K_M \frac{N_M'(s)}{D_M(s)}$$

where N_M', N_P' are monic and K_M, K_P are called the high-frequency gain of W_M, P respectively.

An MRC is then constructed by using a feedforward gain $C_0 = \frac{K_M}{K_P}$ to pre-multiply r.

The general MRC closed-loop is shown below:



Rem The assumptions that the plant is "minimum phase" (N_P hurwitz) and that the relative degree of the plant ($\partial D_P - \partial N_P$) is to the relative degree of the reference model are the main drawbacks of the otherwise "convenient" MRC schemes.

LQ Let us consider the following problem:

Given: $\dot{x} = Ax + bu \qquad x(t_0) = x_0$

$y = cx \qquad t \geq t_0$

Choose $u(\cdot)$ to minimize the cost

$$J = \int_0^\infty \left[y^2(t) + r u^2(t) \right] dt \qquad r > 0.$$

For simplicity we will assume that $\{A, b, c\}$ is minimal i.e. $\{A, b\}$ is c.c., $\{c, A\}$ is c.o.

Note: This is a very particular case in the extensive optimal control theory.

(References : e.g. Kailath, "Linear Systems",

Bryson + Ho "Applied Optimal Control" (classic)

Anderson + Moore "Optimal Control, Linear Quadratic Methods" Prentice Hall, brand new! etc)

It can be shown that the optimal solution is to use

$$\boxed{u = -\bar{K}x}$$

where $\bar{K}$ is such that

$$\det(sI - A + b\bar{K}) = \prod_1^n (s - z_i)$$

and z_i are the left half-plane roots of

$$\Lambda(s) = a(s)\,a(-s) + r^{-1}\,b(s)\,b(-s).$$

(Note : $\partial\Lambda(s) = 2n$, Roots are symmetric w.r.t. $j\omega$-axis and there are no roots on the $j\omega$-axis)

The optimal $\bar{K}$ can be calculated as

$$\boxed{\bar{K} = b^T r^{-1} \bar{P}}$$

and $\bar{P}$ is the $\underline{\underline{PD}}$ solution of the algebraic Riccati Equation (ARE)

$$\boxed{A^T\bar{P} + \bar{P}A - \bar{P}br^{-1}b^T\bar{P} + c^Tc = 0}$$

Note that 1) $\bar{P}$ is symmetric

2) The ARE has more than one solutions but there is only one which yields a stabilizing $\bar{K}$ and that is the positive definite one.

Rem Suppose that we wish to minimize the cost $J_\alpha = \int_0^\infty (y^2 + r u^2) e^{2\alpha t}\,dt$.

This problem can be reduced to the standard one by introducing $x_\alpha = e^{\alpha t}x$, $u_\alpha = e^{\alpha t}u$.

Moreover the closed loop poles of the system minimizing J_α will have real parts less than $-\alpha$. The solution of this problem can be expressed as $u = -\bar{K}_\alpha x$ where

$$\bar{K}_\alpha = b^T r^{-1} \bar{P}_\alpha$$

and $\bar{P}_\alpha$ is the solution of the ARE with

A being replaced by $A+\alpha$.

(Guaranteed Stability margin).

Ref: Anderson + Moore "Linear Optimal Control" Prentice Hall 1971.

QUADRATIC REGULATOR: A simple example

Consider
$$\dot{x} = ax + u \qquad x(0) = x_0$$
$$y = x$$

and
$$J = \int_0^\infty (y^2 + ru^2)\,dt$$

Suppose $u = -kx$

$\therefore \quad x(t) = x(0)\exp[(a-k)t]$

$$\therefore \quad J(k) = \begin{cases} -\dfrac{(1+rk^2)(a-k)^{-1}}{2} x_0 & \text{if } a-k<0 \\ \infty & \text{if } a-k \geq 0 \end{cases}$$

Differentiating $J(k)$ wrt K we get

$$\frac{\partial J}{\partial k} = \left(\frac{2rk}{k-a} - \frac{1+rk^2}{(k-\alpha)^2} \right) \frac{x_0}{2}$$

For an extremum $\frac{\partial J}{\partial k} = 0$ and since $k-a > 0$

we get that the optimum k should satisfy

$$rk^2 - 2ark - 1 = 0$$

or
$$k = a \pm \sqrt{a^2 + r^{-1}}$$

Furthermore, $\frac{\partial^2 J}{\partial k} = \frac{ra^2+1}{(k-a)^3} > 0 \Rightarrow K-a > 0$

which is exactly the stabilizability condition.

Now $K-a>0 \Rightarrow \pm\sqrt{a^2+r^{-1}} > 0$ i.e. we must

choose $K = a + \sqrt{a^2+r^{-1}}$

Asymptotic results : as $r \to \infty$ (expensive control)

$K \to 2a$. Note that if $a>0$ the closed loop

pole will be the mirror image of a w.r.t. $j\omega$ axis.

As $r \to 0$, (cheap control) $K \to \sqrt{\frac{1}{r}}$ i.e. the

closed loop pole is moved deep in the LHP to produce a fast decaying closed loop response.

For intermediate values of r, the results are often most transparently presented in terms of a root-locus plot.

<u>Comments</u>: The weights of y and u in J $\left(= \int y^T Q y + u^T R u\right.$ for the multivariable $\left.\text{case}\right)$ are the "tuning" parameters of the closed loop response. In general, "cheap controls" tend to produce a Butterworth pattern in the c.l. pole locations wich has some advantages and some disadvantages (e.g. high overshoot). Additional problems "show up" when the state vector x is not measured directly but it is

estimated via an observer. In this case the "classic" robustness ppty of the LQ regulator : $\begin{cases} \text{Gain Margin} = \infty \\ \text{Phase Margin} \geq 60° \end{cases}$,

(Note : $|1 + \bar{K}(j\omega I - A)^{-1} b| \geq 1$)

is lost. Tuning the LQ $\overset{+\text{observer}}{\text{weights}}$ using singular value theory or Loop-transfer recovery methods has been shown to give "good" results.

<u>Ref</u>: • SAFONOU, LAUB, HARTMANN : "Feedback Properties of Multivariable Systems : The Role and Use of the Return Difference Matrix", IEEE AC FEB. 1981 (**)

• DOYLE + STEIN : "Multivariable Feedback Design: Concepts for a classical / Modern Synthesis" IEEE AC, FEB 1981 (**).

- STEIN + ATHANS : "The LQG/LTR Procedure for Multivariable Feedback control Design" IEEE AC, FEB 1987

- <u>OBSERVER BASED LQ CONTROL</u>.

If the states of the plant are not directly accessible, the implementation of the LQ Regulator requires the construction of an observer to estimate the plant state vector. E.g. (Full order observer)

Consider the plant:

$$\dot{x} = Ax + bu$$
$$y = cx$$

And design the filter (Kalman-Bucy)

$$\dot{\hat{x}} = A\hat{x} + bu + \lambda(\hat{y} - y)$$
$$\hat{y} = c\hat{x}$$

Then, letting $\hat{e} = \hat{x} - x$, $(\hat{e}_1 = c\hat{e})$ we have

$$\dot{\hat{e}} = A(\hat{x}-x) + bu - bu + \lambda c(\hat{x}-x)$$
$$= \underbrace{(A+\lambda c)}_{\Lambda}\hat{e}$$
$$\therefore \quad \hat{e}(t) = e^{\Lambda(t-t_0)}\hat{e}(0)$$

Assuming that (c, A) is c.o. we can find λ s.t. the eigenvalues of Λ are placed on arbitrary locations i.e. $\forall \alpha > 0$ we can find $\lambda(a)$ s.t. $\|e^{\Lambda(t-t_0)}\| \leq k\, e^{-\alpha(t-t_0)}$ $\forall\, t \geq t_0 \geq 0$ and some $K > 0$.

<u>Rem</u> : The design of such a λ is extremely simplified if (A, b, c) is in the observable canonical form.

We will close this note on controller design for LTI systems noting that other observer constructions are also possible eg. <u>Kreisselmeier</u> "On Adaptive State Regulation" IEEE AC Feb 82 and "Adaptive observers with Exponential Rate of convergence" IEEE AC Feb. 77. Such constructions follow similar principles and will be mentioned in the future, as necessary.

PLANT PARAMETRIZATIONS

Let $P(s) = \dfrac{N_p(s)}{D_p(s)}$; $y = P(s)u$.

<u>OBJECTIVE</u> : Find a parametric model

s.t.

$$y(t) = \theta_*^T \xi(t)$$

where y, ξ are signals available for measurement

and θ_* contains the plant parameters i.e.

coefficients of N_p, D_p.

We have :

$$D_p(s)\, y = N_p(s)\, u.$$

Let $D(s)$ be a Hurwitz polynomial

of degree $\eta = \partial D_p(s)$

Then,

$$\frac{D_p(s)}{D(s)}\, y = \frac{N_p(s)}{D(s)}\, u.$$

$$\therefore \left(1 + \frac{D_p(s) - D(s)}{D(s)}\right) y = \frac{N_p(s)}{D(s)}\, u$$

$$\therefore y = \frac{N_p(s)}{D(s)}\, u + \left[\frac{D(s) - D_p(s)}{D(s)}\right] y.$$

Note that $\partial N_p(s) < \partial D_p = \partial D$

$$\partial[D(s) - D_p(s)] < \partial D$$

$\therefore \dfrac{N_p}{D}, \dfrac{D-D_p}{D}$ can be realized as filters of the form

$$\dot{w} = Fw + bu \quad (\text{or } Fw + by).$$

$$v = \theta w$$

where $\det(sI - F) = D(s)$

(F, b) is a completely controllable pair

θ is Uniquely determined by N_p (or $D - D_p$)

For example,

$$\frac{N_P}{D} = \frac{\beta_0 s^m + \beta_1 s^{m-1} + \dots + \beta_{m+1}}{s^n + d_1 s^{n-1} + \dots + d_{n+1}}$$

$$= \left[0, \dots 0, \beta_0, \beta_1, \dots \beta_{m+1}\right] \begin{bmatrix} \frac{s^{n-1}}{D(s)} \\ \vdots \\ \frac{1}{D(s)} \end{bmatrix}$$

$$\frac{D - D_P}{D} = \frac{(d_1 - a_1)s^{n-1} + \dots + (d_{n+1} - a_{n+1})}{s^n + d_1 s^{n-1} + \dots + d_{n+1}}$$

$$\therefore \quad \dot{w}_1 = \underbrace{\begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -d_{n+1}, & \dots & & -d_1 \end{bmatrix}}_{F} w_1 + \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{b} u$$

$$v_1 = \left[\beta_{m+1}, \dots, \beta_0, 0, \dots 0\right] w_1$$

Similarly for $\dot{w}_2 = F w_2 + bu$

$$v_2 = \left[(d_{n+1} - a_{n+1}), \dots, (d_1 - a_1)\right] w_2$$

Hence, $\boxed{y = v_1 + v_2 = \theta_*^T w}$

where $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, $\theta_* = \begin{bmatrix} \beta_{m+1} \\ \vdots \\ \beta_0 \\ 0 \\ 0 \\ d_{n+1} - a_{n+1} \\ \vdots \\ d_1 - a_1 \end{bmatrix}$

PARAMETRIC MODELS w/ DYNAMIC UNCERTAINTY

1. ADDITIVE UNCERTAINTY.

$$y = P(s)\,u = \left[P_0(s) + \Delta(s)\right] u$$

$P_0(s)$: NOMINAL PLANT $= \frac{N_P}{D_P}$

$\Delta(s)$: ADDITIVE UNCERTAINTY.
POLES IN LHP. PROPER / STRICTLY PROPER

$$D_P(s)\, y = N_P(s)\, u + D_P(s)\, \Delta(s)\, u$$

$$\therefore \quad y = \theta_1^{*T} w_1 + \theta_2^{*T} w_2 + \underbrace{\frac{D_P}{D} \Delta(s)\, u}_{\eta}$$

(see previous example).

$\therefore$ For "successful" identification we need η to be small in some sense, i.e. the corresponding induced gain of $\frac{D_P}{D} \Delta$ to be small.

2. MULTIPLICATIVE UNCERTAINTY

$$y = P(s)\,u = P_0(s)\big(1+\Delta(s)\big)\,u.$$

$P_0(s)$: NOMINAL PLANT

$\Delta(s)$: MULTIPLICATIVE UNCERTAINTY
POLES IN LHP.

$$D_p(s)\,y = N_p(s)\,u + N_p(s)\,\Delta(s)\,u$$

$$\therefore \quad y = \theta_1^{*T} w_1 + \theta_2^{*T} w_2 + \underbrace{\frac{N_p(s)}{D(s)}\,\Delta(s)\,u}_{\eta}$$

Note : If P is proper (strictly proper)

then $\frac{N_p}{D}\Delta$ is proper (strictly proper).

but Δ is not necessarily proper.

Again η should be "small" i.e. the corresponding induced gain of $\frac{N_p}{D}\Delta$ should be small.

3. STABLE FACTOR PERTURBATIONS

The previous uncertainty models do not change the RHP poles of the plant (i.e. P, P_0 have the same RHP poles).

However, if we consider

$$P(s) = \frac{D_1 N_p + \Delta N}{D_1 D_p + \Delta D} \quad ; \quad D_1 \text{ Hurwitz s.t. } \partial D_1 D_p = \partial \Delta D$$

$P(s)$ may have different or even different number of RHP poles than $P_0 = \frac{N_p}{D_p}$.

This type of models arises when we consider an ODE of the form

$$\left(\dots + a_2\frac{d^2}{dt^2} + a_1\frac{d}{dt} + a_0\right) y = \left(\dots + \beta_2\frac{d^2}{dt^2} + \beta_1\frac{d}{dt} + \beta_0\right) u$$

where $\alpha_m \approx 0$, $\beta_m \approx 0$ for $m \geq M$

Let $\hat{D}_1$ be a Hurwitz polynomial of degree $\partial(D_1 D_p + \Delta D)$, s.t. D_1 is a factor of $\hat{D}_1$.

Then

$$\frac{D_1 D_p + \Delta D}{\hat{D}_1}\, y = \frac{D_1 N_p + \Delta N}{\hat{D}_1}\, u$$

$$\therefore \quad \frac{D_p}{D}\, y = \frac{N_p}{D}\, u + \frac{\Delta N}{\hat{D}_1}\, u - \frac{\Delta D}{\hat{D}_1}\, y$$

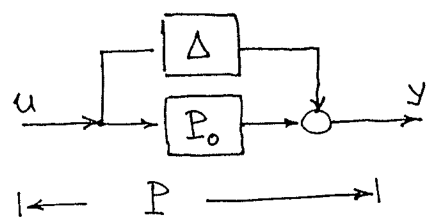
where $\mathcal{D} = \frac{\hat{D}_1}{D_1}$ $\quad (\partial D = \partial D_p)$

$$\therefore \quad y = \theta_1^{*T} w_1 + \theta_2^{*T} w_2 + \underbrace{\left[\frac{\Delta N}{\hat{D}_1}\, u - \frac{\Delta D}{\hat{D}_1}\, y \right]}_{\eta}$$

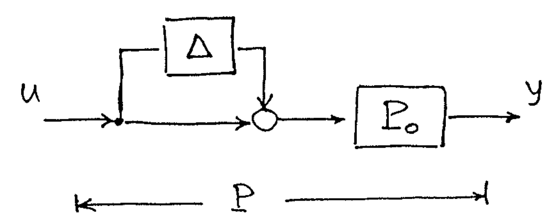
And η should be small i.e. $\frac{\Delta N}{\hat{D}_1}$, $\frac{\Delta D}{\hat{D}_1}$ should be small in terms of the corresponding induced gain.

Pictorially

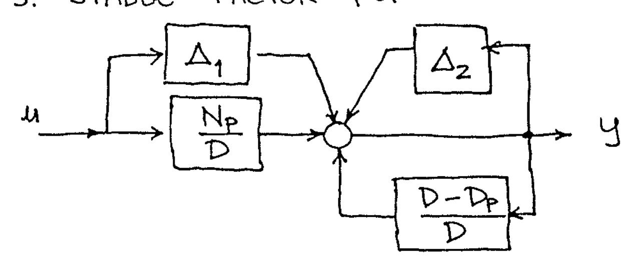
1. ADDITIVE UNCERTAINTY



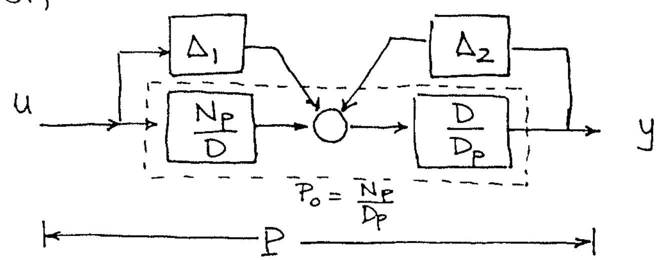
2. MULTIPLICATIVE UNCERTAINTY



3. STABLE FACTOR PERTURBATIONS



or,



$$\Delta_1 = \frac{\Delta N}{\hat{D}_1} \quad , \quad \Delta_2 = - \frac{\Delta D}{\hat{D}_1} y$$

Note that y can be expressed as:

$$y = \theta_*^T w + [\ \boldsymbol{\Delta}\ ] \begin{bmatrix} u \\ y \end{bmatrix}$$

where for the additive uncertainty

$$\boldsymbol{\Delta} = \left[\frac{D_P(s)}{D(s)} \Delta(s) \ , \ 0 \right]$$

for the multiplicative uncertainty

$$\boldsymbol{\Delta} = \left[\frac{N_P(s)}{D(s)} \Delta(s), \ 0 \right]$$

and for the stable factor perturbations

$$\boldsymbol{\Delta} = [\ \Delta_1 \ , \ \Delta_2 \ ]$$

Application of SPR

In MRAC we have that the tracking error can be expressed as

$$e_1 = y_P - y_m = W_M(s)(\phi^T w).$$

where $\phi = \theta - \theta_*$, w is a vector of auxiliary signals, $W_M(s)$ is the transfer function of the reference model.

With this as a motivation, let us consider the system

$$(*) \quad \begin{aligned} \dot{e} &= A e + b(\phi^T w) \\ e_1 &= c^T e \\ \dot{\phi} &= -\gamma e_1 w \end{aligned} \qquad c^T(sI-A)^{-1} b = W_M(s) \ , \quad \gamma > 0$$

and assume that $W_m(s)$ is SPR.

Then $\exists\ P = P^T > 0$ s.t. $\begin{array}{l} A^TP + PA = -qq^T - \varepsilon L \\ Pb = c \end{array}$

for some $L = L^T > 0$, $\varepsilon > 0$ and $q \in \mathbb{R}^n$.

Choose $V = e^T P e + \phi^T \frac{1}{\gamma} \phi$.

Then,
$$\begin{aligned} \mathring{V} &= e^T (A^T P + PA) e + 2 e^T P b\, \phi^T w \\ &\quad - 2 e_1 \phi^T w. \\ &= - \|q^T e\|^2 - \varepsilon\, e^T L e + 2 \underbrace{e^T c}_{e_1} \phi^T w \\ &\quad - 2 e_1 \phi^T w \\ &= - \|q^T e\|^2 - \varepsilon\, e^T L e \\ &\leq -\varepsilon_L \|e\|^2. \end{aligned}$$

where $\varepsilon_L = \varepsilon \cdot \min\{\lambda(L)\}$.

Hence, (*) is U.S. , V is UB, $\|e\| \in L_2$

etc.

80
87

SIMPLE ADAPTIVE CONTROL SCHEMES

1. DIRECT ADAPTIVE CONTROL

In direct adaptive control, the controller parameters are estimated / updated directly on line without using any explicit information about the plant parameters or their estimates.

1.1. Adaptive regulation

Consider the scalar plant

$$\mathring{x} = ax + u \quad ; \quad x(0) = x_0.$$

where a is constant but unknown. The control objective is to determine a bounded function $u = f(t,x)$ s.t. the state $x(t)$ is bounded and converges to 0 as $t \to \infty$, for any given initial condition x_0.

Let $a_m > 0$ be an a priori selected constant and suppose that $-a_m$ is the desired closed loop pole. If a were known then

$$u = -k^* x , \quad k^* = a + a_m$$

could be used to achieve the control objective. Since a is unknown let us use

$$u = - K(t)\, x$$

and search for an appropriate law to generate $K(t)$. Such laws can be developed by applying various parameter identification techniques to appropriate "parametric models" which are linear in the unknown k^*.

For example, from $k^* = a + a_m$ we have that $a = -a_m + k^*$. Hence the

plant can be expressed as

$$\dot{x} = -a_m x + k^* x - \underbrace{K(t)\, x}_{u} \qquad (*)$$

Let $k(t) = \tilde{k}(t) + k^*$ where $\tilde{k}(t)$ is the parameter error.

Hence, $\dot{x} = -a_m x + \left(-\tilde{k}(t)\, x\right)$

$$\Rightarrow \quad x = \frac{-1}{s+a_m} \left(\tilde{k} x \right)$$

A similar result can be obtained with a slightly different approach: Rewrite $(*)$ as

$$x = \frac{1}{s+a_m} \left(k^* x + u \right)$$

$$= \left(k^* \frac{1}{s+a_m} x + \frac{1}{s+a_m} u \right)$$

Obtain an estimate of x using K:

$$\hat{x} = \frac{1}{s+a_m} \left(K(t)\, x + u \right)$$

which motivates the definition of an estimation error:

$$\epsilon = \hat{x} - x = \frac{1}{s+a_m}(Kx+u) - x$$

$$= \frac{1}{s+a_m}(kx+u) - \frac{1}{s+a_m}(k^*\hat{x}+u)$$

$$= \frac{1}{s+a_m}\left[(K-K^*)x\right].$$

Note that $\epsilon = -x + \varepsilon_t$ where $\varepsilon_t = e^{-a_m t}\hat{x}(0)$. Such signals often appear when transfer functions are used to describe signals. And although they should be taken into account, their effect appears–usually– during the transient periods only, without altering the final stability / boundedness result.

Next, let us consider an update law for $K(t)$ $(\to \tilde{K}(t))$.

1. Using Lyapunov techniques,

let $\mathring{K} = \dot{\tilde{K}} = g(t, x, u, k, \epsilon)$ to be determined.

Choose $V(\epsilon, \phi) = \frac{\epsilon^2}{2} + \frac{\tilde{K}^2}{2\gamma}$

$$\therefore \mathring{V} = -a_m \epsilon^2 + \tilde{K} \cdot \left\{ \frac{1}{\gamma} g(\cdot) + x\epsilon \right\}$$

An obvious choice for $g(\cdot)$ is $g = -\gamma \epsilon x$

which gives

$$\dot{V} = -a_m \epsilon^2 \le 0$$

Hence, with $\boxed{\mathring{K} = -\gamma \epsilon x}$; $K(0) = K_0$

$\epsilon_e = 0$, $\tilde{k}_e = 0$ is a U.S. equilibrium.

Further, $\epsilon, \tilde{K}$ are U.B.

Since $\mathring{\epsilon} = -\alpha_m \epsilon + \tilde{k} x$

$$\therefore \quad = -\alpha_m \epsilon + \tilde{k}\varepsilon_t - \tilde{k}\epsilon$$

$\Rightarrow \dot{\epsilon}$ is U.B $(\in L_\infty)$

Next, $V(T) - V(0) = \int_0^T \dot{V}(t)\,dt$

$$\Rightarrow \int_0^T \epsilon^2 dt = \frac{V(0) - V(T)}{\alpha_m} < \infty \quad \text{since } V(T) \text{ is UB}$$

$\therefore$ since $\int_0^T \epsilon^2 dt$ is a nondecreasing function of T, $\int_0^\infty \epsilon^2 dt$ exists and is finite

$\therefore \epsilon \in L_2$.

Note that in this case we also have that $V(T)$ is non increasing function of T ($\dot{V} \le 0$), $V \ge 0$

hence $\inf_T V(T) = V(\infty)$ which exists

$$\therefore \int_0^\infty \epsilon^2 dt = \frac{V(0) - V(\infty)}{\alpha_m}$$

Further $\epsilon, \dot{\epsilon} \in L_\infty \Rightarrow \lim_{t\to\infty} \epsilon(t) = 0$

$\therefore x \in L_\infty$ and $x \to 0$ as $t \to \infty$.

And therefore $u \in L_\infty$ ($k, x \in L_\infty$)

$u \to 0$ as $t \to \infty$.

$\therefore u$ satisfies the control objective

An important question to ask at this point is whether $K(t) \to K^*$ as $t \to \infty$.

1. $\epsilon, x \in L_\infty$ $\epsilon, x \to 0$ as $t \to \infty$ $\Rightarrow$

$\dot{K} \to 0$ as $t \to \infty$

ie. adaptation switches off asymptotically with time. This fact alone, however, does not guarantee that $K \to$ const. let alone $K \to K^*$ (see HW #2).

For this simple example, though, $|\dot{K}| = \gamma|\epsilon x|$

$$\Rightarrow \int |\dot{K}| = \gamma \int |\epsilon x| \le \gamma \sqrt{\int |\epsilon|^2} \sqrt{\int |x|^2} < \infty$$

(Holder's Ineq.)

(In a more general case where $x \in L_\infty$, $\epsilon \in L_2$ and slightly different arguments should be used)

$\therefore$ (HW #2) K converges since $\int |\dot{K}| < \infty$.

To find the limit, note that

$$\lim_{t\to\infty} V(t) = V(\infty) = \lim \frac{\epsilon^2}{2} + \frac{\tilde{k}^2}{2\gamma}$$

$$\therefore \lim_{t\to\infty} \tilde{k}(t) = \pm\sqrt{2\gamma V(\infty)}.$$

or $\quad K(t) \to K^* \pm \sqrt{2\gamma V(\infty)}$

In addition to this, the simplicity of the example allows for the explicit derivation of the solution $\epsilon(t), x(t), k(t)$ i.e. for $\hat{x}(0)=0$

$$\epsilon(t) = \frac{2ce^{-ct}}{c+k_0-a+(c-k_0+a)e^{-2ct}} \cdot \epsilon(0)$$

$$K(t) = a + \frac{c\left[(c+k_0-a)e^{+2ct}-(c-k_0+a)\right]}{(c+k_0-a)e^{2ct}+(c-k_0+a)}$$

where $\quad c^2 = \gamma x_0^2 + (k_0-a)^2.$

Hence, if $c>0$, $\lim_{t\to\infty} K(t) = a+c$

$c<0 \quad \lim K(t) = a-c$

$$\therefore \lim_{t\to\infty} K(t) = K_\infty = \alpha + \sqrt{\gamma x_0^2 + (k_0-a)^2}$$

$\therefore$ for $x_0 \neq 0$, $K(t)$ converges to a stabilizing gain whose value depends on γ and the initial conditions x_0, k_0.

Furthermore in the limit as $t\to\infty$, the closed loop pole is $-K_\infty+\alpha$ which may be different from α_m. Since the control objective is to achieve signal boundedness and regulation of x to zero, the convergence of K to k^* is not crucial.

Note that when $x_0=0$ the system is at rest ($x_0=0$ $K=K_0=$ const.) and no adaptation takes place.

Finally, the adaptive gain γ affects both the transient behavior of the closed loop and the limiting value of $K(t)$. For a given $k_0, x_0 \neq 0$, Large γ will lead to large c $\therefore$ fast convergence of x to zero, but k_∞ will be large as well.

A different approach is to use a modified estimation error :

$$\epsilon = \frac{1}{s+a_m}\left[Kx + u - \epsilon x^2 \right] - x = \frac{1}{s+a_m}\left(\tilde{k}x - \epsilon x^2\right)$$

(The additional term $-\epsilon x^2$ is crucial for stability in the higher order case as well as for robustness)

Note that in this case ϵ UB $\not\Rightarrow$ x UB.

Nevertheless, consider

$$V = \frac{\epsilon^2}{2} + \frac{\tilde{k}^2}{2\gamma}$$

(this is not a Lyapunov function for the closed loop since it does not involve x.)

Then $\dot{V} = -a_m\epsilon^2 - \epsilon^2 x^2 + \epsilon\tilde{k}x + \frac{\tilde{k}\dot{\tilde{k}}}{\gamma}$

With $\dot{\tilde{k}} = -\gamma\epsilon x$

$\Rightarrow \quad \dot{V} = -a_m\epsilon^2 - \epsilon^2 x^2 \leq 0$

$\therefore \quad V \in L_\infty \Rightarrow \epsilon, \tilde{k} \in L_\infty$

& $\epsilon, \epsilon x \in L_2$.

& $\dot{\tilde{k}} \in L_2$

Independent of the boundedness of x.

Next, consider

$$\dot{x} = -a_m x - \tilde{k}x$$

Since $\tilde{k} \in L_\infty$, $x(t)$ cannot grow or decay faster than an exponential $\therefore$ x is continuous.

Further, since $\dot{\epsilon} = -a_m\epsilon + \tilde{k}x - \epsilon x^2$,
$\tilde{k}x = \dot{\epsilon} + a_m\epsilon + \epsilon x^2$

$\therefore \quad \dot{x} = -a_m(x+\epsilon) - \dot{\epsilon} - \epsilon x^2$.

Now let $\bar{x} = x + \epsilon$ (Note : $\epsilon \in L_\infty$, $\epsilon, \epsilon x \in L_2$)

$\therefore\ \dot{\bar{x}} = -a_m \bar{x} - \epsilon x^2$

$= -a_m \bar{x} - \epsilon x \bar{x} + \epsilon^2 x.$

i.e. $\dot{\bar{x}} = -a_m \bar{x} + \gamma_1 \bar{x} + \gamma_2$

where $\gamma_1 \in L_2$ (Note $\epsilon x \in L_2$)

$\gamma_2 \in L_2$ ($\epsilon \in L_\infty$, $\epsilon x \in L_2$)

$$\therefore\ \bar{x}(t) = e^{-\alpha_m t}\bar{x}_0 + \int_0^t e^{-\alpha_m(t-\tau)}\gamma_1 \bar{x} + \int_0^t e^{-\alpha_m(t-\tau)}\gamma_2$$

We will now use a slightly different form of pr # $\frac{3}{4}$, HW #2 :

A. $\int_0^t e^{-\alpha_m(t-\tau)}\gamma_2 \in L_2 \cap L_\infty$:

1. $\int_0^t e^{-\alpha_m(t-\tau)}\gamma_2 \le \sqrt{\int_0^t e^{-2\alpha_m(t-\tau)}}\sqrt{\int_0^t \gamma_2^2} < \frac{1}{\sqrt{2\alpha_m}}\|\gamma_2\|_2$

(Holder's Ineq.)

i.e. the $L_2 \to L_\infty$ induced gain of $\frac{1}{s+a_m}$ $\left(\int_0^t e^{-a_m(t-\tau)}\cdot\right)$ is finite

2. Let $y = \int_0^t e^{-\alpha_m(t-\tau)}\gamma_2 = \frac{1}{s+\alpha_m}\gamma_2$

The $L_2 \to L_2$ induced gain of $\frac{1}{s+a_m}$ is finite and equal to $\frac{1}{\alpha_m}$ $\therefore\ \|y\|_2 \le \frac{1}{\alpha_m}\|\gamma_2\|_2$

$$\therefore\ |\bar{x}(t)| \le e^{-\alpha_m t}|\bar{x}_0| + \Gamma_2(t) + \int_0^t e^{-\alpha_m(t-\tau)}|\gamma_1||\bar{x}|$$

$\Gamma_2 \in L_2 \cap L_\infty$, $\Gamma_2(t) \ge 0$.

$$|\bar{x}(t)|^2 \le \cancel{X}\left(e^{-2\alpha_m t}\bar{x}_0^2 + \Gamma_2^2(t)\right) + \cancel{X}\left(\int_0^t e^{-\alpha_m(t-\tau)} \cdot \int_0^t e^{-\alpha_m(t-\tau)} \cdot \gamma_1^2 \bar{x}^2 d\tau\right)$$

B. Bellman - GRONWALL LEMMA (B').

$$y(t) \le c(t) + \int_a^t \mu(s) y(s)\, ds \qquad a \le t \le b$$

$\mu, c \ge 0$ integrable over $[a,b]$

Then $y(t) \le c(t) + \int_a^t c(s)\mu(s)\exp\left[\int_s^t \mu(\tau)d\tau\right] ds$

Thus,

$$e^{\alpha_m t}\bar{x}^2 \le \lambda\left(e^{-\alpha_m t}\bar{x}_0^2 + e^{\alpha_m t}\Gamma_2^2\right) + \frac{\lambda}{a_m}\int_0^t e^{\alpha_m \tau}\gamma_1^2 \bar{x}^2\, d\tau$$

B-G.8'

$$\Rightarrow \bar{x}^2 \le \lambda\left(e^{-2\alpha_m t}\bar{x}_0^2 + \Gamma_2^2\right) + \frac{\lambda^2}{\alpha_m}\int_0^t e^{-\alpha_m(t-s)}\left(e^{-2a_m s}\bar{x}_0^2 + \Gamma_2^2\right)\cdot\gamma_1^2 \cdot e^{\frac{\lambda}{a_m}\int_s^t \gamma_1^2}\, ds$$

1. $\gamma_1 \in L_2 \Rightarrow \int_s^t \gamma_1^2 \le c_1$
2. $\Gamma_2 \in L_2 \cap L_\infty \Rightarrow \Gamma_2\gamma_1 \in L_2$

From which, $\bar{x} \in L_2 \cap L_\infty$.

$\therefore\ x \in L_2 \cap L_\infty$

Further $\dot{x} = -\alpha_m x - \tilde{k}x \in L_\infty$

$\therefore\ x \to 0$ as $t \to \infty$.

and $u = Kx \in L_\infty$

A GRADIENT METHOD.

Consider the parametric model of the plant

$$x = k^* \hat{x} + \hat{u}$$

where $\hat{x} = \frac{1}{s+a_m}x$, $\hat{u} = \frac{1}{s+a_m}u$. $(u = -kx)$

Then, $e = K\hat{x} + \hat{u} - x = \tilde{K}\hat{x}$

is the estimation error.

Let m be a normalizing signal, whose purpose will become obvious in the following. In this case we will simply take $m = 1 + |\hat{x}|^2$.

Consider now the cost

$$J(\phi,t) = \frac{\epsilon^2}{2m} = \frac{(\tilde{K}\hat{x})^2}{2m}$$

Note that $\frac{1}{m}$ is well defined and $\frac{\hat{x}^2}{m}$ is U.B.

Furthermore, $J \to \infty \Rightarrow \tilde{K} \to \infty \therefore$ for large

$\tilde{k}$, $-\nabla J$ will give the direction along which J decreases w.r.t. $\tilde{k}$.

Assuming that $\hat{x}(t)$ does not depend on $\tilde{k}(t)$, the gradient method gives

$$\dot{\tilde{k}} = -\gamma \frac{\partial J}{\partial \tilde{k}} = -\gamma \frac{\epsilon \hat{x}}{m} \quad , \quad \gamma > 0$$

Hence, the adaptively controlled plant becomes

$$\dot{x} = -\alpha_m x - \tilde{k} x$$
$$\dot{\hat{x}} = -\alpha_m \hat{x} + x$$
$$\dot{\tilde{k}} = -\gamma \frac{\epsilon \hat{x}}{m} = -\gamma \frac{\tilde{k} \hat{x}^2}{m} .$$

We now proceed in two steps: 1. establich boundedness of $\tilde{k}$ & $\epsilon/\sqrt{m} \in L_2$

2. boundedness + Convergence of $x, \hat{x}$.

1). Let $V = \frac{\tilde{k}^2}{2\gamma}$

Then $\dot{V} = -\frac{\epsilon \tilde{k} \hat{x}}{m} = -\frac{\epsilon^2}{m} \leq 0$

$\therefore V, \tilde{k} \in L_\infty$, $\frac{\epsilon}{\sqrt{m}} \in L_2$

Moreover, since $\frac{\hat{x}}{\sqrt{m}}$ is U.B $\Rightarrow$ $\frac{\epsilon \hat{x}}{m} \in L_2$.

From $\dot{\tilde{k}} = -\gamma \frac{\epsilon \hat{x}}{m}$ we have that $\dot{\tilde{k}} \in L_2$. Further $\tilde{k} \in L_\infty \Rightarrow x, \hat{x}$ cannot grow or decay faster than an exponential.

2). For the second part, observe that

$$x = \frac{1}{s+a_m} (-\tilde{k} x) \quad (+ \varepsilon_t)$$

$$\varepsilon_t = x_0 e^{-\alpha_m t} .$$

which can be written as

$$x = \frac{-1}{s+a_m} \left[\tilde{k} \left(\frac{s+a_m}{s+a_m} x \right) \right]$$

Thus

$$|\hat{x}|^2 \le ce^{-2\alpha t} + \frac{\lambda_1}{\alpha_m} \int_0^t e^{-\alpha_m(t-\tau)} \epsilon^2 + \frac{\lambda_2}{\alpha} \int_0^t e^{-\alpha(t-\tau)} (\dot{\tilde{K}} \hat{x})^2$$

$$\le ce^{-2\alpha t} + \lambda \int_0^t e^{-\alpha(t-\tau)} \left[\frac{\epsilon^2 + (\dot{\tilde{K}}\hat{x})^2}{m}\right] . m$$

$$\le \left[\frac{1}{\alpha} + ce^{-2\alpha t}\right] + \lambda \int_0^t e^{-\alpha(t-\tau)} \left[\frac{\epsilon^2 + (\dot{\tilde{K}}\hat{x})^2}{m}\right] \hat{x}^2$$

Using the Bellman-gronwall lemma B′

$$e^{\alpha t} |\hat{x}|^2 \le \left[\frac{1}{\alpha} e^{\alpha t} + ce^{-\alpha t}\right]$$

$$+ \int_0^t \left(\frac{e^{\alpha s}}{\alpha} + ce^{-\alpha s}\right) \lambda \left(\frac{\epsilon^2 + (\dot{\tilde{K}}\hat{x})^2}{m}\right)(s) \left[\exp \int_s^t \frac{\epsilon^2 + (\dot{\tilde{K}}\hat{x})^2}{m}(\tau) d\tau\right] ds$$

Now, $\frac{\epsilon}{\sqrt{m}} \in L_2 \cap L_\infty$, $\frac{\dot{\tilde{K}}\hat{x}}{\sqrt{m}} \in L_2 \cap L_\infty$

$\therefore |\hat{x}| \in L_\infty \Rightarrow m \in L_\infty$

Since $m \in L_\infty$, we have $\epsilon \in L_2 \cap L_\infty$

And from $x = \frac{1}{s+a_m} \epsilon \Rightarrow x \in L_2 \cap L_\infty$

From the properties of differentiation,

$$\tilde{K} s[x] = s[\tilde{K} x] - (s\tilde{K}) x$$

$$\left(\frac{d}{dt}(xy) = \frac{dx}{dt} y + x \frac{dy}{dt}\right)$$

$$\therefore \quad x = -\frac{1}{s+a_m} s+a_m \tilde{K} \frac{1}{s+a_m} x + \frac{1}{s+a_m} \dot{\tilde{K}} \frac{1}{s+a_m} x$$

$$\boxed{x = -\tilde{K}\hat{x} + \frac{1}{s+a_m} \dot{\tilde{K}}\hat{x}.} \quad (+\varepsilon_t)$$

$$\therefore \quad \hat{x} = -\frac{1}{s+a_m} \epsilon + \frac{1}{(s+a_m)^2} \dot{\tilde{K}}\hat{x}. \quad (+\varepsilon_t')$$

$$\therefore \quad |\hat{x}|^2 \le \cancel{\lambda}\left[\int_0^t e^{-\alpha_m(t-\tau)} \epsilon\right]^2 + \lambda_2 \left[\int_0^t e^{-\alpha(t-\tau)} \dot{\tilde{K}}\hat{x}\right]^2 + ce^{-2\alpha t}$$

where λ_1, λ_2, c are positive constants and $0 < \alpha < \alpha_m$.

Note that $\left(\int_0^t e^{-\alpha_m(t-\tau)} \epsilon\right)^2 \le \int_0^t e^{-\alpha_m(t-\tau)} \int_0^t e^{-\alpha_m(t-\tau)} \epsilon^2$

$$\le \frac{1}{\alpha_m} \int_0^t e^{-\alpha_m(t-\tau)} \epsilon^2$$

and $\hat{x} \in L_2 \cap L_\infty$., $u \in L_2 \cap L_\infty$
(Further $\dot{x}$, bounded $\Rightarrow$ $x \to 0$ etc.)

<u>Comments</u> In this simple example we have used some techniques which seem to be applicable to more general cases. i.e.,

- the boundedness of the parameter estimates was obtained using, essentially, only the fact that $\frac{\hat{x}}{\sqrt{m}}$ is U.B. This part also yielded $\frac{\epsilon^2}{m}$ is integrable and $\dot{\tilde{k}} \in L_2$.

- We were then able to write an integral inequality for $\hat{x}$ (part of m) of the form

$$|\hat{x}|^2 \leq C + \lambda \int_0^t e^{-\alpha(t-\tau)} \mu(\tau)(\hat{x})^2 d\tau$$

Where $\mu(\tau)$ is integrable. Such an inequality is in a suitable form to apply the B-G lemma

since $\mu(\cdot) \in L_2 (\cap L_\infty)$. As a matter of fact the boundedness of $\hat{x}$ and m could be established under weaker conditions, namely

$$\int_{t_0}^{t_0+T} \mu^2(t)dt \leq c + \beta T \quad \forall\, T \geq 0,\ t_0 \geq 0$$

where $\beta < \alpha$.

A physical interpretation of the above proof is that the perturbation $\epsilon + \dot{\tilde{k}} x$ "adds" energy to the m-system which should be dissipated in order to preserve boundedness.

A more general + more compact derivation of this result will be pursued in the following.

SYSTEM IDENTIFICATION + PERSISTENT EXCITATION.*

Consider the linear model

$$y_p = \theta^{*T} w(t) + \varepsilon_t$$

(ε_t is an exponentially decaying term due to initial Conditions)

and the identifier

$$y_i = \theta^T w$$

$\therefore$ Identification error:

$$e_1 = y_i - y_p = \phi^T w + \varepsilon_t$$

where

$$\phi = \theta - \theta^*$$

$w(t) \in \mathbb{R}^{2n}$ available for measurement

* REF Sastry + Bodson: "Adaptive Control: Stability Convergence + Robustness" Prentice Hall, 1989.

Properties of Identification Algorithms.

$$\dot{\phi} = \dot{\theta} = -\gamma \frac{\epsilon_1 w}{m} \qquad \gamma > 0$$

$m = 1$: gradient

$m = 1 + \gamma w^T w$: normalized gradient.

($m = 1 + \|w\|^2_{2,\delta}$: modified normalized gradient)

$$\epsilon_1 = \phi^T w \quad ; \quad w, m \in L^e_\infty.$$

THM : • with the gradient algorithm + w pw cont.

$\Rightarrow$ $e_1 \in L_2$

$\phi \in L_\infty$

• with the normalized gradient + w : pw cont.

$\Rightarrow$ $\frac{e_1}{\sqrt{m}} \in L_2 \cap L_\infty$

$\phi \in L_\infty$, $\dot{\phi} \in L_2 \cap L_\infty$

$$\frac{\phi^T w(t)}{\|\sqrt{m}_t\|_\infty} \in L_2 \cap L_\infty$$

truncation

Effects of initial conditions

When $\epsilon_1 = \phi^T w + \varepsilon_t$; $\epsilon_t \le k \cdot e^{-\alpha t}$ $\alpha > 0$.

then the conclusions of the previous theorem are valid.

Projections

Assume that it is known a priori that

$$\theta^* \in \Theta$$

Θ : closed convex set with smooth boundary $\partial\Theta$

e.g. $\theta_i^* \in [\alpha_{min}(i), a_{max}(i)]$ $i = 1, 2 ..$

or $\Theta = \{\theta \mid (\theta - \theta_c)^T P^{-1} (\theta - \theta_c) \le 1\}$

P : symmetric positive definite matrix

θ_c : known constant vector.

(Generalized ellipsoids with center θ_c radius P.)

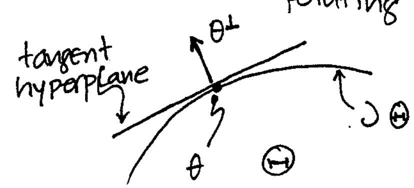
Then, a normalized gradient w/ projection algorithm is defined by:

$$\dot{\theta} = -\gamma \frac{\epsilon_1 w}{m} \qquad \theta \in int(\Theta) \quad (\text{interior of } \Theta)$$

$$= Pr\left[-\gamma \frac{\epsilon_1 w}{m}\right] \quad \text{if } \theta \in \partial\Theta \text{ and } \epsilon_1 w^T \theta^{\perp} \ge 0$$

where Pr : Projection onto the hyperplane tangent to $\partial\Theta$ at θ

$\theta^{\perp}$: unit vector perpendicular to the hyperplane (tangent to Θ at θ) pointing outward



e.g. if $\theta_i^* \in [\theta_{i\,min}^*, \theta_{i\,max}^*]$

then the update law becomes

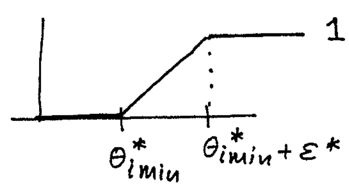
$$\dot{\theta}_i = \begin{cases} -\gamma \dfrac{\epsilon_1 w_i}{m} & \theta_i \in (\theta_{i\,min}^*, \theta_{i\,max}^*) \\ & \theta_i = \theta_{i\,min}^* ; \dot{\theta}_i \ge 0 \text{ or } \theta_i = \theta_{i\,max} ; \dot{\theta}_i \le 0 \\ 0 & \text{if } \theta_i = \theta_{i\,min}^* \text{ and } \dot{\theta}_i < 0 \\ & \text{or } \theta_i = \theta_{i\,max}^* \text{ and } \dot{\theta}_i > 0 \end{cases}$$

The previous law is not Lipschitz continuous and some mild modifications are required to guarantee existence + uniqueness of solutions (smooth projections) e.g.

if $\theta^* \geq \theta^*_{i\,min} + \varepsilon^*$; $\varepsilon^* > 0$

take $\dot{\theta}_i = -\gamma\, \sigma_p(\theta, \epsilon_1 w_i) \frac{\epsilon_1 w_i}{m}$

where σ_p :



$\epsilon_1 w_i > 0$

$\sigma_p = 1$ when $\epsilon_1 w_i < 0$.

When a (smooth) projection is used it can be shown that the derivative of the Lyapunov function at the boundary (boundary region) is $\leq$ to its value with the original ODE.

$\therefore$ the results of the previous theorem are still valid and in addition we have that, starting inside Θ, $\theta \in \Theta$ for all t.

Similar properties can be shown to hold for other estimation algorithms, eg. Least-squares w/ covariance resetting etc.

Assuming that the signals w, $\dot{w}$ are bounded (i.e. input + output of the plant are bounded) we have the following results:

- The estimation error $\epsilon_1 \in L_2 \cap L_\infty$, $\epsilon_1 \to 0$ as $t \to \infty$ and $\phi, \dot{\phi} \in L_\infty$

  Further, $\dot{\phi} \in L_2 \cap L_\infty$ and $\dot{\phi} \to 0$ as $t \to \infty$.

In order to relax the conditions on w we may use the definition of <u>regular</u> signals which avoids certain "pathological" cases (Such a condition is not necessary in discrete-time systems).

Def Let $z : \mathbb{R}_+ \to \mathbb{R}^n$ s.t. $z, \dot{z} \in L_\infty^e$.
z is called regular if there exist $K_1, K_2 \geq 0$ s.t.

$$|\dot{z}(t)| \leq K_1 \|z_t\|_\infty + K_2 \quad \forall\, t \geq 0.$$

(Subscript 't' denotes truncation).

e.g. e^t is regular but $\sin(e^t)$ is not.

Lemma Let $\phi, w : \mathbb{R}_+ \to \mathbb{R}^{2n}$ s.t. $w, \dot{w} \in L_\infty^e$ and $\phi, \dot{\phi} \in L_\infty$. If w is regular and

$$\beta = \frac{\phi^T w}{1 + \|w_t\|_\infty} \in L_2$$

Then, $\beta, \dot{\beta} \in L_\infty$ and $\beta \to 0$ as $t \to \infty$.

Application When w is possibly unbounded, but regular, the relative error $e_1 / 1 + \|w_t\|_\infty$ tends to 0 as $t \to \infty$. This will prove useful in establishing stability of adaptive controllers where w is not known a priori to be bounded

Persistent Excitation + Exponential Parameter Convergence.

The issue of parameter convergence is related to the asymptotic stability of the ODE:

$$\dot{\phi} = -\gamma\, w w^T \phi$$

which is of the form

$$\dot{\phi} = -A(t)\,\phi$$

$A(t)$: Symmetric Positive Semi-definite $\forall\, t \geq 0$.

(Note that when $w(t)$ is a vector, $\text{rank}(A(t)) \leq 1$.

Def Persistency of Excitation

$w : \mathbb{R}_+ \to \mathbb{R}^n$ is persistently Exciting (PE)

if $\exists\ \alpha_1, \alpha_2, \delta > 0$ s.t.

$$\alpha_2 I \geq \int_{t_0}^{t_0+\delta} w(\tau)\, w^T(\tau)\, d\tau \geq \alpha_1 I \quad \forall\, t_0 \geq 0.$$

Although $w(\tau) w^T(\tau)$ is singular $\forall\, \tau$, PE just requires that $w(t)$ "rotates" sufficiently in $\mathbb{R}^n$ so that the integral is uniformly positive definit over <u>any</u> interval of <u>some</u> length δ.

<u>THM</u> PE + ES

Let $w : \mathbb{R}_+ \to \mathbb{R}^n$ be piecewise continuous

If w is PE

Then $\dot{\phi} = -\gamma w w^T \phi \;;\; \gamma > 0$ is globally ES. ▲▲

An interesting proof was given by Anderson (77) noting that PE is a UCO condition on the system

$$\dot{\theta}_* = 0$$

$$y(t) = w^T(t)\, \theta^*(t)$$

In other words PE is an "Identifiability" condition on the above system.

For the proof of the theorem we will use the following Lemmas:

<u>Def</u> Uniform Complete Obrervability UCO

$[C(t), A(t)]$ is UCO if $\exists\, \beta_1, \beta_2, \delta > 0$ st. $\forall\, t_0 \geq 0$

$$\beta_2 I \geq N(t_0, t_0+\delta) \geq \beta_1 I$$

where $N(t_0, t_0+\delta)$ is the so called observability Grammian

$$N(t_0, t_0+\delta) = \int_{t_0}^{t_0+\delta} \Phi^T(\tau, t_0)\, C^T(\tau)\, C(\tau)\, \Phi(\tau, t_0)\, d\tau$$

and $\Phi(\cdot,\cdot)$ is the STM of $\dot{x} = A(t)\, x$ △△

Note that if $(C(t), A(t))$ are UCO and

$$\dot{x} = A(t)\, x$$
$$y = C(t)\, x$$

$x(t_0)$ can be found from the knowledge of $y(t)$, $t \in [t_0, t_0+\delta]$ as

$$x(t_0) = N^{-1}(t_0, t_0+\delta) \int_{t_0}^{t_0+\delta} \Phi^T(\tau, t_0)\, C^T(\tau)\, y(\tau)\, d\tau.$$

Lemma ES of LTV systems

Consider $\dot{x} = A(t)x \quad ; \quad x_0 \in \mathbb{R}^n \quad (*)$

Then the following are equivalent:

a). $x=0$ is an ES equilibrium of $(*)$

b). $\forall\ C(t) \in \mathbb{R}^{m\times n}$ (m arbitrary) s.t.

$[C(t), A(t)]$ is UCO, $\exists$ symmetric $P(t) \in \mathbb{R}^{n\times n}$

and $\gamma_1, \gamma_2 > 0$ s.t.

1). $\gamma_2 I \geq P(t) \geq \gamma_1 I. \qquad (\#)$

2. $-\dot{P}(t) = A^T(t)P(t) + P(t)A(t) + C^T(t)C(t) \qquad (\#\#)$

c) For some $C(t) \in \mathbb{R}^{m\times n}$ s.t. $[C(t), A(t)]$ is UCO

$\exists\ P(t) \in \mathbb{R}^{n\times n}$ symmetric and $\gamma_1, \gamma_2 > 0$

s.t. $(\#), (\#\#)$ are satisfied.

LEMMA UCO under output Injection

Assume that $\forall\ \delta > 0\ \exists\ K_\delta \geq 0$ s.t.

$$\int_{t_0}^{t_0+\delta} \|K(\tau)\|^2 d\tau \leq K_\delta \qquad \forall\ t_0 \geq 0$$

Then $[C, A]$ is UCO iff $[C, A+KC]$ is UCO.

Moreover if the observability grammian of $[C, A]$ satisfies

$$\beta_2 I \geq N(t_0, t_0+\delta) \geq \beta_1 I$$

Then the observability grammian of $[C, A+KC]$ satisfies

$$\beta_2' I \geq N'(t_0, t_0+\delta) \geq \beta_1' I \qquad (\text{same } \delta)$$

where $\beta_1' = \beta_1 / (1 + \sqrt{K_\delta \beta_2})^2$

$\beta_2' = \beta_2 \exp[K_\delta \beta_2]$.

LEMMA (ES) Suppose $\dot{x} = f(t,x)$; x_0

and there exist $V(t,x)$, and $a_1, a_2, a_3, \delta > 0$

s.t. $\forall x \in B$, $t \geq 0$

$$\alpha_1 |x|^2 \leq V(t,x) \leq \alpha_2 |x|^2$$

$$\dot{V}(t,x) \leq 0 \quad \text{along the trajectories of } \dot{x} = f(t,x)$$

$$\int_t^{t+\delta} \dot{V}(\tau, x(\tau))\, d\tau \leq -a_3 |x(t)|^2$$

Then $x(t)$ converges to 0 exponentially.

<u>Proof of the PE+ES thm.</u>

Let $V = \phi^T\phi$. Hence $\dot{V} = -2\gamma(w^T\phi)^2 \leq 0$ along the trajectories of the ODE $\dot{\phi} = -\gamma w w^T \phi$.

Then,

$$\int_{t_0}^{t_0+\delta} \dot{V}\, d\tau = -2\gamma \int_{t_0}^{t_0+\delta} [w^T \phi(\tau)]^2\, d\tau \quad \forall\, t_0 \geq 0$$

By PE, $[w^T, 0]$ is UCO $\therefore$ under output injection, $(K = -\gamma w)$ the system becomes, $[w^T, -\gamma w w^T]$ with

$$K_\delta = \int_{t_0}^{t_0+\delta} |\gamma w(\tau)|^2\, d\tau = \gamma^2 \operatorname{trace}\left\{\int_{t_0}^{t_0+\delta} w(\tau) w^T(\tau)\, d\tau\right\}$$

$$\leq n\gamma^2 \beta_2 \qquad (n = \dim(w))$$

Using the previous lemma (UCO) $[w^T, -\gamma w w^T]$ is UCO

$\therefore$ Using $\phi(\tau) = \Phi_{ww^T}(\tau, t_0)\phi(t_0)$ we get.

$$\int_{t_0}^{t} \dot{V}\, d\tau \leq -\frac{2\gamma\beta_1}{(1+\sqrt{n}\,\gamma\beta_2)^2} |\phi(t_0)|^2$$

Exp. Convergence follows from the lemma (ES).

THM EXPONENTIAL CONVERGENCE OF THE IDENTIFIER :

For the previous Identification problem +

assuming u & y are bounded ($w, \dot{w}$ UB)

If w is PE
Then the identifier parameter θ converges to the nominal parameter θ^* ($\phi \to 0$) exponentially fast.

(gradient, normalized gradient, LS/Covariance resetting)

▲▲

REMARKS

1) Exponential Convergence Rates. : Can be found from the results in the proof of previous thm – e.g. for the standard gradient algorithm, $\phi \le k e^{-\alpha t}$ with

$$\alpha = \frac{1}{2\delta} \ln \left\{ \frac{1}{1 - \frac{2\gamma a_1}{(1+\sqrt{n}\,\gamma a_2)^2}} \right\}$$

γ : adaptive gain ; a_1, a_2, δ : as in PE definition

n : Number of adjustable parameters

When γ, reference input u are small, rate of conv. $\to \gamma \alpha_1/\delta$. ; rate of conv. $\propto$ (amplitude of ref. inp)^2

However, large adaptive gains + reference inputs will 'saturate' the convergence rate which may even decrease.

Furthermore, the rate of convergence depends on a complex manner on the input signal & the plant to be identified via a_1, a_2, δ.

What is particularly hard is to establish PE based on conditions on the input signal instead of w. For this, it is necessary that the plant, parametrized by $y = \theta^{*T} w$, is minimal so that the number of parameters to be identified is the minimum required.

(see also discussion below).

- Initial Conditions in the plant do not affect the exponential stability of the identifier. They do however affect the rate of convergence if the rate of decay of the I.C. transients is 'slower' than the rate of convergence of the algorithm (Kreisselmeier).

- ES of the Identifier also guarantees some robustness properties w.r.t. disturbances. A typical result (Narendra + Annaswamy) is that a gradient scheme without modifications will guarantee boundedness provided that the level of PE is large compared to the size of the disturbance (in MRAC, needs Relative degree 1 plants + SPR reference model).

In general, a <u>Local</u> robustness result can be obtained through Malkin's theorem:

<u>THM</u>: Let $\dot{x} = f(x,t)$ with equilibrium $x=0$ and assume that 0 is ES. Consider

$$\dot{x} = f(x,t) + g(x,t) \qquad (*)$$

s.t. $\|g(x,t)\| \leq b\|x\|$ whenever $\|x\| < \delta$ for some $\delta > 0$. Then 0 of $(*)$ is ES.

This is not very practical however since it requires the perturbation $g(x,t)$ to be 0 at the origin. A weaker result is given via the definition of total stability:

<u>DEF</u> Equil $x=0$ of $\dot{x} = f(x,t)$ is <u>totally stable</u> if $\forall \varepsilon > 0$, $\exists\ \delta_1(\varepsilon), \delta_2(\varepsilon) > 0$ s.t. every solution $x(t; x_0, t_0)$ of $(*)$ satisfies $\|x(t; x_0, t_0)\| < \varepsilon \quad \forall\ t \geq t_0$ provided that $\|x_0\| < \delta_1$, $\|g(x,t)\| < \delta_2$.

<u>THM</u> (MALKIN) If the equilibrium state of $\dot{x} = f(x,t)$ is U.A.S. then it is totally stable

- The ES property of an identifier can be shown to hold in the case the parameter update is of the form

$$\dot{\phi} = -\gamma\, w \cdot W_M(s)[\phi^T w] = -\gamma e_1 w.$$

provided that $W_M(s)$ is SPR. Although $\dot{\phi}$ is not given by a 'true' gradient algorithm, using the properties of SPR transfer functions, it can be shown that $e_1 = W_M(s)[\phi^T w] \in L_2$

$$e_1, \phi \in L_\infty$$

Also, when e_1 is taken as

$$e_1 = W_M(s)\,[\phi^T w - \bar{\gamma} w^T w e_1] \qquad \bar{\gamma} > 0$$

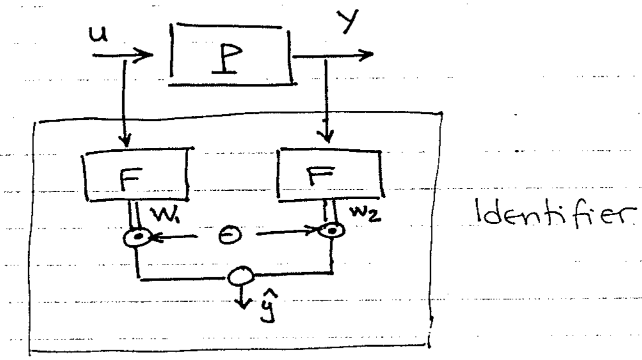
and $\dot{\phi} = -\gamma e_1 w$ then $W_M(s)$ SPR $\Rightarrow$

$$e_1, \dot{\phi} \in L_2, \quad e_1, \phi \in L_\infty.$$

Further ES stability of the identifier is guaranteed (for both schemes) provided that w is PE and $w, \dot{w} \in L_\infty$.

These results are useful in MRAC and especially in the relative degree 1 case where $W_M(s)$ can be selected SPR and $y_P - y_M = e_1 = W_M(s)[\phi^T w]$

CONDITIONS ON THE REFERENCE INPUT



Problem: what are the conditions on u wich guarantee that w $(=(w_1, w_2))$ is PE ?

1) w can be viewed as the output of a linear system with input u:

$$w = H_{wu}(s)\, u = \begin{bmatrix} (sI-F)^{-1} q \\ (sI-F)^{-1} q\, P(s) \end{bmatrix} u$$

F, q auxiliary filters

$P(s)$ plant.

<u>Lemma:</u> Let w be stationary and R be its autocovariance $R_w(t) = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} w(\tau)\, w^T(t+\tau)\, d\tau$

Then w is PE iff $R_w(0) > 0$.

Note that in frequency domain $R_w(0) = \int S_w(d\omega)$

S_w : spectral measure of w.

$\therefore \quad R_w(0) = \int H_{wu}^*(j\omega) H_{wu}^T(j\omega) S_u(d\omega)$

<u>DEF</u>. A stationary signal $u : \mathbb{R}_+ \to \mathbb{R}$ is <u>sufficiently rich</u> of order K if the support of the spectral density of u — $S_u(d\omega)$ — contains at least K points.

eg. A single sinusoid contributes 2 points to the spectrum: $+\omega_0, -\omega_0$

DC contributes 1 point.

<u>THM</u> PE + Sufficient Richness.

Let $w(t) \in \mathbb{R}^n$ be the output of a stable LTI system with transfer function $H_{wu}(s)$ and stationary input u. Assume that $H_{wu}(j\omega_1) \dots H_{wu}(j\omega_n)$ are linearly independent on $\mathbb{C}^{n \times n}$ $\forall\, \omega_1, \dots, \omega_n \in \mathbb{R}$

Then w is PE iff r is suff. rich of order n

The net result of this analysis can be described as follows:

The exponential convergence of the parameter error to zero is guaranteed provided that the plant of order n is minimal and the reference input u contains a sufficient number of sinusoids of different frequencies. At least n frequencies are needed to identify $2n$ parameters. These frequencies must not be zeros of $H_{wu}(j\omega)$.

Furthermore, "sufficient" separation between the input frequencies is necessary in order to guarantee a "not-arbitrarily-small" rate of exponential convergence

Finally, if u is suff. rich of order $k < 2n$ will converge to the subspace : $R_w(0)[\theta - \theta^*(t)] = 0$ (not necessarily to a constant)

MODEL REFERENCE ADAPTIVE CONTROL

1). PLANT ASSUMPTIONS

$$y_p = k_p \frac{N_p(s)}{D_p(s)} u_p = P_o(s)\, u_p$$

(SISO - LTI) D_p, N_p are monic coprime polynomials of degree n, m respectively (n, m known)

$P_o(s)$ is strictly proper

$N_p(s)$ is Hurwitz (Minimum phase assumption)

The sign of the high frequency gain k_p is known. W.o.l.o.g. $k_p > 0$.

2) REFERENCE MODEL ASSUMPTIONS.

$$y_M = W_M(s)\, r = k_m \frac{N_M}{D_M} r.$$

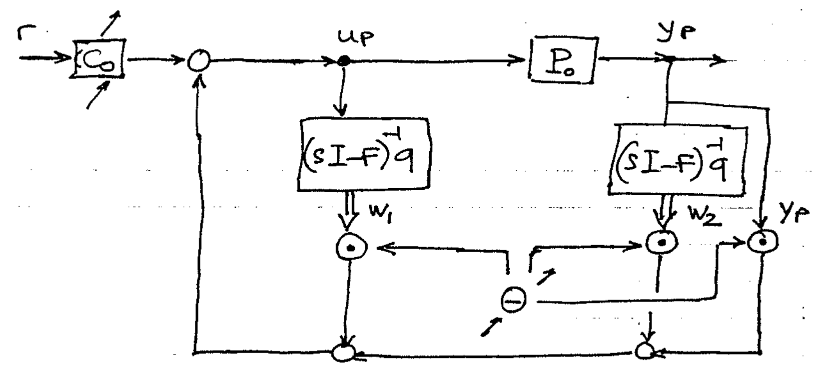
D_M, N_m are monic, coprime Hurwitz polynomials of degree n_m, m_m respectively ; $n_m \leq n$

$$n_m - m_m = n - m$$

$k_m > 0$.

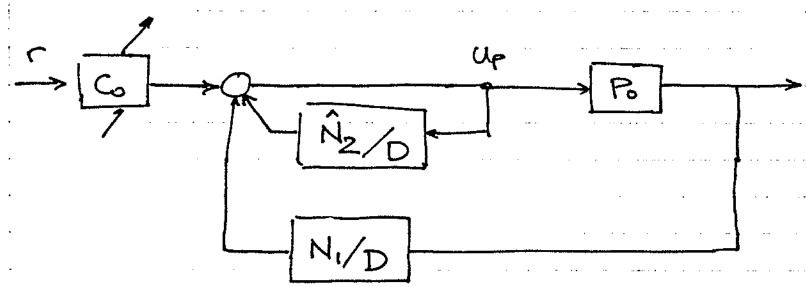
3) REFERENCE INPUT : r : PW continuous UB.

CONTROLLER STRUCTURE



$F \in \mathbb{R}^{n-1 \times n-1}$, (F, q) cc

OR



where $\hat{N}_2$, N_1 are polynomias of degree $n-2$, $n-1$ respectively and coefficients completely determined by θ (and vice-versa)

$D(s)$: Hurwitz, of degree $n-1$: $D(s) = \det(sI-F)$

REMARK

$F_1(D(s))$ should be selected s.t. $N_m(s)$ is a factor of $D(s)$. $\square$

Then,

$$u_p = c_0 r + \frac{\hat{N}_2}{D} u_p + \frac{N_1}{D} y_p$$

$$= \frac{D}{N_2}\left[c_0 r + \frac{N_1}{D} y_p\right] \qquad \ni \; N_2 = D - \hat{N}_2$$

$$\therefore \; y_p = \frac{c_0 k_p D N_p}{N_2 D_p - N_1 k_p N_p} r$$

And therefore $\exists$ unique N_2 ($\Rightarrow \hat{N}_2$), N_1, c_0 s.t. $y_p = W_m(s) r$.

(see previous lectures)

$\therefore$ $\exists$ unique set of controller parameters $\theta \in \mathbb{R}^{2n-1}$ and c_0 s.t. $y_p = W_m(s) r$.

Furthermore this controller guarantees internal / exponential stability of the closed loop.

REMARK

Note that a lot of cancellations occur in MRC. : The plant + controller have $3n-2$ states (modes) $n-1$ of them are cancelled in the controller $\left(\frac{D}{N_2} \cdot \frac{N_1}{D}\right)$. $(n-m)$ correspond to the cancellation of D/N_m by the c-l. denom. in (y_p/r). And m correspond to the cancellation of N_p in the t.f. (y_p/r). Of course, all of the cancelled modes are stable either by assumption (N_p) or by design (D). $\square$

Another representation of the controller in state space form is

$$u_p = c_0 r + \theta^T w$$

where $w = \begin{pmatrix} w_1 \\ w_2 \\ y_p \end{pmatrix}$

Note that this expression is <u>linear</u> in the (usually unknown) controller parameters.

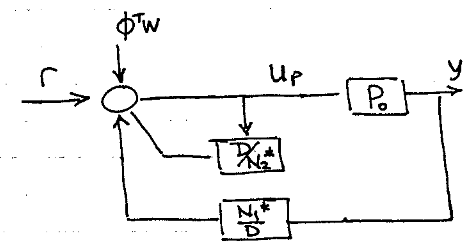
and ω, r are signals available for measurement.
(The nominal or desired controller parameters for which the MRC objective is achieved will be denoted by $()^*$ ie. θ^*, c_0^*.)

Furthermore, in this formulation, the controller contains a direct throughput $y_p \to u_p$

A strictly proper controller $(y_p \to u)$ can be obtained by using n-th order filters F and letting $\omega = \begin{pmatrix} (sI-F)^{-1} q\, u \\ (sI-F)^{-1} q\, y \end{pmatrix}$. Such a controller may have improved 'robustness' properties, in terms of high frequency noise or unmodeled dynamics, but it does require the update of an additional parameter.

<u>Case 1</u> K_p = Known ; $c_0^* = 1$ w.o.l.o.g.

In this case the MRAC loop can be written as



where $()^*$ denotes the nominal polynomials for the compensator and $\phi = \theta - \theta^*$.

Let S_u denote the transfer function $r \to u_p$ for the nominal closed loop.

Then,

$$\begin{pmatrix} u_p \\ y_p \end{pmatrix} = \begin{bmatrix} S_u \\ W_M \end{bmatrix} (r + \phi^T w) + \varepsilon_t$$

ε_t : Exp. Decaying terms

S_u, W_M : ES

Parameter Update law

We can construct the signal

$$e_1 = y_P - y_M = W_M(\phi^T \omega).$$

Let $\epsilon_1 = e_1 + \theta^T W_M[\omega] - W_M[\theta^T \omega]$

$$= e_1 + \phi^T \zeta - W_M(\phi^T \omega)$$

$$= \phi^T \zeta \qquad ; \quad \zeta = W_M\, \omega .$$

(AUGMENTED ERROR)

ϵ_1 can be constructed from known signals.

∴ The unknown controller parameters satisfy the linear model equation

$$\epsilon_1 \neq \phi^T \zeta \qquad (+\ \varepsilon_t)$$

∴ they can be directly estimated by:

$$\dot{\theta} = \dot{\phi} = -\gamma\, \epsilon_1 \zeta / m$$

m : Normalizing signal : $1 + \zeta^T \zeta \left[e^{-\delta_0 t} \left\| \begin{pmatrix} u_P \\ q_m y_{Pt} \end{pmatrix} \right\|_{2,\delta_0} + q_\varepsilon \right]^2$

$\delta_0, q_m, q_\varepsilon > 0.$

From the adaptive law we have:

$$V = \frac{1}{2\gamma} \phi^T \phi$$

$$\dot{V} = -\epsilon_1 \phi^T \zeta / m = -\epsilon_1^2 / m \leq 0$$

∴ V is UB, $\epsilon_1 / \sqrt{m} \in L_2$

Since $\|\zeta\|^2 / m$ is UB * and ϕ is UB

ϵ_1^2 / m is UB ∴ $\dot{\phi}$ is UB. and $\|\dot{\phi}\| \in L_2$

---

* ζ can be written as $H_z(s)[u_P\, y_P]$ where $H_z(s)$ is a strictly proper, known transfer function (matrix) which depends on $W_m(s)$ and (F, q). Using the properties of $\|\cdot\|_{2,\delta}$ $\frac{\|\zeta\|^2}{m}$ will be UB provided that $g_{z,\delta_0}[H_z] \leq c < \infty$ i.e. δ_0 must be chosen so that $\mathrm{Re}[\lambda_i[F]]$ and $\mathrm{Re}[\mathrm{poles}(W_M)] < -\delta_0$

Rem. Similar properties can be obtained if the adaptive law is $\dot{\phi} = -\gamma\, W_M[\phi^T \omega]\, w$ and $W_M(s)$ is SPR. For this however we must require that the plant has rel. degree $= 1$ $(n-m = n_m - m_m)$. This assumption, although it is met in several applications, it is quite restrictive. The price paid to remove the SPR condition, is the use of the "augmented" error signal ϵ_1 and the auxiliary vector ζ which increase the dimensionality of the controller.

The properties of SPR functions can be used to produce a variety of adaptive laws via different constructions of the augmented error: For example a general adaptive law would be $\dot{\phi} = -\gamma\, W_L[\epsilon_1'] \cdot \zeta'$ where W_L is SPR and $\epsilon_1' = \phi\, \zeta'$. (In our case $W_L = 1$).

It can be argued that the extra degrees of freedom offered by W_L can be used to improve

the robustness properties of the adaptive controller and its behavior in the presence of external noise. (The issue however is still unclear).

For Details: * Narendra + Valavani, IEEE AC, Aug. 1978 AND JUNE 1980

OR. * Narendra + Annaswamy: Stable Adaptive Systems, Prentice Hall 1989.

<u>AUGMENTED ERROR CONCEPT</u>: MONOPOLI, IEEE AC, Oct. 74.

In order to establish the stability properties of the MRAC we need to derive some properties for the signal $\phi^T w$ which "perturbs" the nominal closed loop $\begin{bmatrix} u_p \\ y_p \end{bmatrix} = \begin{bmatrix} S_u \\ W_M \end{bmatrix} [r + \phi^T w]$

(Note that the adaptive law provides information about $\phi^T \zeta$ <u>only</u>).

This has been - traditionally - the difficult step in analyzing the properties of a Direct MRAC scheme. Several approaches have been developed, e.g. Narendra + Valavani + Lin (1980) Kreisselmeier + Narendra IEEE AC Dec 1982 etc.

We will proceed in a somewhat different way.

Consider the operator identity:

$$[1-\Lambda(s)] + \Lambda(s) = 1$$

Λ is stable, minimum phase, with rel. degree $\geq n-m$ and DC gain $= 1$.

E.g. $\Lambda(s) = \dfrac{a^k}{(s+a)^k}$; $k \geq n-m$, $a > 0$

Then,

$$\phi^T w = \underbrace{[1-\Lambda(s)]}_{\Lambda_1(s)}\ \underbrace{\phi^T w}_{(\dot{\phi^T w})} + \Lambda(s)\, W_M^{-1}(s)\, W_M(s)\, \phi^T w.$$

Since $\Lambda(s)$ has DC gain 1 we can write

$$\Lambda_1(s) = \frac{1-\Lambda(s)}{s} = \frac{N_\Lambda(s)}{D_\Lambda(s)}$$

where $\partial(N_\Lambda) \leq \partial(D_\Lambda) - 1$

Furthermore, from the operator identity

$$s\,\phi(t) = \phi(t)\, s + \dot{\phi}(t)$$

we can express $W_M(s)[\phi^T w]$ as

$$W_M(s)[\phi^T w] = \phi^T W_M(s)[w] + W_{M1}(s)\left\{ \left(W_{M2}(s)[w]\right)^T \dot{\phi} \right\}$$

where the poles of $W_{M1}(s)$, $W_{M2}(s)$ belong to the set of poles of $W_M(s)$.

E.g. Let $W_M(s) = \dfrac{1}{s+K}$. Then

$$\frac{1}{s+K}[\phi^T w] = \frac{1}{s+K}\phi^T\left[\frac{s+K}{s+K}\right]w = \frac{1}{s+K}\, s\,\phi^T \frac{1}{s+K} w - \frac{1}{s+K}\dot{\phi}^T \frac{1}{s+K} w + \frac{1}{s+K} K \phi^T \frac{1}{s+K} w$$

$$= \underbrace{\frac{s+K}{s+K}\left(\phi^T \frac{1}{s+K} w\right)}_{\phi^T W_M[w]} - \underbrace{\frac{1}{s+K}}_{W_{M1}}\left[\dot{\phi}^T \underbrace{\frac{1}{s+K}}_{W_{M2}} w\right]$$

Thus,

$$\phi^T\omega = \Lambda_1(s)\left\{\dot{\phi}^T\omega + \phi^T\dot{\omega}\right\} + \Lambda W_M^{-1}(s)\left[\phi^T Z\right] + \Lambda W_M^{-1} W_{M1}(s)\left\{\left(W_{M2}(s)[\omega]\right)^T \dot{\phi}\right\} + \varepsilon_t$$

Let $\quad \sqcup = \begin{bmatrix} u_p \\ q_m y_p \end{bmatrix}$

$$G(s) = (sI - F)^{-1} q \; ; \quad \hat{G}(s) = \begin{bmatrix} G(s) & 0 \\ 0 & G(s) \end{bmatrix}$$

$$Q_m = \begin{pmatrix} 1 & 0 \\ 0 & q_m \end{pmatrix} \; ; \quad H(s) = \begin{bmatrix} S_u(s) \\ q_m W_M(s) \end{bmatrix}$$

Then:

$$\sqcup = H(s)\left[\phi^T\omega + r\right] + \varepsilon_t$$

$$\omega = \bar{G}(s)\, Q_m^{-1} \sqcup + \varepsilon_t$$

where $\bar{G}(s) = \begin{bmatrix} \hat{G}(s) \\ 0 \; 1 \end{bmatrix}$ if y_p is included in ω

and $\bar{G}(s) = \hat{G}(s)$ if a 'strictly proper' controller is used.

Further, Let $m_f(t) = \left[\varepsilon_{-\delta} \|\sqcup_t\|_{2,\delta} + q_f\right]^2$

$$= \left[\left(\int_0^t e^{-2\delta(t-\tau)} \sqcup^T \sqcup\right)^{1/2} + q_f\right]^2$$

and denote by C_w the constant s.t.

$$\|\omega\|^2 \le C_w^2\, m_f + C_R + \varepsilon_t \qquad C_R > 0$$

- $\|\omega\|^2 \le \|\hat{G} Q_m^{-1} \sqcup + \varepsilon_t|^2 + |W_M(\phi^T w + r) + \varepsilon_t|^2$

1). $|\hat{G}(s) Q_m^{-1} \sqcup|^2 \le g_{2\delta}^2(\hat{G}(s) Q_m^{-1})\, m_f$

due to y_p throughput $\leftarrow$

2) $\|W_M(\phi^T\omega)\| \le C_\phi^2 \cdot g_{2\delta}^2(W_M(s)) \cdot \gamma_{2\delta}^2(\bar{G}(s) Q_m^{-1}) \cdot m_f$

3) $|W_m[r]|^2 \le \gamma_\infty^2(W_m) \cdot \|r\|_\infty^2$.

4). CAUCHY's Inequality: $2|ab| \le \varepsilon |a|^2 + \frac{1}{\varepsilon}|b|^2 \quad \forall \varepsilon > 0$

$$\Rightarrow \quad C_w^2 = g_{2\delta}^2(\hat{G}(s) Q_m^{-1}) + C_\phi^2\, g_{2\delta}(W_M(s))\, \gamma_{2\delta}^2(\bar{G}(s) Q_m^{-1}) + \varepsilon$$

ε : arbitrarily small ; C_ϕ : bound of ϕ.

Next, from the closed loop equation:

$$\sqcup = H(s)\left[r + \phi^T\omega\right] + \varepsilon_t$$

$$= H(s)(\phi^T\omega) + R + \varepsilon_t$$

$$\sqcup = H(s)\Big\{ \Lambda_1(s)(\mathring{\phi}\omega) + \Lambda_1(s)(\phi\mathring{\omega}) + \Lambda W_M^{-1}(s)(\phi^T\zeta_1)$$

$$+ \Lambda W_M^{-1} W_{M1}(s)\left[(W_{M2}(s)[\omega])^T \dot{\phi}\right]\Big\}$$

$$+ R + \varepsilon_t$$

R: the (bounded) effect of r on $\sqcup$ via $H(s)$

ε_t: Exp. decaying terms due to I.C. $\leq c e^{-\alpha t}$

α: "stability margin" of closed loop system.

Let $\delta < \alpha$ and take $2,\delta$ norms of truncated signals

$$\|\sqcup_t\|_{2\delta} \leq \gamma_{2\delta}(H\Lambda_1)\,\|(\mathring{\phi}\omega)_t\|_{2\delta}$$

$$+ \gamma_{2\delta}(H\Lambda_1)\,\|(\phi\mathring{\omega})_t\|_{2\delta}$$

$$+ \gamma_{2\delta}(H\Lambda W_M^{-1})\,\|(\phi^T\zeta_1)_t\|_{2\delta}$$

$$+ \gamma_{2\delta}(H\Lambda W_M^{-1}W_{M1})\,\|(W_{M2}(s)[\omega])^T\dot{\phi}_t\|_{2\delta}$$

$$+ \|R_t\|_{2\delta} + c$$

1) $$\|(\mathring{\phi}\,\omega)_t\|_{2\delta} \leq \left\{\int_0^t e^{2\delta\tau}\|\mathring{\phi}(\tau)\|^2 \cdot \|\omega(\tau)\|^2 d\tau\right\}^{1/2}$$

$$\leq \left\{\int_0^t e^{2\delta\tau}\|\mathring{\phi}(\tau)\|^2 \cdot \left[C_\omega^2 m_f + C_R + \varepsilon_t\right](\tau)\, d\tau\right\}^{1/2}$$

$$\leq \Big\{C_\omega^2\int_0^t e^{2\delta\tau}\|\mathring{\phi}(\tau)\|^2 m_f(\tau)\, d\tau$$

$$+ C_R\int_0^t e^{2\delta\tau}\|\dot{\phi}(\tau)\|^2 d\tau + c\Big\}^{1/2}$$

$$\leq \left\{C_\omega^2\,\|[\|\mathring{\phi}\|\sqrt{m_f}]_t\|_{2\delta}^2 + \underbrace{C_R\,\|\dot{\phi}_t\|_{2\delta}^2}_{\leq C_R' \cdot e^{2\delta t}} + c\right\}^{1/2}$$

2) $$\|(\phi\mathring{\omega})_t\|_{2\delta} \leq \left\{\int_0^t e^{2\delta\tau}\|\phi(\tau)\|^2 \cdot \|\mathring{\omega}(\tau)\|^2 d\tau\right\}^{1/2}$$

$$\leq C_\phi\left\{\int_0^t e^{2\delta\tau}\|\mathring{\omega}(\tau)\|^2 d\tau\right\}^{1/2} = C_\phi\,\|\mathring{\omega}_t\|_{2\delta}$$

$$\mathring{\omega}(\tau) = s\,\omega(\tau) = \begin{bmatrix} s\hat{G}Q_m^{-1}\sqcup \\ s\,W_M\,[\phi^T\omega + r] \end{bmatrix} + \varepsilon_t = \begin{bmatrix} s\hat{G}Q_m^{-1}\sqcup \\ s\,W_M[\phi^T\omega] + sW_M[r] \end{bmatrix} + \varepsilon_t$$

$$\|\mathring{\omega}_t\|_{2\delta} \leq \left\{\gamma_{2\delta}^2[sGQ_m^{-1}]\,\|\sqcup_t\|_{2\delta}^2 + \left[\gamma_{2\delta}[sW_M]\|(\phi^T\omega)_t\|_{2\delta} + \|(sW_M[r])_t\|_{2\delta}\right]^2\right\}^{1/2} + c$$

$$\| \phi^T \omega_t \|_{2\delta} = \left\{ \int_0^t e^{2\delta\tau} \left(\phi^T \omega(\tau)\right)^2 d\tau \right\}^{1/2}$$

$$\leq C_\phi \cdot \| \omega_t \|_{2\delta} \leq C_\phi \cdot \gamma_{2\delta}\left(\bar{G} Q_m^{-1}\right) \| U_t \|_{2\delta} + C$$

$$\therefore \| \mathring{\omega}_t \|_{2\delta} \leq \left\{ \gamma_{2\delta}^2 \left[s \hat{G} Q_m^{-1} \right] \| U_t \|_{2\delta}^2 \right.$$

due to y_p throughput. $\longleftarrow$

$$\left. + \left[C_\phi \gamma_{2\delta} [s W_M] \, \gamma_{2\delta} \left[\bar{G} Q_m^{-1} \right] \| U_t \|_{2\delta} + C + \| R_t' \|_{2\delta} \right]^2 \right\}^{1/2} + C$$

Using Caucy's inequality again we have that

$$\forall \, \varepsilon > 0 \qquad (a+b)^2 \leq (1+\varepsilon) a^2 + \left(1+\tfrac{1}{\varepsilon}\right) b^2$$

$$\therefore \| \mathring{\omega}_t \|_{2\delta} \leq \left\{ \left[\gamma_{2\delta}^2 \left(s \hat{G} Q_m^{-1} \right) + C_\phi^2 \gamma_{2\delta}^2 (s W_m) \, \gamma_{2\delta}^2 \left[\bar{G} Q_m^{-1} \right] (1+\varepsilon) \right] \cdot \| U_t \|_{2\delta}^2 \right.$$

$$\left. + \; C_\varepsilon^2 + \| (R_\varepsilon)_t \|_{2\delta}^2 \right\}^{1/2} + C$$

$$\leq \left\{ C_{\mathring{\omega}}^2 \, \| U_t \|_{2\delta}^2 + C_\varepsilon^2 + \| (R_\varepsilon)_t \|_{2\delta}^2 \right\}^{1/2} + C$$

where C_e^2, C are some constants due to I.C.
R_ϵ is a signal related to r
and $C_{\mathring{\omega}}^2$ is arbitrarily close to
$\gamma_{2\delta}^2 (s \hat{G} Q_m^{-1}) + C_\phi^2 \, \gamma_{2\delta}^2 (s W_M) \, \gamma_{2\delta}^2 (\bar{G} Q_m^{-1})$

Finally,

$$\| (\phi \mathring{\omega})_t \|_{2\delta} \leq C_\phi \left\{ C_{\mathring{\omega}}^2 \, \| U_t \|_{2\delta}^2 + C_\varepsilon^2 + R_\varepsilon^2 e^{2\delta t} \right\}^{1/2} + C$$

3) $\| W_{M2}(s) [\omega]^T \mathring{\phi}_t \|_{2\delta}$: (Note: $W_{M2}(s)$ is strictly proper)

Let C_m^2 : $[W_{m2}(\omega)|_{(t)}^2 \leq C_M \; m_f + \varepsilon_Z$

i.e. $C_m^2 = g_{2\delta}^2 \left(W_{M2} \bar{G} Q_m^{-1} \right) + \varepsilon$

Then $\| W_{M2}(s) [\omega]^T \mathring{\phi}_t \|_{2\delta} \leq \left\{ C_M^2 \, \| (\mathring{\phi} \sqrt{m_f})_t \|_{2\delta}^2 + C \right\}^{1/2}$

$(C > 0)$

<u>Comment</u> Despite their 'dreadful' appearance these bounds are actually quite nice. They indicate that $\| U_t \|_{2\delta}$ is "bounded" in terms of $\| (\mathring{\phi} \sqrt{m_f})_t \|_{2\delta}$, $A_1 \| \phi Z_t \|_{2\delta}$ and $A_2 \| U_t \|_{2\delta}$ where A_1 is $O(a^{n-m})$ and A_2 is $O(\frac{1}{a})$ 'a' being the arbitrary constant used in the ficticious filter $\Lambda(s)$.

Choosing 'a' sufficiently large we can obtain a bound on $\|U_t\|_{2\delta}$ using $\|(\dot{\phi}\sqrt{m_f})_t\|_{2\delta}$ and $A_1\|(\phi\xi_1)_t\|_{2\delta}$ which, in turn, will be in a convenient form to apply the Bellman Gronwall lemma since $\dot{\phi}$ and $\phi\xi/\sqrt{m_f} = \phi\xi/\sqrt{m}\cdot\sqrt{m/m_f}$ are L_2 (or small in the mean) provided of course that m/m_f is UB.

All that remains now is some algebraic calculations

Substituting the previous expressions in the bound for $\|U_t\|_{2\delta}$ and letting

$$\Gamma_0 = \gamma_{2\delta}(H\Lambda_1)$$

$$\Gamma_1 = \gamma_{2\delta}(H\Lambda W_M^{-1})$$

$$\Gamma_2 = \gamma_{2\delta}(H\Lambda W_M^{-1}W_{M1})$$

we find

$$\|U_t\|_{2\delta} \leq \Gamma_0\left\{C_W^2\,\|(|\dot{\phi}|\sqrt{m_f})_t\|_{2\delta}^2 + C_R' e^{2\delta t} + c\right\}^{1/2}$$

$$+\Gamma_0 C_\phi\left\{C_{\dot{\omega}}^2\,\|U_t\|_{2\delta}^2 + R_\epsilon^2 e^{2\delta t} + C_e\right\}^{1/2}$$

$$+\Gamma_1\,\|(\phi^T\xi_1)_t\|_{2\delta}$$

$$+\Gamma_2\cdot\left\{C_M^2\,\|(|\dot{\phi}|\sqrt{m_f})_t\|_{2\delta}^2 + c\right\}^{1/2}$$

$$+\|R_t\|_{2\delta} + C'$$

or, using $\sqrt{A+B} \leq \sqrt{A}+\sqrt{B}$; $A, B > 0$ and combining the constants for simplicity:

$$\|U_t\|_{2\delta} \leq (\Gamma_0 C_W + \Gamma_2 C_M)\,\|(|\dot{\phi}|\sqrt{m_f})_t\|_{2\delta}$$

$$+\Gamma_1\,\|(\phi^T\xi)_t\|_{2\delta}$$

$$+\Gamma_0 C_\phi C_{\dot{\omega}}\,\|U_t\|_{2\delta}^2$$

$$+C_R e^{\delta t} + C$$

Observe that $\Gamma_0 = O(\gamma_{2\delta}(\Lambda_1)) = O(\frac{1}{a})$

Hence, imposing the constraint $\Gamma_0 C_\phi C_{\dot{\omega}} < 1$ (i.e. a suff. large) we obtain

$$\|U_t\|_{2\delta} \le \frac{\Gamma_{\dot\phi}}{\Gamma_0'} \|(\dot\phi|\sqrt{m_f})_t\|_{2\delta}$$
$$+ \frac{\Gamma_1}{\Gamma_0'} \|(\phi^T\zeta)_t\|_{2\delta}$$
$$+ C_R e^{\delta t} + C$$

where $\Gamma_0' = 1 - \Gamma_0 C_\phi C_{\dot\omega} > 0$

$\Gamma_{\dot\phi} = \Gamma_0 C_\omega + \Gamma_2 C_m$

and C_R, C are modified accordingly.

and thus,

$$\|U_t\|_{2\delta} + q_f e^{\delta t} = \left[e^{2\delta t} m_f(t)\right]^{1/2} \le$$

$$\frac{\Gamma_{\dot\phi}}{\Gamma_0'} \|(|\dot\phi|\sqrt{m_f})_t\|_{2\delta} + \frac{\Gamma_1}{\Gamma_0'} \|(\phi^T\zeta)_t\|_{2\delta}$$
$$+ (C_R + q_f) e^{\delta t} + C$$

Use, again, Cauchy's inequality to square both sides. Note that $(A+B)^2 \le (1+\varepsilon)A^2 + (1+\frac{1}{\varepsilon})B^2$ and therefore if A represents a signal of interest e.g. $\|\phi^T\zeta\|_{2\delta}$ and B is a constant, ε can be taken

arbitrarily small.

Thus

$$e^{2\delta t} m_f(t) \le (1+\varepsilon)\left\{ P_1(q)\left(\frac{\Gamma_{\dot\phi}}{\Gamma_0'}\right)^2 \|(\dot\phi\sqrt{m_f})_t\|_{2\delta}^2 + P_2(q)\left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \|\phi\zeta_t\|_{2\delta}^2\right.$$
$$+ (1+\tfrac{1}{\varepsilon})\left[(C_R+q_f)e^{2\delta t} + C^2 + 2C(C_R+q_f)e^{\delta t}\right]$$
$$\le (1+\varepsilon)\int_0^t e^{2\delta\tau}\left\{ P_1(q)\left(\frac{\Gamma_{\dot\phi}}{\Gamma_0'}\right)^2 \|\dot\phi\|^2 + P_2(q)\left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \frac{(\phi^T\zeta)^2}{m_f}\right\}$$
$$\cdot m_f(\tau)\, d\tau$$
$$+ (1+\tfrac{1}{\varepsilon})\, C_R' e^{2\delta t} + C'$$

where $P_1(q) = (1+q)$; $P_2(q) = (1+\frac{1}{q})$, $q > 0$

Applying the B-G lemma we find that m_f will be UB provided that

$$2\delta(t-\tau) > (1+\varepsilon)\int_\tau^t P_1(q)\left(\frac{\Gamma_{\dot\phi}}{\Gamma_0'}\right)^2 \|\dot\phi\|^2 + P_2(q)\left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \frac{(\phi^T\zeta)^2}{m_f}\, d\tau$$

Note that $\frac{(\phi^T\zeta)^2}{m_f} = \frac{(\phi^T\zeta)^2}{m} \cdot \frac{m}{m_f}$ and $\frac{m}{m_f}$ is UB :

1. $m = 1 + \zeta^T\zeta$; $|\zeta|^2/m_f \le g_{2\delta}(H_z) < \infty$
2. $m = (e^{\delta_0 t}\|U\|_{2\delta_0} + q_e)^2$; $\delta_0 > \delta \Rightarrow m/m_f \le c < \infty$

Hence, since $\|\mathring{\phi}\|$, $\frac{\phi\xi}{\sqrt{m}} \in L_2$ the inequality is trivially satisfied, for any $\delta > 0$. $\therefore$ m_f is UB.,

Remarks 1) Notice that in the ideal case

$$\frac{\phi\xi}{\sqrt{m}} \in L_2 \Rightarrow \int_{\tau}^{t} \left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \frac{(\phi\xi)^2}{m_f} d\tau < \infty$$

for any constants Γ_1, Γ_0'. In this case we did not have to impose any new constraints on 'a' and the ficticious filter $\Lambda(s)$ other than $\delta_0' > 0$. It will not be so in the case of unmodelled dynamics where $\frac{(\phi\xi)^2}{m}$ is simply "small in the mean" i.e.

$$\int_{\tau}^{t} \left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \frac{(\phi\xi)^2}{m_f} d\tau \leq \left(\frac{\Gamma_1}{\Gamma_0'}\right)^2 \mu (t-\tau) + c$$

2) The bounds are valid provided that $u_p(t), y_p(t)$ etc. exist. (see small gain theorem) To show existence + uniqueness of solutions we rely on the independent result, that ϕ is bounded and therefore none of the closed loop signals can grow faster than an exponentials (i.e. they belong in L_∞^e)

3) Finally we need one more step to conclude the analysis, that is to show that u_p and y_p are U.B. (All we have is $\epsilon_\delta \|U_t\|_{2\delta}$ is UB). For this, simple calculations show that

$\|\dot{U}_t\|_{2\delta} = O[\|U_t\|_{2\delta}]$ (Notice that $\mathring{\phi}$ is UB) and therefore the result follows (see related lemma in previous handouts).

From the boundedness of u_p, y_p we also conclude that m and all the closed loop signals are bounded $\therefore$ $\phi\xi \in L_2$. Further from $\mathring{\phi} \in L_\infty \cap L_2 \Rightarrow \phi\xi \to 0$, $e_1 = y_p - y_m \in L_2$ & $\dot{e}_1 \in L_\infty \Rightarrow e_1 \to 0$ as $t \to \infty$.

Further remarks

In the presence of arbitrary initial conditions $e_1 = \phi^T \xi + \varepsilon_t$ and a slight modification of the previous proof is required. In this case it is easy to show that ε_t (corresponding to IC applied on the nominal ($\Rightarrow$ ES) loop) can be described as

$$\dot{\xi} = A\xi \;;\; \xi(0) \;,\quad \epsilon_t = c\xi \qquad A: \text{Hurwitz}.$$

Hence we can take

$$V = \frac{1}{2\gamma}\phi^T\phi + \beta\,\xi^T P \xi$$

with $P = P^T > 0$ s.t. $A^T P + PA = -I$; $\beta > 0$ (suff. large)

Then
$$\dot{V} = -\epsilon_1 \frac{\phi^T \xi}{m} - \beta\,\xi^T\xi$$
$$= -\epsilon_1^2/m + \frac{c\,\xi\,\epsilon_1}{m} - \beta\,\xi^T\xi$$
$$\le -\frac{\epsilon_1^2}{m} \underbrace{- \beta\|\xi\|^2 + \frac{\|c\|\|\xi\||\epsilon_1|}{m}}_{\text{complete the squares}}$$
$$\le -\frac{\epsilon_1^2}{m} + \frac{|c|^2}{4\beta m}\cdot\frac{\epsilon_1^2}{m}$$

Choosing β, which is an arbitrary constant, to be suff. large and since $\frac{1}{m} \le c < \infty$

we can make $\frac{\|c\|^2}{4\beta m}$ arbitrarily small, say less than ε.

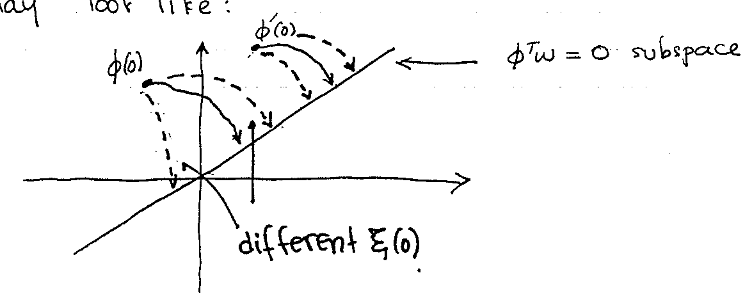
Then $\dot{V} \le -\frac{\epsilon_1^2}{m}\cdot(1-\varepsilon) \le 0$

Hence, the previous arguments are still valid.

Notice, however, that the bound on the parameter error depends on the initial conditions $\xi(0)$ and $\phi(0)$. In other words $\phi(t)$ is not uniformly bounded w.r.t. initial conditions (it is not Unif. Ultimately bounded) and

$$\|\phi\|^2 \le \|\phi(0)\|^2 + \underbrace{2\gamma\beta\,\lambda_{max}(P)\,\|\xi(0)\|^2}_{\substack{O(\gamma)\text{, but may be} \\ \text{arbitrarily large}}}$$

Pictorally, without PE, the possible trajectories of ϕ may look like:



This observation indicates that our stability result is not uniform w.r.t. I.C. which, at this point, is of no concern since $\epsilon/\sqrt{m}$, $\dot{\phi} \in L_2$ but it may destroy stability / boundedness when disturbances / unmodeled dynamics are present.

This is actually the case, and in general the vector field for ϕ should be modified to guarantee Ultimate boundedness and some robustness ppties in non ideal cases.

Such modifications fall under the general characterization of "projections" and two of them are:

1). Soft Projection or leak or σ-modification (Ioannou + Kokotovic)

$$\dot{\phi} = \dot{\theta} = -\gamma \frac{\epsilon_1 \xi}{m} - \sigma\theta \quad ; \quad \sigma > 0$$

In this case $\dot{V} = \phi^T \dot{\phi} = -\gamma \frac{\epsilon_1 \phi^T \xi}{m} - \sigma \theta \phi$
$(V = \frac{1}{2}\phi^T\phi)$

for which it can be easily shown that when $\|\phi\|$ is large $\sigma\theta^T\phi = O(\|\phi\|^2)$

$$-\sigma\theta\phi = -(\sigma\theta^{*}\phi + \sigma\|\phi\|^2) \leq -\frac{\sigma}{2}\|\phi\|^2 + \frac{\sigma}{2}\|\theta^{*}\|^2.$$

Using the previous techniques we also have that $-\gamma\frac{\epsilon_1\xi}{m} \leq$ const. $\therefore$ $\dot{V} \leq -\frac{\sigma}{2}\|\phi\|^2 + \text{const.}$

$\therefore$ V is uniformly ultimately bounded.

Furthermore, $\|\phi\|$ converges exponentially fast (with rate $O(\sigma)$) to a residual set $\|\phi\| < \text{const.}$ where the constant is independent of initial conditions

In other words $\|\phi\| \leq C_\phi + \varepsilon_t$; C_ϕ independent of IC and, after taking care of ε_t in the usual way, the constant C_ϕ can be used in our stability condition, instead of $\|\phi\|_\infty$.

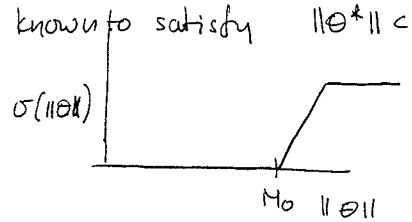
This modification has the drawback of introducing a bias in the parameter estimates and the tracking / estimation error. (see HW#3)

At the expense of requiring some additional a priori information on θ^* (usually available by the nature of the problem, physical constraints etc) this situation can be remedied by using smooth projections or "smooth switching-σ" (soft projection) modifications :

2) switching - σ . $\mathring{\theta}$ as before, where now,

$$\sigma = \begin{cases} 0 & \text{if } \|\theta\| < M_0 \\ \sigma_0 > 0 & \text{if } \|\theta\| \geq M_0(1+\varepsilon) \\ \sigma_0 \frac{\|\theta\| - M_0}{\varepsilon M_0} & \text{otherwise} \end{cases}$$

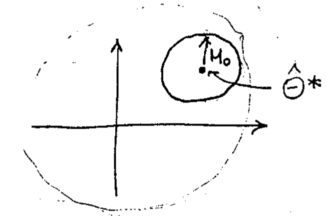
and $\|\theta^*\|$ is known to satisfy $\|\theta^*\| < M_0$



In this case, $\sigma \theta^T \phi \geq 0 \quad \forall \phi$ and it is $O(\|\phi\|^2)$ for $\|\phi\| > 2M_0(1+\varepsilon)$

The advantage, in this case, is that the equilibrium of the unperturbed system is $e_1 = 0$ $\phi = 0$ and the adaptive controller guarantees convergence of e_1 to zero in the absence of unmodeled dynamics / disturbances.

Notice that, in practice, the information on θ^* is given in terms of ellipsoids, or hypercubes not necessarily centered at 0. e.g. $\theta^* \in \{ \|\theta - \hat{\theta}^*\| < M_0 \}$ where $\hat{\theta}^*$ is an initial guess / constant bias

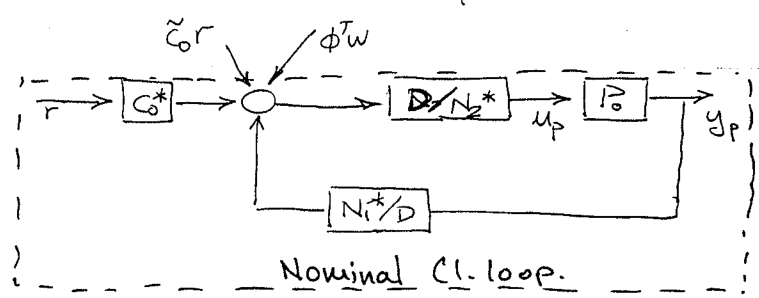


In such a case we should replace θ in the adaptive law and the $\sigma(\|\theta\|)$ expressions by $\theta - \hat{\theta}^*$. The previous analysis holds with $C_\phi = M_0(1+\varepsilon)$.

Finally, smooth projections also achieve the same result, i.e. to ~~keep~~ $\|\phi\|$ inside an ellipsoid of fixed radius without altering the ppties of $\epsilon_1/\sqrt{m}$ or $\dot{\phi}$ (L_2 functions).

Case 2. K_p unknown (positive)

When K_p is unknown, so is c_0^* and the MRAC closed loop becomes



i.e. the perturbation entering the loop due to the unknown parameters is now

$$\tilde{c}_0 r + \phi^T w$$

As in the previous case, there are two basic steps in the analysis / design of a MRAC:

1). Establish the error equation which will determine the parameter update law. Derive the adaptation properties.

2). Describe the closed loop signals in terms of the quantities involved in the adaptive law. Use the properties of adaptation to establish stability / convergence.

Step 1). From the previous closed loop description and the definition of c_0^*, N_1^*, N_2^* we have

$$y_p = W_M(s)\cdot\frac{1}{c_0^*}(\bar{\phi}^T\bar{w}) + W_M(s)\, r$$

$$\bar{\phi} = \begin{bmatrix}\tilde{c}_0\\ \phi\end{bmatrix};\quad \bar{w} = \begin{bmatrix} r\\ w\end{bmatrix}$$

Hence, $e_1 = y_p - y_m = W_M(s)\cdot\frac{1}{c_0^*}(\bar{\phi}^T\bar{w})$

(Exp. decaying terms due to IC. are omitted for simplicity).

Note that there is a fundamental difference between this description of the tracking error and the one obtained when c_0^* was known. Namely, the perturbation $\bar{\phi}^T \bar{w}$ is now filtered by the (partially) unknown transfer function $W_M(s) \frac{1}{c_0^*}$, instead of $W_M(s)$.

We will therefore have to modify our construction of the augmented error.

There are several ways of doing this,

1). Introduce an auxiliary parameter, say ψ_0, to estimate $\frac{1}{c_0^*}$. Note that $\frac{1}{c_0}$ cannot be used directly as an estimate of $\frac{1}{c_0^*}$. Without special provisions, the adaptive law does not guarantee that c_0 will be bounded away from 0 and therefore $\frac{1}{c_0}$ may become arbitrarily large or even undefined (e.g. $c_0(t)=0$ at some t).

The introduction of the additional parameter ψ_0 to estimate $\left(\frac{1}{c_0^*}\right)$ solves this problem (see Narendra Lin + Valavani, IEEE AC 1980). It has the disadvantage, however, that parameter convergence is not possible even if $r(t)$ has more than $\frac{2n+1}{2}$ frequencies.

2) Constructing an "input error" equation (Sastry + Bodson).

3). Use an a priori known lower bound of c_0^* to constraint the estimate c_0.

(This approach will be presented here).

Consider the known signal $c_0 e_1$.

$$c_0 e_1 = c_0^* e_1 + \tilde{c}_0 e_1 = W_M(s)\, \bar{\phi}^T \bar{w} + \tilde{c}_0 e_1$$

Construct the augmented error

$$\epsilon_1 = \underbrace{c_0 e_1 + c_0 y_m}_{= c_0 y_p} + \underbrace{\theta^T \zeta}_{W_M(s)(w)} - W_M(s)\,[\underbrace{c_0 r + \theta^T w}_{u_p}]$$

Then,

$$\epsilon_1 = c_0 e_1 + \tilde{c}_0 y_m + \phi^T \xi - W_M(s)\left[\tilde{c}_0 r + \bar{\phi}^T \bar{w}\right]$$

$$= \tilde{c}_0 y_m + \phi^T \xi + \tilde{c}_0 W_M(s) \frac{1}{c_0^*}(\bar{\phi}^T \bar{w}).$$

$$= \tilde{c}_0 \underbrace{W_M(s)\left[r + \frac{1}{c_0^*}\tilde{c}_0 r + \frac{1}{c_0^*}\phi^T w\right]}_{\frac{c_0 r}{c_0^*}} + \phi^T \xi$$

$$= \tilde{c}_0 y_P + \phi^T \xi \triangleq \bar{\phi}^T \bar{\xi} \quad ; \quad \bar{\xi} = \begin{bmatrix} y_P \\ W_M(s)\, w \end{bmatrix}$$

(Note $\bar{\xi} \neq W_M(s)\bar{w}$)

Hence:

Construct: $\epsilon_1 = c_0 y_P + \theta^T \xi - W_M(s)[u_P]$

It follows that

$$e_1 = \bar{\phi}^T \bar{\xi} \quad ; \quad \bar{\xi} = \begin{bmatrix} y_P \\ W_M(s)\, w \end{bmatrix}$$

Estimate $\bar{\theta}^* = \begin{pmatrix} c_0^* \\ \theta^* \end{pmatrix}$ with $\bar{\theta} = \begin{bmatrix} c_0 \\ \theta \end{bmatrix}$ as

$$\dot{\bar{\theta}} = \dot{\bar{\phi}} = -\gamma \mathrm{Pr}\left[\frac{\epsilon_1 \bar{\xi}}{m}\right] \qquad \left(\mathrm{Pr}: \text{Projection}\right.$$ s.t. $c_0 > c_{0min}$.$)$

A simpler form of the adaptation above can be written as follows

$$\dot{c}_0 = \dot{\tilde{c}}_0 = -\gamma \mathrm{Pr}\left[\frac{\epsilon_1 y_P}{m}\right] - \sigma_c c_0$$

$$\dot{\theta} = \dot{\phi} = -\gamma\left[\frac{\epsilon_1 \xi}{m}\right] - \sigma\theta$$

where, Pr: smooth projection of c_0 in the interval $[c_{0min}, c_{0max}]$ (see p. 205).

may be ∞

σ_c : switching σ-modification to constraint the upper bound of c_0

σ : switching σ-modification for θ.

Needless to say, that Projections can be used to obtain bounded parameter estimates for both c_0 and θ, and the adaptive law can be particularly simple to implement if $\bar{\theta}^*$ belongs to a hypercube or even ellipsoids. On the other hand, the σ-modification offers some advantages in the event the θ^* constraints have been under estimated.

It is now straightforward to apply our Lyapunov techniques and show that $|\phi|$ is UB $\epsilon_1/\sqrt{m}$, $\|\dot{\phi}\| \in L_2$ and $\frac{\epsilon_1}{\sqrt{m}}, \dot{\phi}$ are UB

Step 2).

Closed loop equation:

$$\begin{bmatrix} u_p \\ q_m y_p \end{bmatrix} = \sqcup = H(s)\left[r + \frac{1}{c_0^*}\bar{\phi}^T\bar{w}\right] = H(s)\frac{1}{c_0^*}[\bar{\phi}^T\bar{w}] + R$$

$$H(s) = \begin{bmatrix} S_u(s) \\ q_m W_M(s) \end{bmatrix} \qquad \begin{matrix} S_u(s)\text{: tr.f. } r \to u_p \\ W_M(s) \text{ tr.f. } r \to y_p \end{matrix} \left\{ \begin{matrix} \text{Nominal} \\ \text{cl. loop.} \end{matrix} \right.$$

Let $\hat{w} = W_M^{-1}(s)\bar{\xi} = \begin{bmatrix} W_M^{-1} y_p \\ W_M^{-1}\xi \end{bmatrix} = \begin{bmatrix} \frac{c_0}{c_0^*} r + \frac{1}{c_0^*}\phi^T w \\ w \end{bmatrix}$

Hence, $$\bar{\phi}^T\hat{w} = \tilde{c}_0\left[\frac{c_0}{c_0^*} r\right] + \underbrace{\frac{\tilde{c}_0}{c_0^*}\phi^T w + \phi^T w}_{\frac{\tilde{c}_0 + c_0^*}{c_0^*}\phi^T w}$$

$$= \frac{c_0}{c_0^*}\left(\tilde{c}_0 r + \phi^T w\right) = \frac{c_0}{c_0^*}\bar{\phi}^T\bar{w}$$

or $$\frac{1}{c_0}\bar{\phi}^T\hat{w} = \frac{1}{c_0^*}\bar{\phi}^T\bar{w}$$

Thus,

$$\sqcup = H(s)\left[\frac{1}{c_0}\bar{\phi}^T\hat{w}\right] + R$$

where, due to the projection, $\frac{1}{c_0(t)}$ is well defined for all t and UB.

We may now repeat the previous steps to decompose $\|\frac{1}{c_0}\bar{\phi}^T\hat{w}\|_{2\delta}$ as a weighted sum of three terms

1) a term depending on $\|\dot{\bar{\phi}}\sqrt{m_f}\|_{2\delta}$
2) a term depending on $\|\bar{\phi}^T W_M(s)\hat{w}\|_{2\delta} = \|\bar{\phi}^T\bar{\xi}\|_{2\delta}$
3) a term depending on $\|\sqcup\|_{2\delta}$

where the weight of term #2 is $O(a^{n-m})$
" " " - #3 is $O(\frac{1}{a})$

and a, δ are "free" parameters used only in the analysis and s.t. $\gamma_{2\delta}(H)$ is well defined and $a > \delta$. With similar arguments as before we may conclude that $\sqcup(t)$ is UB and $e_1 \to 0$

An important observation is that the bound of $\frac{1}{c_0(t)}$ will now appear in the expressions of the various weights. This is, of course, of no consequence in the ideal case.. It will be crucial, however, in the analysis of any application of a MRAC scheme where disturbances/unmodeled dynamics are present. Allowing $c_0(t)$ to approach 0 will severely limit any robustness properties of the adaptive controller.

We will conclude the presentation of the stability analysis of a MRAC in the ideal case by noting that parameter convergence ($\phi \to 0$) can be achieved provided that $r(t)$ is sufficiently rich.
The result is <u>non-trivial</u> and for details, see Narendra + Annaswamy and/or Sastry + Bodson. (and references therein)

ROBUSTNESS OF MRAC

CASE 1 : DYNAMIC UNCERTAINTY / UNMODELED DYNAMICS

A crucial assumption in the previous development was that the plant was described as

$$y_p = P_0(s)\, u_p$$

where $P_0(s)$ was a $\overset{\text{min. phase}}{\overline{\text{transfer}}}$ function of known order and relative degree. However, more often than not, such assumptions are only "approximately" valid in applications where the true plants are non-linear and/or infinite dimensional. Typically, the best we can hope for in practical situations, is to have some information on a nominal, approximate plant transfer function which is related to the true plant description by means of a dynamic uncertainty operator.

For example consider the plant;

$$y_p = P(s)\,u_p$$

where $P(s) = P_o(s)\,(1 + \Delta(s))$

$\Delta(s)$: multiplicative uncertainty.

(Note: Other uncertainty models – additive, stable factor perturbations – can be similarly analyzed and will be omitted from the present discussion).

A measure of "how well $P(s)$ is approximated by $P_o(s)$" can then be given in terms of the "size" of the operator $\Delta(s)$. It should be mentioned that this statement makes sense only when the input and output spaces of the operators are defined. (e.g. $L_2 \to L_2$, $L_{2\delta} \to L_{2\delta}$, $L_2 \to L_\infty$ etc).

This being the case, the "size" of $\Delta(s)$ is the induced gain of $\Delta(s)$ from its input space to its output space.

The robustness problem can now be formulated as follow.

Consider the plant

$$y_p = P(s)\,u_p$$

and let $P(s)$ be described as

$$P(s) = P_o(s;p)\,[\,1 + \Delta(s;p)]$$

$P_o(s;p)$ denotes the "nominal" plant which is parametrized by a set of parameters p.

$\Delta(s;p)$ denotes a dynamic uncertainty operator which describes the effects of unmodeled dynamics not included in P_o and which, in general, depends on p..

All operators are assumed to be causal and exponentially stable. Furthermore, with some extra work, we can allow Δ to include mild (sector bounded) nonlinearities.

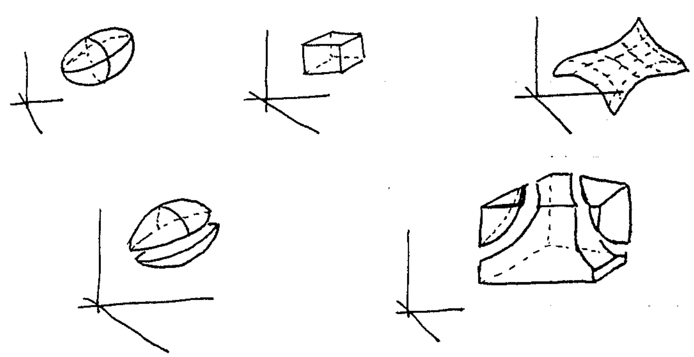
Further, suppose that there exists a nominal parameter vector $p*$ for which $\Delta(s;p*)$ is "small" in some sense. Although the smallness requirements on $\Delta(s;p*)$ will be defined precisely after we solve the problem, it suffices, for the moment, to think of $p*$ as the vector for which $\gamma_{2\delta}[\Delta(s;p*)]$ is small and $g_{2\delta}[\Delta(s;p*)]$ is small.

<u>REMARK</u> This is similar to ~~the~~ standard model-order reduction / approximation problem where, given $P(s)$ and $P_o(s;p)$, e.g. n-th order t.f. with coefficients . we seek a vector $p*$ s.t. $\gamma_2[\Delta(s;p*)]$ is minimum. The problem does not necessarily admit a unique solution nor it is easy. A variant of this problem was solved in [Glover, Int. Journal of Control 1984 "All optimal Hankel-norm approximations of linear multivariable systems and their L_∞ error bounds".

In the adaptive control formulation however we are not required to find $p*$; we just need to know that it exists and is s.t. $P_o(s;p*)$ has certain properties (controllability, observability, order + relative degree known, min-phase; whichever necessary). This is exactly one of the advantages of adaptive control. $p*$ may be "too expensive" or even impossible to determine, or it may change with time (the time-varying problem requires some additional tools and will be omitted from the present discussion)

In addition, even if $P(s) = \bar{P}_o(s;p*)$ the order of $\bar{P}_o$ may be too large, requiring a very "expensive" controller. In such a case, it may be advantageous to consider a low-order approximation with one of the previously mentioned uncertainty models.

A natural assumption at this point is that $p*$ belongs to a (usually) bounded set $\mathcal{P}$ for which we have some a priori information.
(e.g. physical constraints of the problem).
Let us denote by $D_{\mathcal{P}}$ the diameter of the set $\mathcal{P}$
i.e. $D_{\mathcal{P}} = \sup \{ \|p_1 - p_2\| ; p_1, p_2 \in \mathcal{P} \}$.
Such sets can be ellipsoids or hypercubes in the simplest cases, or surfaces or even a collection of disjoint sets.



Note that $D_{\mathcal{P}}$ is a measure of the size of the parametric uncertainty in the description of $P(s)$

Also, suppose that the non-parametric uncertainty $\Delta(s; p*)$ satisfies

$$\gamma_{2\delta}[W_1(s)\Delta(s;\theta*)] \leq \mu_1$$
$$g_{2\delta}[W_2(s)\Delta(s;\theta*)] \leq \mu_2$$

where μ_1, μ_2 are known constants and
W_1, W_2 are known ES, min-phase, weighting operators;
that is μ_1, μ_2 define — in a weighted sense — the size of non-parametric (dynamic) uncertainty.

For the plant $P(s)$, the control objective (MRC) can be defined as: "design $u_p = f(y_p, r)$ such that y_p tracks asymptotically the output of a reference model $y_M = W_M(s)$ with input r, as close as possible and all closed loop signals remain UB" —
In this setup we can now state a variety of MRC / MRAC related problems.

1. MRC : Ideal case.

Suppose $D_p = \mu_1 = \mu_2 = 0$. Design u_p s.t. $y_p - y_m \to 0$ and the closed loop is internally stable.

The solution of this problem was given in an earlier part of these notes (controller design) under the conditions that $P_o(s; p^*)$ satisfies the MRC assumptions and $W_M(s)$ is appropriately selected. (these conditions are also assumed to hold in the following)

2. MRC : **Robustness**

Suppose $D_p = 0$; $\mu_1, \mu_2 > 0$. Design u_p to satisfy the MRC objective.

(Synthesis Problem)

or, suppose $D_p = 0$ and consider an MRC designed for $\mu_1 = \mu_2 = 0$.
Determine μ_1 and/or μ_2 s.t. closed loop stability is preserved.
(Analysis Problem)

Solutions for the analysis problems can be given in terms of the small gain theorem. The synthesis problem is, of course, more complicated. For solutions, see Francis "A course on H_∞ control theory" Springer-Verlag.
Note that if D_p is allowed to be non-zero the solution of the problem using a linear controller becomes very hard and/or conservative (conservatism of SGT).

3. **MRAC : Ideal case**

Suppose $\mu_1, \mu_2 = 0$. Design u_p ~~to~~. satisfy the MRC objective.

This problem was solved previously using a special form of a nonlinear controller:

Linear control + Estimation.

This control law – termed MRAC –

was able to satisfy the control objective (under the MRAC assumption) for an arbitrary finite value of $p*$. That is MRAC has "infinite" robustness margin w.r.t. parametric uncertainty.

4. MRAC : Robustness

1. Let $D_p > 0$, $\mu_1, \mu_2 > 0$. Design u_p to satisfy the MRAC objective.
(Synthesis Problem)

2. Let $\mu_1, \mu_2 > 0$. Design u_p to satisfy the MRAC objective and maximize D_p (or vice-versa)

3. Consider a MRAC designed as in part 3. Given D_p, find μ_1, μ_2 for which the closed loop signals remain U.B. (or vice-versa)
(Analysis Problem : the classical MRAC robustness)

This problem will be discussed in some detail next.

A MRAC Robustness Problem.

Assume that $P_o(s; p*)$ and $W_M(s)$ satisfy the standard MRAC assumptions and let $C(s; \bar{\theta})$ be the linear time invariant controller s.t. $C(s; \bar{\theta}*)$ meets the MRC objective for $P_o(s; p*)$.

Further assume that the plant parametric uncertainty expressed in terms of a controller parametric uncertainty, is s.t.

$$\bar{\theta}* \in \Theta$$

where Θ is a closed, bounded convex set.

Wolog, and in order to simplify our discussion let us normalize Θ s.t.

$$\Theta = \{ \bar{\theta} \mid \|\bar{\theta}\| \leq M_o \}$$

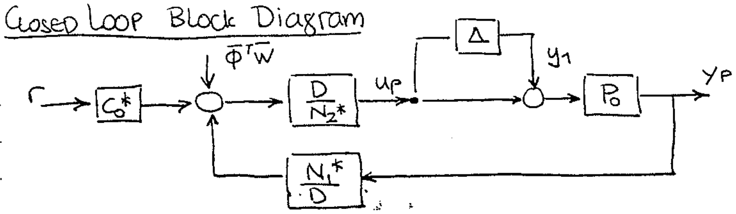
This is always possible by passing the necessary translations / scalings into the auxiliary filter transfer functions of the MRC design. We will keep M_0 however, as a measure of the size of Θ or, in other words, the size of parametric uncertainty reflected on the controller parameters.

We seek to establish "realistic" conditions on $\Delta(s;p^*)$ s.t. a MRAC designed along the lines of the sol'n of problem 3 guarantees the boundedness of all closed loop signals.

In this approach we will allow for possible modifications of the adaptive laws and a solution which:

1. is global w.r.t. I.C.
2. does not require PE
3. employs direct adaptation of the controller parameters.

Closed Loop Block Diagram



$$\bar{\phi} = \begin{bmatrix} \tilde{c}_0 \\ \phi \end{bmatrix} \qquad \bar{w} = \begin{bmatrix} r \\ w \end{bmatrix}$$

$$w = \bar{G}(s)\, Q_m^{-1}\, U \qquad U = \begin{bmatrix} u_p \\ q_m y_p \end{bmatrix}$$

Let $\tilde{y} = W_1 \Delta u_p = W_1 y_1$; W_1 a frequency weighting factor ; poles + zeros in lhp.

S_u : $\bar{\phi}^T\bar{w} \to u_p$ I/O operator.

S_T : Complementary Sensitivity $(y_1 \to u_p)$

S_y : $y_1 \to y_p$ I/O operator $= \frac{1}{c_0^*} W_M \frac{N_2^*}{D}$

Then

$$u_p = S_u\left(\bar{\phi}^T\bar{w} + c_0^* r\right) + S_T\, y_1$$

$$y_p = W_M\left(\frac{1}{c_0^*}\bar{\phi}^T\bar{w} + r\right) + S_y\, y_1$$

$$\begin{bmatrix} u_p \\ q_m y_p \end{bmatrix} = U = \begin{bmatrix} S_u & S_T W_1^{-1} \\ W_M \cdot \frac{q_m}{c_0^*} & q_m S_y W_1^{-1} \end{bmatrix} \begin{bmatrix} \bar{\phi}^T\bar{w} + c_0^* r \\ \tilde{y} \end{bmatrix}$$

where q_m : the weighting used in the normalizing signal m.

or $u = H(s)\begin{bmatrix} \bar{\phi}^T\bar{w} \\ \tilde{y} \end{bmatrix} + R$

where, for $H(s)$ to be proper, W_1 should have relative degree at most $n-m$.

ADAPTATION

1. Augmented error.

$$\epsilon_1 = c_0 y_p + \theta^T \zeta - W_M(u_p)$$
$$= (\tilde{c}_0, \phi^T)\begin{pmatrix} y_p \\ \zeta \end{pmatrix} + W_M F_1 \Delta[u_p]$$
$$= \bar{\phi}^T \bar{\zeta} + \underbrace{W_M F_1 \Delta(u_p)}_{n}. \quad (+\epsilon_t)$$

$$\bar{\phi} = \begin{pmatrix} \tilde{c}_0 \\ \phi \end{pmatrix}, \quad \bar{\zeta} = \begin{pmatrix} y_p \\ \zeta \end{pmatrix}$$
$$F_1 = \frac{N_2^*}{D}.$$

Note: ϵ_t : exponentially decaying terms

min. rate of decay : stability margin of $H(s)$

i.e. $\epsilon_t \le K_0 \cdot e^{-\alpha t}$;

K_0 : constant depending on I.C.

α : $H(s)$ analytic in $Re(s) > -\alpha$

(W_1 chosen appropriately).

2. Parameter Update

$$\dot{\bar{\phi}} = -\gamma\, P_r \frac{\epsilon_1 \bar{\zeta}}{m} \qquad \text{(Projection)}$$

or

$$\dot{\bar{\phi}} = -\gamma\, P_r \frac{\epsilon_1 \bar{\zeta}}{m} - \gamma\sigma\bar{\theta} \qquad \text{(Mixed projection and switching } \sigma\text{-modif.)}.$$

When projection is used for all the parameters in $\bar{\phi}$;

$$V = \tfrac{1}{2}\bar{\phi}^T\bar{\phi} \Rightarrow \dot{V} = -\gamma\, P_r \frac{\epsilon_1 \bar{\phi}^T \bar{\zeta}}{m}$$
$$\le \frac{-1}{2\gamma}(1-\varepsilon)\frac{\epsilon_1^2}{m} + \frac{1}{2\gamma}\frac{n^2}{m} + \frac{\|m^{-1}\|_\infty}{8\gamma\varepsilon}\epsilon_t^2$$

$\varepsilon > 0$ arbitrarily small

ϵ_t expon. decaying terms ; rate of decay : at least as fast as the stability margin of H.

Due to projection, V is UUB $\Rightarrow$

$$(1-\varepsilon)\int_{t_0}^{t_0+T}\frac{\epsilon_1^2}{m} \le \frac{1}{\gamma}\underbrace{[\bar{\phi}^T\bar{\phi}(t_0) - \bar{\phi}^T\bar{\phi}(t_0+T)]}_{\le C} + \underbrace{\frac{\|m^{-1}\|_\infty}{4\varepsilon}\int_{t_0}^{t_0+T}\epsilon_t^2}_{\le C_\varepsilon}$$
$$+ \int_{t_0}^{t_0+T}\frac{n^2}{m}.$$

Similarly for $\frac{(\bar{\phi}^T\bar{\zeta})^2}{m}$.

When the switching σ modification is used, $\sigma\bar{\theta}^T\bar{\phi} \ge 0$.

Assuming that $g_{2\delta_0}(W_M F_1 \Delta) < \infty$, $\exists\, V_0 > 0$ s.t. $V > V_0$

$\Rightarrow \dot{V} < 0 \Rightarrow V$ is UUB and similar bounds are obtained

Further, $\|\mathring{\phi}\|^2 = \gamma^2 P_r^2 \dfrac{\epsilon_1^2 \|\xi\|^2}{m^2}$ (Projection)

$$\leq \gamma^2 C_2^2 \frac{\epsilon_1^2}{m} + \gamma^2 \frac{\epsilon_1^2}{m}\frac{\epsilon_t^2}{m}$$

Finally, if $g_{2\delta_0}(W_H F_1 \Delta) < \infty$ we have

1. $\bar{\Phi}, \dot{\bar{\Phi}}$ UUB
2. $\bar{\theta} \in \Theta \Rightarrow \|\bar{\Phi}\| \leq 2M_0 = C_\phi$
3. $\displaystyle\int_{t_0}^{t_0+T} \frac{\epsilon_1^2}{m}, \int_{t_0}^{t_0+T} \frac{(\bar{\Phi}^T\bar{\xi})^2}{m} \leq C + g_{2\delta_0}^2 (W_H F_1 \Delta)\, T / (1-\varepsilon)$
4. $\displaystyle\int_{t_0}^{t_0+T} \|\dot{\bar{\Phi}}\|^2 \leq C + \gamma^2 C_2^2\, g_{2\delta_0}^2 (W_H F_1 \Delta)\, T/(1-\varepsilon).$

REM: • ε : arbitrarily small. Used for technical reasons and will not appear in the final result.

- Similar expressions can be obtained for the switching σ-modification.
- C : Constants depending on M_0, γ, ε.

Next, with $Q_d = \begin{pmatrix} 1 & \\ & q_d \end{pmatrix}$, $q_d > 0$ and

$\delta > 0$ s.t. $H(s)$ is analytic in $Re(s) > -\delta$.

$$U = H Q_d^{-1} \begin{bmatrix} \bar{\Phi}^T \bar{w} \\ q_d \tilde{y} \end{bmatrix} + R + \varepsilon_t$$

$$\|U_t\|_{2\delta} \leq \gamma_{2\delta}(HQ_d^{-1}) \left\| \begin{matrix} \bar{\Phi}^T \bar{w}_t \\ q_d \tilde{y}_t \end{matrix} \right\|_{2\delta} + \|R_t\|_{2\delta} + C$$

$$\leq \gamma_{2\delta}(HQ_d^{-1}) \cdot \sqrt{\|\bar{\Phi}^T \bar{w}_t\|_{2\delta}^2 + q_d^2 \gamma_{2\delta}^2 (W_1 \Delta) \|U_t\|_{2\delta}^2}$$

$$+ \|R_t\|_{2\delta} + C$$

- Decomposition of $\bar{\Phi}^T\bar{w}$ in terms of $\bar{\Phi}^T\bar{\xi}$ and $\dot{\bar{\Phi}}$.

(A slightly different approach is necessary to avoid the requirement of a strictly proper $\Delta(s)$)

Let $\hat{w} = \begin{bmatrix} \frac{1}{c_0^*} \bar{\Phi}^T \bar{w} + r \\ w \end{bmatrix}$. Then, $\frac{1}{c_0} \bar{\Phi}^T \hat{w} = \frac{1}{c_0^*} \bar{\Phi}^T \bar{w}$.

or $\boxed{\bar{\Phi}^T \bar{w} = \dfrac{c_0^*}{c_0} \bar{\Phi}^T \hat{w}}$

where by the ppties of the projection algorithm,

$\left\|\dfrac{c_0^*}{c_0}\right\|_\infty \leq \dfrac{c_0^*}{c_{0min}} < \infty$, $\left\|\dfrac{c_0}{c_0^*}\right\|_\infty \leq \dfrac{c_{0max}}{c_0^*} < \infty$.

Now, decompose $\bar{\Phi}^T\bar{w}$ using the previous techniques, as follows

$$\dot{\overline{\phi}^T\overline{w}} = \Lambda_1 (\overline{\phi}^T\overline{w}) + \Lambda W_M^{-1} W_M \left(\frac{c_0^*}{c_0} \overline{\phi}^T \hat{w} \right)$$

$$= \Lambda_1(\dot{\overline{\phi}}^T \overline{w}) + \Lambda_1(\overline{\phi}^T \dot{\overline{w}}) + \Lambda W_M^{-1} \frac{c_0^*}{c_0} \overline{\phi}^T W_M \hat{w}$$

$$+ \Lambda W_M^{-1} W_{M1}\, c_0^* \left(\frac{\dot{\overline{\phi}}^T}{c_0}\right) W_{M2} \hat{w}$$

$$\therefore \| \overline{\phi}^T \overline{w}_t \|_{2\delta} \le (\Gamma_0 C_{\overline{w}} + \Gamma_2 C_M) \| (\|\dot{\overline{\phi}}\| / \sqrt{m_f})_t \|_{2\delta}$$

$$+ \Gamma_1 \| (\overline{\phi}^T W_M \hat{w})_t \|_{2\delta}$$

$$+ \Gamma_0 C_{\overline{\phi}} C_{\hat{w}} \| U_t \|_{2\delta}$$

$$+ C_R e^{\delta t} + C.$$

$\Gamma_0 = \gamma_{2\delta}(\Lambda_1)$

$\Gamma_1 = \gamma_{2\delta}(\Lambda W_M^{-1}) \cdot \| \frac{c_0^*}{c_0} \|_\infty$

$\Gamma_2 = \gamma_{2\delta}(\Lambda W_M^{-1} W_{M1}) \| \frac{c_0^*}{c_0} \|_\infty$

$C_{\overline{w}}, C_M$: similar expressions as in the previous development; more complicated due to the appearance of $\overline{\phi}^T\overline{w}$ as a part of $\hat{w}$.

Technical remark : need $\dot{r}$ to be bounded or, at least, $\| \dot{r}_t \|_{2\delta} \le c \cdot e^{\delta t}$. $\Rightarrow$ r should not change abruptly too often ; use a prefilter to make r smooth.

Next, notice that

$$W_M(s)\, \hat{w} = \begin{pmatrix} W_M\left(\frac{1}{c_0^*} \overline{\phi}^T \overline{w} + r\right) \\ \xi \end{pmatrix} = \begin{pmatrix} y_p - W_M F_1 \Delta u_p \\ \xi \end{pmatrix}$$

$$\therefore \overline{\phi}^T W_M(s)\, \hat{w} = \overline{\phi}^T \overline{\xi} - \tilde{c}_0 W_M F_1 \Delta u_p.$$

$$\Rightarrow \| (\overline{\phi}^T W_M(s) \hat{w})_t \|_{2\delta} \le \| (\overline{\phi}^T \overline{\xi})_t \|_{2\delta} + \| \tilde{c}_0 \|_\infty \cdot \gamma_{2\delta}(W_M F_1 \Delta) \| U_t \|_{2\delta}$$

$$\therefore \| U_t \|_{2\delta}^2 \le (1+\varepsilon)\, \gamma_{2\delta}^2 (H Q_d^{-1}) \Big\{ q_d^2\, \gamma_{2\delta}^2 (\mathbf{W_1 \Delta}) \| U_t \|_{2\delta}^2$$

$$+ p_1(q) (\Gamma_0 C_{\overline{\phi}} C_{\hat{w}})^2 \| U_t \|_{2\delta}^2$$

$$+ p_2(q)\, \Gamma_1^2 \| \tilde{c}_0 \|_\infty^2\, \gamma_{2\delta}^2 (W_M F_1 \Delta) \| U_t \|_{2\delta}^2$$

$$+ p_3(q)\, \Gamma_1^2 \| \overline{\phi}^T \overline{\xi}_t \|_{2\delta}^2$$

$$+ p_4(q) (\Gamma_0 C_{\overline{w}} + \Gamma_2 C_M)^2 \| (\|\dot{\phi}\| \sqrt{m_f})_t \|_{2\delta}^2 \Big\}$$

$$+ \left(1 + \frac{1}{\varepsilon}\right)\left(C_R e^{2\delta t} + C\right)$$

$p_i(q)$: Cauchy constants.

Simplifying the notation:

$$\Gamma_H = \gamma_{2\delta}(H Q_d^{-1})$$

$$\Gamma_\Delta = \gamma_{2\delta}(W_1 \Delta)$$

$$\hat{\Gamma}_0 = \Gamma_0 C_\phi C_{\dot{\omega}} \quad ; \quad \hat{\Gamma}_1 = \Gamma_1 \|\tilde{c}_0\|_\infty \gamma_{2\delta}(W_H F_1 \Delta)$$

$$\hat{\Gamma}_2 = \Gamma_0 C_{\bar{\omega}} + \Gamma_2 C_H$$

$$\begin{aligned}
\frac{1}{(1+\varepsilon)\Gamma_H^2}\|U_t\|_{2\delta}^2 \leq\ & q_d^2 \Gamma_\Delta^2 \|U_t\|_{2\delta}^2 \\
& + p_1(q)\, \hat{\Gamma}_0^2 \|U_t\|_{2\delta}^2 \\
& + p_2(q)\, \hat{\Gamma}_1^2 \|U_t\|_{2\delta}^2 \\
& + p_3(q)\, \Gamma_1^2 \|\bar{\phi}^T \bar{\varepsilon}_{1t}\|_{2\delta}^2 \\
& + p_4(q)\, \hat{\Gamma}_2^2 \|(\|\dot{\phi}\|\sqrt{m_f})_t\|_{2\delta}^2 . \\
& + \left(1+\frac{1}{\varepsilon}\right)\left(c_R e^{2\delta t} + c\right)
\end{aligned}$$

Finally, using B-G lemma and selecting the cauchy constants to maximize the stability region we obtain :

THM The closed loop signals will be U.B. provided that.

$$\sup \left\{ \frac{1}{\Gamma_H^2} - q_d^2 \Gamma_\Delta^2 - \frac{1}{2\delta}\left[\sqrt{2\delta}\,(\hat{\Gamma}_0 + \hat{\Gamma}_1) + \Gamma_1\, g_{2\delta_0}(W_H F_1 \Delta) + \gamma \hat{\Gamma}_2 C_z\, g_{2\delta_0}(W_H F_1 \Delta)\right]^2 \right\} > 0$$

where the supremum is taken w.r.t.

$\delta \in (0, \alpha)$; α : the stability margin of $H(s)$ $(< \bar{\delta}_0)$

$q_d > 0$

$\Lambda(s), W_1(s)$. the auxiliary weighting filters.

Notes:
- The constants ε are absorbed by the strict inequality sign
- $\gamma \to 0$: maximizes "robustness" of MRAC
- $\frac{1}{\Gamma_H^2} - q_d \Gamma_\Delta^2 > 0$: Robustness of the nominal LTI controller. (SGT)

- Recovery of LTI Robustness :

$\gamma \to 0$, $M_0 \to 0$, $\delta_0 \to 0$.

<u>REM</u> : $\hat{\Gamma}_0 = O(\frac{1}{a}) \cdot O(M_0)$

$\Gamma_1 = O(a^{n-m}) \cdot O(\frac{1}{c_{min}})$

$\hat{\Gamma}_1 = O(a^{n-m})\, O(\frac{1}{c_{min}}) \cdot O(\gamma_{2\delta}(w_1 \Delta))$

$\Gamma_\Delta = O(\gamma_{2\delta}(w_1 \Delta))$

$\hat{\Gamma}_2 =$ complicated expression. $O(\frac{1}{a})$, $O(M_0)$, $O(a^{n-m})$, $O(\frac{1}{c_{min}})$

<u>OTHER QUALITATIVE REMARKS</u>

- Given any $M_0 < \infty$, $c_{min} > 0$ there exist $\mu_1 , \mu_2 > 0$

s.t. $\gamma_{2\delta}(w_1 \Delta) < \mu_1$

$g_{2\delta}(w_2 \Delta) < \mu_2$

guarantees boundedness of all closed loop signals.

- $M_0 \uparrow$ $\mu_1, \mu_2 \downarrow$: "Robustness w.r.t. unmodeled dynamics decreases as the parametric uncertainty increases

- $c_{min} \to 0$ $\mu_1 , \mu_2 \to 0$: If c_{min} is too small it may be advantageous to use an additional parameter to estimate $\frac{1}{c_0^*}$ (Narendra Lin Valavani).

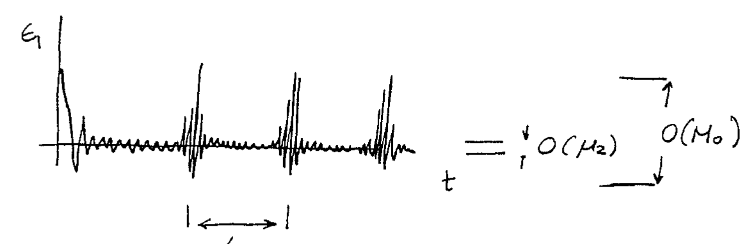
- $n-m \uparrow$ $\mu_1, \mu_2 \downarrow$: Poor robustness if the plant has high relative degree.

- "Performance" Characterization

$$\int_t^{t+T} \frac{e_1^2}{m} \le C + K \cdot g_{2\delta_0}^2 (W_H F_1 \Delta) \cdot T$$

due to swapping terms $O(1+\gamma)$

'Good' performance on the average. But may contain Bursts $O(M_0)$



Bursts can be avoided if $\phi^T\phi(t_0) - \phi^T\phi(t_0+T) < \varepsilon$ $\forall T$; t_0 ~~large~~ enough.

e.g. $\phi^T\phi \to$ const. (see p. 277)

A popular way of achieving this is through the use of dead-zones (discussed later).

SOME DESIGN GUIDELINES

- The previous stability theorem offers some general design guidelines although some caution should be exercised in interpreting the results. The theorem gives a conservative condition for global boundedness. It will not necessarily produce the "best" MRAC if μ_1, μ_2 are maximized w.r.t. the various design parameters ($W_M(s)$, $D(s)$, q_m, δ_0, γ etc). It does indicate however that $W_M(s)$ should be used more as a tuning parameter of the closed loop sensitivities and less as a tracking specification.

Afterall, tracking can be modified by using a prefilter.

Other tools should be used as well.

E.g. local analysis [Anderson et al. "Stability of Adaptive Systems" MIT press, 1986]

The design of adaptive (Non Linear) controllers is not straightforward and extensive and careful analysis is required. General theorems can give a rough idea of what a good design should look like. The adaptive controller should then be tailored to the needs of the specific problem. (Further comments on design guidelines will be given later).

DEAD ZONE MODIFICATION

This is a popular and quite intuitive modification, motivated by the idea to stop adaptation when the signal-to-noise-ratio becomes small.

(Ref: Peterson + Narendra 'Bounded Error Adaptive Control' IEEE AC Dec. 1982)

Briefly described, an adaptive law with dead zone is

$$\dot{\theta} = -\gamma d_z \frac{\epsilon_1 \bar{\zeta}}{m}$$

(or, adaptive law with dead-zone and projection:

$$\dot{\theta} = -\gamma d_z P_r \frac{\epsilon_1 \bar{\zeta}}{m} \Big)$$

where, for the case of unmodeled dynamics

$$\epsilon_1 = \bar{\phi}^T \bar{\zeta} + \eta \; ; \quad \eta = W_H F_1 \Delta u_p .$$

and d_z is taken as

$$d_z = \begin{cases} 0 & \text{if } \epsilon_1^2/m < M_d^2 \\ \dfrac{\epsilon_1^2/m - M_d^2}{M_d^2 E} & \text{if } M_d^2 \le \epsilon_1^2/m < M_d^2(1+E) \\ 1 & \text{if } \epsilon_1^2/m \ge M_d^2(1+E) \end{cases}$$

Note: This type of a dead-zone switches adaptation off when ϵ_1^2/m is less than some threshold. For this reason it is usually referred to as a relative dead zone.

It can be easily shown that if $M_d > g_{2\delta_0}(W_H F_1 \Delta)$

$$d_z \frac{\epsilon_1 \bar{\phi}^T \bar{\zeta}}{m} \ge d_z |\epsilon_1| \frac{\left(|\epsilon_1| - W_H F_1 \Delta(u_p)\right)}{m} \ge 0$$

(modulo exponentially decaying transients).

Additional calculations yield

$$\|\dot{\bar{\phi}}\| \in L_2 \; , \quad d_z \frac{\epsilon_1 \bar{\phi}^T \bar{\zeta}}{m} \in L_1$$

and

$$\|\dot{\bar{\phi}}\| \to 0 \; , \quad \frac{|\epsilon_1|}{m} , \frac{|\bar{\phi}^T \bar{\zeta}|}{m} \le M_d + g_{2\delta}(W_H F_1 \Delta) + h(t)$$

where $h(t) \in L_2$ and $h(t) \to 0$ as $t \to \infty$.

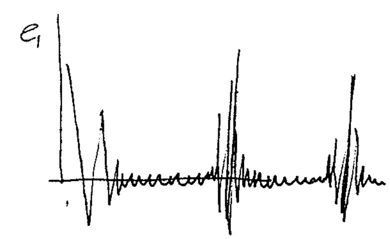
In this case, similar conditions for global boundedness can be obtained with the following additional remarks:

• 1. Global boundedness is guaranteed for sufficiently small M_d. In other words the dead-zone should not be too conservative (Notice that M_d will enter the various expressions together with $\mu_2 = g_{2\delta}(\omega_M F, \Delta)$)

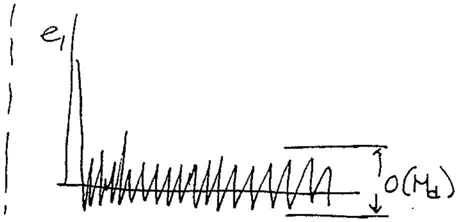
• 2. Asymptotic $\overset{\text{tracking}}{\text{performance}}$ is always $O(M_d^2)$ in the mean square sense i.e. $\int_{t_0}^{t_0+T} \frac{\epsilon_1^2}{m} \leq c + kM_d^2 T$

However, $\left|\frac{\epsilon_1^2}{m}\right|_{t \geq t_0} \leq O(M_d^2)$. as well.

That is, dead-zones can guarantee a uniform bound of the normalized tracking error and any bursts will be limited by $O(M_d)$



without dead zone



with dead-zone

The price paid is that asymptotic tracking is lost in the ideal case. i.e. even if $\mu_1, \mu_2 = 0$ $\epsilon_1^2/m = O(M_d^2)$.

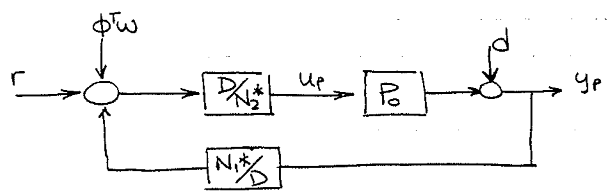
• 3 θ will converge to a constant provided that $M_d > g_{2\delta}(\omega_M F, \Delta)$ ... (Asymptotically LTI controller).

• 4. Typically, the best behavior of relative dead-zones is with fast unmodeled dynamics, e.g. $\frac{1}{\mu s+1}, \mu \ll 1$

• 5 Although $|\epsilon_1|_{t \geq t_0} = O(M_d)$ the "order of" may (and will) include the bound of m which is not necessarily small. In general, m depends on r and the control input of the "tuned" controller $(\theta = \theta^*)$.

ROBUSTNESS OF MRAC.

CASE 2. Bounded disturbances



$$\begin{bmatrix} u_p \\ q_m y_p \end{bmatrix} = \begin{pmatrix} S_u \\ q_m W_M \end{pmatrix} (\phi^T w) + R + \begin{pmatrix} S_{du} \\ q_m S_{dy} \end{pmatrix} d.$$

Augmented error

$$\epsilon_1 = \phi^T \zeta + S_{dy} d$$

where S_{dy} is the usual Sensitivity to output disturbances $: (1+CP)^{-1}$.

REM Good sensitivity properties of the nominal controller – e.g. using internal models – can make $S_{dy} d$ very small.

Following a similar analysis as in the case of dynamic uncertainty, ϕ UUB

$$\int_{t_0}^{t_0+T} \epsilon_1^2/m \leq C + \int_{t_0}^{t_0+T} \frac{|S_{dy} d|^2}{m} \quad \text{etc.}$$

The difficulty in this case is that $\int_{t_0}^{t_0+T} \frac{|S_{dy} d|^2}{m}$ is not necessarily small in the mean. Typical analytical arguments that have been used in this case can be classified in three categories.

1) An internal model of the disturbances is available and has been incorporated in the controller.

$\therefore S_{dy} d \in L_2$ and $\int \frac{|S_{dy} d|^2}{m} < C$

(Too restrictive)

Similarly, using an internal model and a large constant q_ϵ in the normalizing signal $m = (\mathcal{E}_{-\delta_0} \| u \|_{2\delta_0} + q_\epsilon)^2$ we have that

$$\int_{t_0}^{t_0+T} \frac{|S_{dy}d|^2}{m_f} \leq c + \underbrace{\frac{|S_{dy}d|^2}{q_\epsilon^2}}_{\text{suff. small, say } \leq \mu_3^2} \cdot T$$

The previous analysis can now produce global boundedness for μ_3 small enough. The "disadvantages" of this approach are:

1. Some stability margin $(2\bar{\delta})$ is (unnecessarily) traded-off for disturbances
2. $\|d\|_\infty$ should be "small". s.t. $\frac{|S_{dy}d|^2}{q_\epsilon^2} \leq \mu_3^2$.
3. Adaptation may become too slow due to the large values of q_ϵ.

3). The "general" proof by contradiction.

[Egardt Stability of Model Reference Adaptive + Self Tuning Regulators Springer Verlag 1979, Kreisselmeier + Narendra, IEEE AC Dec. 1982]

<u>Idea</u>: ϕ is UB $\Rightarrow$ no signal grows faster than an exponential. Assume $m \to \infty$. Then for any $M > 0$ arbitrarily large, $m > M$

in a time interval of length $\ln M$

Inside this interval $\int_{t_0}^{t_0+\ln M} \frac{|S_{dy}d|^2}{m} \leq \frac{|S_{dy}d|^2}{M} \cdot \ln M$

Using the Bellman Gronwall lemma — or some Lyapunov function candidate for the closed loop states — we obtain a contradiction, that is m should become smaller than M in $[t_0, t_0+\ln M]$.

<u>"Disadvantage" of the approach</u>: Although this technique can be used for any $\overset{\text{finite}}{\text{size}}$ of bounded disturbances it indicates that m may have to become very large inside some interval. Intuitively, it can be argued that the signals should become suff. large (ϵ_1 large) in order for the signal to noise ratio (ϵ_1/d) to become large and the adaptation to produce good estimates of the controller parameters.

4) Use some internal model design together with a small <u>absolute</u> dead zone.

i.e. $d_z = 0$ if $|\epsilon_1| < M_1$

where, now, $M_1 > |S_{dy} d|$

<u>Disadvantage</u>: $\|d\|_\infty$ should be "small"

<u>Advantage</u>: No stability margin (2δ) needs to be traded-off for bounded disturbances.

<u>CONCLUDING REMARKS IN THE CASE OF BOUNDED DISTURBANCES</u>:

- BOUNDED DISTURBANCES WILL NOT DESTROY GLOBAL BOUNDEDNESS OF THE ADAPTIVE CLOSED LOOP. THE SIGNAL BOUNDS HOWEVER MAY BECOME EXCESSIVELY LARGE EVEN IF THE SIZE OF THE DISTURBANCE IS $O(1)$ (REMEMBER: THE CLOSED LOOP SYSTEM IS NONLINEAR).

- EVEN RELATIVELY SMALL DISTURBANCES MAY PRODUCE BURST PHENOMENA.
- CONTRARY TO UNMODELED DYNAMICS, THE WORST EFFECT OF DISTURBANCES TAKES PLACE IN THE CASES OF INSUFFICIENT (DOMINANT) EXCITATION.

e.g. Consider the regulation case.

When the state of the system is driven to zero, the unmodeled dynamics-terms also go to zero. However, disturbances remain non zero. The adaptation is trying to minimize ϵ_1^2

$$\epsilon_1 = \phi^T w + d$$

w: small $\rightsquigarrow$ ϕ: large so that $\phi^T w + d$: small.

The estimated parameters must be restricted inside a bounded set via σ-modifications / projections or dead zones.

MORE ON DESIGN GUIDELINES

- **Reference Model** $W_M(s)$, **Auxiliary filters** $D(s)$:

  DESIGN IN ORDER TO OBTAIN GOOD CLOSED LOOP SENSITIVITY FUNCTIONS. (NOMINAL PLANT)

  $(1+CP)^{-1}$, $(1+CP)^{-1}CP$

- **INTERNAL MODELS** : HIGHLY RECOMMENDED FOR DISTURBANCE ATTENUATION.

  CAUTION : SHOULD BE IMPLEMENTED WITHOUT AFFECTING THE RELATIVE DEGREE OF THE PLANT.

  e.g. AUGMENT THE PLANT by $\frac{Q_1(s)}{L(s)}$

  $L(s)$: Internal Model
  $Q_1(s)$: Hurwitz Polynomial

  1. $\text{Deg.}(Q_1) = \text{Deg.}(L)$
  2. $\text{Deg}(L)$ = as small as possible in order to keep the dimension of the parameter space small.

- **DIMENSION** OF THE CONTROLLER + PARAMETER SPACE
  (Indirectly determined by the assumed order of P_0)
  In general should be kept low.

TRADE-OFF'S : $(n \downarrow)$

(+) FASTER ADAPTATION ; LOWER EXCITATION REQUIREMENTS ; BETTER BEHAVIOR WRT DISTURBANCES AND A FIXED SIZE OF DYNAMIC UNCERTAINTY.

(−) INCREASED DYNAMIC UNCERTAINTY.

A RULE OF THUMB :

SELECT THE ORDER OF $P_0(s)$ AS TO OBTAIN (FOR SOME PARAMETER VECTOR) A GOOD APPROXIMATION OF $P(s)$ IN THE LOW FREQUENCY RANGE (TYPICALLY RELATED TO THE FREQUENCY CONTENT OF THE REFERENCE INPUT).

- **ADAPTATION**

  - RESTRICT THE PARAMETER SPACE !!!

    σ- modifications / Projections

  - USE EXTERNAL INFORMATION TO OBTAIN A "GOOD" ESTIMATE OF THE PARAMETRIC UNCERTAINTY SET AT THE CURRENT OPERATING CONDITIONS e.g. From physical models, the parameters of the Linear-System Approximation (P_o) may be affected by external signals which are available for measurement. (External temperature, Mach-numbers, dynamic pressure etc) Such information is typically used to determine the parameter settings of a Gain-Scheduling controller. In the adaptive case it can be used to produce an estimate of Θ (easier in the indirect-adaptive control case). Adaptation will then produce a high-integrity design since "successful control" does not rely completely on the information from such sensors as Gain-scheduling does.

- Dead zones : Especially for disturbances (Absolute dead zones). Presently, the only

means available to prevent bursts in live adaptive controllers.
("Live" meaning adaptive controllers whose gain does not go to zero as $t \to \infty$).
Depending on the problem, relative dead zones may still be used but performance may deteriorate considerably.
Dead-zones thresholds should <u>not</u> be too conservative. (Instability is just around the corner!)

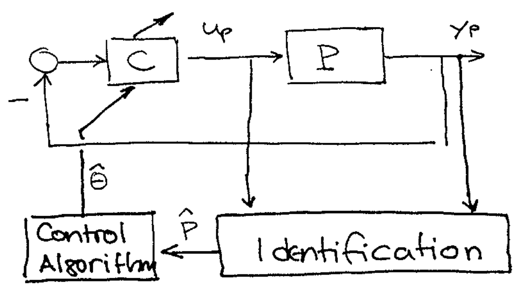
- Normalize signals : Select normalization weights and pole according to the problem. Although normalization will, in general, decrease the speed of adaptation and produce worse transients these problems may be partially fixed by using least-squares types of algorithms rather than simple gradient schemes.

Least squares vs. gradient may speed up parameter convergence by a factor of 100!. Usually, Least squares with covariance resetting and/or Covariance modifications should be used in order to prevent the adaptive gain from going to zero. (Bodson and Sastry) Also, alternative estimation techniques are available to increase the speed of adaptation and shorten adaptation transients e.g. use of multiple models etc. For details see [Narendra + Annaswamy] and references there in. Some Improvements of the overall design may be obtained – depending on the specific problem– by monitoring the level of excitation or the richness of the input signal and switching adaptation on and off accordingly. (Low order nominal plants).

As a final remark the design of adaptive controllers should not be considered as a panacea. It may help to improve performance and stability for plants with large parametric and small non-parametric uncertainty and/or slowly time-varying parameters.. but extensive work needs to be done in order to guarantee that the undesirable effects of adaptation will be avoided.

SOME COMMENTS ON INDIRECT ADAPTIVE CONTROL SCHEMES

The typical block-diagram structure of an indirect adaptive controller is:



The nominal plant parameters p^* are identified using one of the standard estimation algorithms as shown in previous lectures.

For example, starting with the plant description

$$y_p = \frac{N_p}{D} u_p + \frac{D - D_p}{D} y_p + \Delta_1 u_p + \Delta_2 y_p$$

where $\partial D = \partial D_p$

$\frac{N_p}{D_p}$: Nominal Plant transfer function

Δ_1, Δ_2 stable factor perturbations

we may construct the estimation error

$$\epsilon_1 = \hat{p} w - y_p = \tilde{p} w + \eta$$

where : $y_p = p_* w + \eta$

$\hat{p}$: the estimate of p_* ; $\tilde{p} = \hat{p} - p_*$

η : due to $\Delta_1 u_p, \Delta_2 y_p$

w : the states of the auxiliary identification filters $(sI-F)^{-1} q u_p$, $(sI-F)^{-1} q y_p$

with $D(s) = \det(sI - F)$.

The parameters $\hat{p}$ are updated by:

$$\dot{\hat{p}} = \dot{\tilde{p}} = -\gamma d_z P_r \frac{\epsilon_1 w}{m} - \sigma \gamma \hat{p}$$

Your favorite modifications
- dead zones (fixed - relative)
- smooth Projections (soft - hard)

m : Normalizing signal : $\dot{m} = -2\delta_0 m + u^2 + q_m y^2 + 1$
(Egardt, Praly)
s.t. η^2/m is U.B.

As it was previously discussed, this or any similar estimation algorithm (LS - Covariance resetting) guarantees the boundedness of $\hat{p}, \mathring{\hat{p}}$ and

$$\left.\begin{aligned}\int_{t_0}^{t_0+T} \epsilon_1^2/m &\le c + \int_{t_0}^{t_0+T} \eta^2/m \\ \int_{t_0}^{t_0+T} \|\mathring{\hat{p}}\|^2 &\le c + \gamma^2 k \int_{t_0}^{t_0+T} \eta^2/m\end{aligned}\right\} \text{ w/o dead zones}$$

OR

$$\left.\begin{aligned}&\epsilon_1^2/m \le M_1^2 + \varepsilon \; ; \; t \ge t_0 \\ &\mathring{\hat{p}} \in L_2, \; \mathring{\hat{p}} \to 0\end{aligned}\right\} \text{ w/ relative dead-zone}$$

where: M_1 is the dead-zone threshold s.t. $\frac{\eta^2}{m} \le M_1^2 + \varepsilon_t$

ε: arbitrarily small
t_0: suff. large

OR

$$\left.\begin{aligned}&|\epsilon_1| \le d_0 + \varepsilon \; ; \; t \ge t_0 \\ &\mathring{\hat{p}} \in L_2, \; \mathring{\hat{p}} \to 0\end{aligned}\right\} \text{ w/ fixed (absolute) dead zone}$$

where: d_0 is the dead zone threshold s.t. $|\eta| \le d_0 + \varepsilon_t$

ε: arbitrarily small, t_0 suff. large.

(i.e. fixed dead zones should be used with bounded disturbances)

Thus, the plant is effectively described by

$$y_p = \hat{p} w + \epsilon_1$$

or, converting to state space, the plant is described by

$$\begin{aligned}\dot{x} &= A(\hat{p})\, x + b(\hat{p})\, u_p + g_1(\hat{p})\, \epsilon_1 \\ y_p &= c(\hat{p})\, x + g_2(\hat{p})\, \epsilon_1\end{aligned} \qquad (*)$$

which is a time varying system description in terms of the known parameters $\hat{p}$.

Notice that the plant representation (*) is well defined since $\hat{p}, \mathring{\hat{p}}$ are U.B. (due to normalization + projection) and holds irrespective of the boundedness of u_p, y_p.

In other words the identification problem is decoupled from the control problem.

Hence, what remains to be done is to design u_p to stabilize the plant

$$\dot{x} = A(\hat{p})x + b(\hat{p})u_p \;;\; y_p = c(\hat{p})x \quad (1)$$

and guarantee boundedness w.r.t. the perturbations $g_1(\hat{p})\epsilon_1$ and $g_2(\hat{p})\epsilon_1$.

Consider a fixed order compensator

$$\left.\begin{aligned} \dot{w} &= F(\hat{\theta})w + q_1(\hat{\theta})y_p \\ u_p &= q_3(\hat{\theta})w + q_4(\hat{\theta})y_p \end{aligned}\right\} (2)$$

and suppose that we have an algorithm:

" given $\hat{p}$, calculate $\hat{\theta} = f(\hat{p})$;

f : Lipshitz continuous* "

such that the "frozen" closed loop system (1)+(2) is E.S. with stability margin $\delta(\hat{p}) \geq \delta_* > 0$.

*: f differentiable is desired but not necessary

For example, given the plant (1) and $\hat{p}$, the compensator (2) –and $\hat{\theta}$– can be specified by solving a pole-placement or LQ (LQR/KBF) or MRC problem or, even, an H_∞ problem (although the latter requires a lot more computatio which should be performed on line).

The "catch" is that $\hat{\theta}$ must be computable from $\hat{p}$, for any possible value of $\hat{p}$.

Consider for example the plant $P_0 = \frac{\beta}{s+a}$ and suppose that the vector $\hat{p} = \begin{pmatrix}\hat{\beta}\\ \hat{a}\end{pmatrix}$ is updated on line as an estimate of $p_* = \begin{pmatrix}\beta\\ a\end{pmatrix}$. A PPC would then be of the form $u_p = -Ky_p$; $K = \frac{\hat{a} - a_*}{\hat{\beta}}$ i.e. $\hat{\beta}$ MUST be nonzero in order to calculate K.

The available estimation schemes, however, provide no such guarantees. $\hat{\beta}(t) = 0$ for some t is possible to occur or, even, $\hat{\beta}(t) \to 0$ as $t \to \infty$.

In the more general case, $\hat{\theta}$ can be expressed as

$$S_t \hat{\theta} = A_* \leftarrow \text{may be a function of } \hat{p}$$

where S_t is the Sylvester matrix of the estimated plant i.e. S_t depends on $\hat{p}(t)$.

To solve for $\hat{\theta}$, S_t must be nonsingular and, more than that, $|\det(S_t)| \geq c > 0$ $\forall\, t$. (otherwise $\hat{\theta}$ will not be U.B.)

For this to hold the estimated plant must be "strongly" controllable and observable at each time t.

Assuming that $|\det(S_t)| \geq c > 0$, the closed loop can be put in the form

$$\dot{x}_c = A_c(\hat{p}(t)) + b_c(\hat{p}(t))[r]$$

where r are external inputs including

e_1 : the estimation error + uncertainty contributions

A_c, b_c are matrices depending on $\hat{p}(t)$ and $\hat{\theta} = f(\hat{p}(t))$.

and for each fixed time t, the -now- constant matrix $A_c(\hat{p}(t))$ is ES.

It follows (see previous handouts) that the time-varying matrix $A_c(\hat{p}(t))$ will be E.S. if $A_c(\hat{p}(t))$ is UB,

1. $\sup_t \max_i \operatorname{Re}[\lambda_i[A_c]] \leq -\delta_* < 0$

AND

2. $\int_{t_0}^{t_0+T} \|\dot{A}_c\|^2 \leq c + k\mu^2 T$

for suff. small μ.

i.e. $\exists\ \mu_{*1} : \{\mu < \mu_{*1},\ 1, 2\} \Rightarrow$ ES. of the TV matrix A_c

Note that 1 is implied by the assumptions that $|\det S_t| \geq c > 0$ and the stabilizing ppty of the control law while 2 is implied by : n^2/m bounded — and γ suff. small if no dead-zone is used — <u>AND</u> the assumption that $\hat{\theta} = f(\hat{p})$ is Lipschitz in $\hat{p}$.

It is now a quite straightforward procedure to establish boundedness using the B-G lemma, and requiring

$$\int_{t_0}^{t_0+T} \frac{n^2}{m} \le c + \mu^2 T$$

μ : sufficiently small.

(For details, see Middleton et al. "Design Issues in Adaptive Control" IEEE AC 1988).

REMARKS

* In the dead zone case, $\dot{\phi} \to 0$ and the closed loop system behaves more and more as an LTI system as $t \to \infty$. It can be shown that if $\frac{n^2}{m} \le \mu^2 + \epsilon_t$; $\mu < \infty$ and the dead zone threshold is selected strictly greater than μ,

$\hat{p}, \hat{\theta} \to$ constant. That is, the closed loop system can be expressed as an LTI system with an L_2 perturbation due to $\dot{\phi}$ and a state-dependent perturbation $\epsilon_1 \sim \mu\sqrt{m}$.

In this case the robust-stability properties of the closed loop system are determined by those of the frozen (LTI) system 1 + 2 with $\hat{p}_\infty = \lim \hat{p}\,|_{t\to\infty}$. Since $\hat{\theta}$, calculated as $f(\hat{p})$, determines the desired controller for the plant $P_0(s;\hat{p})$, the adaptive controller will be able to tolerate uncertainty of size μ_* s.t.

$$\mu_* \ge \inf_{\hat{p} \in \mathcal{P}} \mu_{LTI}(\hat{p})$$

where $\mu_{LTI}(\hat{p})$ is the uncertainty tolerated by the closed loop of $P_0(s;\hat{p})$ and the corresponding desired LTI controller.

For $\inf\limits_{\hat{p} \in \mathcal{P}} \mu_{LTI}(\hat{p})$ to be nonzero, $\mathcal{P}$, the set of parametric uncertainty in $\hat{p}$, should not contain or be arbitrarily close to points where $P_0(s;\hat{p})$ is uncontrollable or unobservable.

* $\inf\limits_{\hat{p} \in \mathcal{P}} \mu_{LTI}(\hat{p}) > 0$ is a "standard" condition and a typical problem of indirect adaptive schemes (it does not appear in the direct MRAC case where the problem is circumvented by estimating $\hat{\theta}$ directly). Presently, the following solutions are available:

- 1. $p_* \in \mathcal{P}$ and diam $\mathcal{P}$: suff. small.

Since $\left|\det S_t|_{\hat{p}=p_*}\right| > 0$ and $\det S_t(\hat{p})$ is continuous in $\hat{p}$ $\Rightarrow$ $|\det S_t(\hat{p})| > 0$ $\forall \hat{p}$

- 2. If PE is 'available', $\hat{p}$ will converge to a residual set $\mathcal{P}_\infty$ s.t. $p_* \in \mathcal{P}_\infty$ and diam $\mathcal{P}_\infty = O(\mu)$.

$\therefore$ for suff. small μ, $|\det S_t(\hat{p})| \geq c > 0$ $\forall\, t \geq T$, T large enough.

In this case, it can be shown that cl. loop boundedness is preserved by calculating the controller parameters as

$$\hat{\theta} = f(\hat{p}) \;;\; \text{whenever } |\det S_t(\hat{p})| \geq c > 0$$

selected a priori

$$\hat{\theta} = f(\hat{p}_{t_i}) \;;\; \text{whenever } |\det S_t(\hat{p})| < c, \quad t \geq t_i$$

$$\text{and } t_i : |\det S_t(\hat{p})| \geq c$$

(Note that due to PE and for μ suff. small $\exists\, t_i < t_{i+1} < \infty$ s.t. $|\det S_{t_{i+1}}(\hat{p})| \geq c$ and eventually $\exists\, T : t \geq T \Rightarrow |\det S_t| \geq c$).

- 3. Middleton's approach: Use several estimators to estimate $\hat{p}_i$ in several convex closed bounded sets $\mathcal{P}_i$ $i = 1, 2 \ldots N$ s.t. $\exists\, i : p_* \in \mathcal{P}_i$.

$\hat{\theta} = f(\hat{p}_i)$ where i is determined by a suitable criterion s.t. only a finite number of switchings between sets ($\mathcal{P}_i$) will occur (however, N may be large).

see details in Middleton's paper.

OTHER REMARKS

1. TRACKING PERFORMANCE OF INDIRECT SCHEMES

   Use internal models.

2. Bounded Disturbances: As in MRAC case.

   Use internal models + a (small) fixed dead zone.

3. Multivariable Plants (MIMO case).

   The indirect adaptive control case is a straightforward extension of the SISO plant analysis.

   Direct MRAC, however, requires more involved conditions (see refs. in Narendra - Annaswamy).

4. Discrete time Systems.

   A completely analogous analysis can be performed in the case of discrete time adaptive control. An excellent reference for this case is Goodwin + Sin: Adaptive Filtering Prediction and Control, Prentice Hall 86.