Nonlinear Control Theory, the PhD Course 1994

Rantzer, Anders; Bernhardsson, Bo

1994

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Rantzer, A., & Bernhardsson, B. (1994). *Nonlinear Control Theory, the PhD Course 1994*. (Technical Reports TFRT-7519). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
2

ISSN 0280–5316
ISRN LUTFD2/TFRT--7519--SE

Nonlinear Control Theory
The PhD Course 1994

Anders Rantzer
Bo Bernhardsson

Department of Automatic Control
Lund Institute of Technology
June 1994

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
INTERNAL REPORT

Date of issue
June 1994

Document Number
ISRN LUTFD2/TFRT--7519--SE

Author(s)
Anders Rantzer and Bo Bernhardsson

Supervisor

Sponsoring organisation

Title and subtitle
Nonlinear Control Theory. The PhD Course 1994.

Abstract

The course 1994 followed two books: Nijmeijer/van der Schaft's book "Nonlinear Dynamic Control Systems" and Khalil's "Nonlinear Systems". The course was followed by approximately 12 PhD-students. All students had taken an introductory course in nonlinear dynamical systems before.

There was a pronounced desire to study the geometrical approach. We therefore chose to start with the geometrical theory. The two parts of the course are independent so the order could easily be switched.

The impressions of the books are reasonably good. To ease the reading we prepared explanatory notes to be read before each lecture (2 hours). Each lecture was followed by a problem-session (2 hours). There were also 3 take-home problems: 1) Problem Solving, 2) Car-parking with Lie-brackets, 3) Presentation of Nonlinear Article.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
149

Recipient's notes

Security classification

The report may be ordered from the Department of Automatic Control or borrowed through the University Library 2, Box 1010, S-221 03 Lund, Sweden, Fax +46 46 110019, Telex: 33248 lubbis lund.

Summary

The course 1994 followed two books: Nijmeijer/van der Schaft's book "Nonlinear Dynamic Control Systems" and Khalil's "Nonlinear Systems". The course was followed by approximately 12 PhD-students. All students had taken an introductory course in nonlinear dynamical systems before.

There was a pronounced desire to study the geometrical approach. We therefore chose to start with the geometrical theory. The two parts of the course are independent so the order could easily be switched.

The impressions of the books are reasonably good. To ease the reading we prepared explanatory notes to be read before each lecture (2 hours). Each lecture was followed by a problem-session (2 hours). There were also 3 take-home problems: 1) Problem Solving, 2) Car-parking with Lie-brackets, 3) Presentation of Nonlinear Article.

Included in this documentation are

- Course Program
- Errata Nij-vdS
- Errata Khalil
- Session Notes 1-14
- Lecture Slides 1-14 (no figures)
- Article on KYP-Lemma (Rantzer)
- Article on IQC's (Rantzer/Megretski).

Other material that we used:

- Sastry, Nonlinear Systems, pp 288-314
- Glad, T. Nonlinear Control Theory, Chaps 3,6,7
- Willems, 1972, Dissipative Dynamical Systems
- The LMI-lab manual

Lecture 8 on mechanical systems was given by Rolf Johansson and lecture 9 on Volterra series by Sven Spanne. Their contributions are gratefully acknowledged. We also thank S. Sastry and T. Glad for generously sharing unpublished text material.

Lund, June 1994

Bo Bernhardsson, Anders Rantzer

NONLINEAR CONTROL THEORY, 1994

Lecturers

Bo Bernhardsson, Anders Rantzer. Special guest stars.

Literature

— H. NIJMEIJER AND A.J. VAN DER SCHAFT, *Nonlinear Dynamical Control Systems.* Springer Verlag, Englewood Cliffs, NJ, 1991, ISBN 3-540-97234-X, 2nd printing.

— H. K. KHALIL, *Nonlinear Systems*, McMillan Publishing Co., 1992, ISBN 0-2-363541-X.

— Notes, journal papers and other material.

Meetings

There will be one lecture and one exercise per week (2+2 hours). Participants are supposed to prepare for the lectures by reading ahead in the book and to take active part of the exercises.

Lecture Plan:

0 Introduction. Inspiration.
1 Examples. Manifolds. Implicit Function Theorems. Tangent Vectors.
2 Vector Fields. Lie Brackets. Distributions. Frobenius' Theorem.
3 Local Controllability.
4 One-forms. Codistributions. Observability. Nonlinear Kalman Decomposition.
5 State Space Transformations. Canonical Forms. Exact Linearization SISO. Zero Dynamics.
6 Exact Linearization MIMO. Disturbance Decoupling. I/O Decoupling
7 Center Manifold Theory. More on Zero Dynamics.
8 Mechanical Nonlinear Control Systems.
9 Volterra Series.
10 Lyapunov Theorems. Center Manifolds. Regions of Attraction. Nonautonomous Systems.
11 Absolute Stability. KYP Lemma. Circle/Popov Criteria. Simultaneous Lyapunov Functions.
12 Dissipativity. IQCs. Slowly Varying Systems.
13 Oscillations. Describing Functions. Perturbation Theory I.
14 Periodic Perturbations. Averaging. Singular Perturbations.

Examination

3 hand-in problems + take-home exam.

Points

Nominally 8p.

Errata list to Nijmeijer-van der Schaft, 2nd Ed

page		error	correction
5	line 11	$\dot{\varphi}$	$\dot{\phi}$
15	(1.46)	$\dot{\theta} = u$	$\dot{\theta} = -u$
25	-9	p for M_2	p for M_1
32	14	N	P (three places)
36	13	(2.32)	(2.35)
49	(2.96)	X_j	Y_j
51	-2	X_*^{-1}	X_*^{-h}
53	(2.108)	The vector field and Fig. 2.8 are incompatible	
58	-4	$[X_1, X_2] = X_2$	$[X_1, X_2] = -X_2$
59	-1	by Theorem 3.36	by Theorem 2.36
60	(2.140)	change to $\frac{\partial h}{\partial r_\alpha} + \sum_{j=1}^{n} b_{j\alpha}(r,s)\frac{\partial h}{\partial s_j}(r,s)$	
63	-3	$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$	
69	-10	The reference [So] does not exist in the reference list.	
70	2	The references [B] [S] do not exist in the reference list.	
71	Ex. 24b,c	Wrong formulation.	
71	-7	y	Y
77	-8	Should be $g_2(x_0 + hg_1) = g_2(x_0) + h\frac{\partial g_2}{\partial x_0}(x_0)g_1 + \ldots$	
79	7	$k = 0, 1, 2, \ldots$	$k = 1, 2, \ldots$
79	11	Should be "the number of Lie-brackets in it +1."	
91	14	$AB_j - AB_j$	$AB_j - B_jA$
95	-4	Since dim $\mathcal{O}(x_0) = n$	Since dim $d\mathcal{O}(x_0) = n$
97	9	$i \in \underline{p}$	$i \in \underline{k}$
99	-1	(2.154)	(2.164)
101	-7	(3.92)	(3.99)
106	6	CA^{i-1}	CA^i
106	Th. 3.47	system (3.1)	system (3.69)
106	-8	invariant (3.1)	invariant for (3.1)
106	-3	(2.161)	(2.168)
107	Th. 3.49	system (3.1)	system (3.69)
119	6	$x(t, s, x, u)$	$x(t, s, x, 0)$
120	(4.16),(4.18)	missing)	
121	(4.27)	w_{i-1}	w_{k-1}
121	(4.27)	w_i	w_k
121	-11	Then it can checked	Then it can be checked
143	-6	The paper on its turn	The paper in its turn
172	[Kr]	The journal changed name in 1976, before: SIAM J. Contr.	

page		error	correction
209	13	$z_1 = \dots$	$z_2 = \dots$
218	7	(7.38)	(7.37)
253	9	A) is missing	
258	-6	A (is missing	
301	-7	Let (x_0, u_0) an	Let (x_0, u_0) be an
302	-6	af the matrix	of the matrix
307	Ex 10.10	Why use I here when J was	used before.?
310	-8	Should be $z = T(x - x_0)$?	
316	(10.75)	$\dots + b_1\varphi_2\bar{x}_1^3$	$\dots + b_1 p_1 \varphi_2 \bar{x}_1^3$
319	-1	system	systems
334	(11.37)	Should end with $k_j z^j$	
338	(11.53)	$l_1^1 \sin\theta_2^1 + \dots$	$l_1^1 \sin\theta_1^1 + \dots$
360	(12.48)	Change (to {	
361	-12	$X\{q_i, q_j\}$	$X(\{q_i, q_j\})$
361	-13	$X\{p_i, p_j\}$	$X(\{p_i, p_j\})$
363	-9	$X_H, \dots$	$X_{H_0}, \dots$
364	8	en	and
368	10	$\{H_0, \mathcal{F}_O)$	$\{H_0, \mathcal{F}_O\}$
392	15	The reference [TA] does not	exist in the reference list.
392	-7	symetries	symmetries
393	[HvdS]	This reference must have occured now	
439	4	and	an

Discussion: Should probably include the formula $(F \circ G)^* = G^* \circ F^*$.

Why not define local observability at x_0 a little sharper:

$$x_1 \mid^V x_2 \text{ for all } x_1, x_2 \in W \subset V \ni x_0?$$

The main theorem works also with this definition.

In (5.36) I think it should read $k + l = 0, \dots, \kappa_i + \kappa_j - 2$? The same in (5.38). But (5.39) is correct, there it must be $2n - 1$. There should be some more discussion of how (5.36) and (5.39) are obtained.

Please also check (5.69) and (5.78).

I think problem 5.10 is wrong, even in the linear case. If you add an integrator before a SISO system $b(s)/a(s)$ you lose observability iff $b(0) = 0$. I dont see how to change problem 5.10 in the nonlinear case so that it becomes correct.

Errata

NONLINEAR SYSTEMS

Hassan K. Khalil
Michigan State University

The textbook Nonlinear Systems was published in June, 1991. Since then it has been used by the author and a few colleagues in classrooms at a number of universities, and some inadvertent errors have been detected. This errata sheet corrects these errors.

- Page ii: change 'Koktovic' to 'Kokotovic'.
- Page 3, Line 1: change 'the the' to 'the'.
- Page 3, Line 13: change 'the roots' to 'the real roots'.
- Page 15, Line 10: change to If $h(\cdot)$ is differentiable ...
- Page 18, Line 18: change Section 5.3 to Section 5.2.
- Page 38, Line 3 from the bottom: change 'a node' to 'a node with distinct eigenvalues'. A node with multiple eigenvalues could become a focus after a small perturbation, although it will keep its stability type; e.g., a stable node could become a stable focus.
- Page 39, Line 7 of the 2nd paragraph: change 'eigenvalues' to 'eigenvalue'.
- Page 46, Line 5 of the 2nd paragraph: change 'a node' to 'a node with distinct eigenvalues'.
- Page 51, Figure caption: change 'negative' to 'positive'.
- Page 54, Figure 1.37 (b): reverse the arrow heads.
- Page 60, Exercise 1.14: change 'For each for' to 'For each of'.
- Page 60, Exercise 1.15: change 'For each for' to 'For each of'.
- Page 65, Line 6 of the 2nd paragraph: change 'eigenvalue' to 'eigenvalue of'.
- Page 66, Line 4: change '$F : R^n \to R^m$ is differentiable' to '$F : R^n \to R$ is continuously differentiable'. The mean value theorem is valid only for a scalar function of a vector argument. It was erroneously stated for a vector function of a vector argument. At every point in the book where the mean value theorem is used, it should be applied component wise, leading to the same conclusions. Corrections will be made at the respective pages.
- Page 71, Line 2 from the bottom: change $x_k(t) - x$ to $x_k(t) - x(t)$.
- Page 72, Line 4 of the Proof: change $\leq$ to $=$.
- Page 77, Line 5: correct the spelling of the third occurrence of Lipschitz.
- Page 77, Line 2 from the bottom: change $\|\partial f/\partial x\|$ to $\|\partial f/\partial x\|_\infty$.

- Page 78, change the first three lines to:
  z_i on the line segment joining x and y such that

$$|f_i(t,x) - f_i(t,y)| = \left|\frac{\partial f_i}{\partial x}(t,z_i)(x-y)\right| \le L_0\|x-y\|_\infty$$

  Hence

$$\|f(t,x) - f(t,y)\|_\infty \le L_0\|x-y\|_\infty$$

- Page 85, Line 4 from bottom: change '$\le$' to '$=$'.
- Page 91, Exercise 2.2: on the left-hand side of the second inequality, replace $\frac{1}{n}$ by $\frac{1}{\sqrt{n}}$. On the left-hand side of the third inequality, replace $\frac{1}{m}$ by $\frac{1}{\sqrt{m}}$.
- Page 94, Exercise 2.2, Line 2: change $f(x)$ to $f(t,x)$.
- Page 101, last line of the text and first line of the footnote: change 'entirely inside' to 'in the interior of'.
- Page 104, Line 3: replace the phrase 'which is ... (nonnegative).' with a new sentence 'Also, $V(x)$ is positive definite if and only if all the leading principal minors of P are positive'. The corresponding statement for positive semidefinite matrices is only necessary.
- Page 110, Line 3 of the second paragraph: change 'converges' to 'converge'.
- Page 115, Line 9: change $\frac{1}{2}k$ to $1/2k$.
- Page 116, First line of Lemma 3.1: change 'bounded' to 'bounded and belongs to D'.
- Page 117, Line 2: change Ω to $\Omega \subset D$.
- Page 120, Line 9 from the bottom: change R^n to R^2.
- Page 121, last line: insert the integral sign, $\int_0^{x_i}$.
- Page 129: change the last equation to

$$f_i(x) = f_i(0) + \frac{\partial f_i}{\partial x}(z_i)\,x$$

  where z_i is a point ....
- Page 130: change the first eight lines to
  x to the origin lies entirely in D. Since $f(0) = 0$, we can write $f_i(x)$ as

$$f_i(x) = \frac{\partial f_i}{\partial x}(z_i)x = \frac{\partial f_i}{\partial x}(0)x + \left[\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right]x$$

  Hence

$$f(x) = Ax + g(x)$$

where

$$A = \frac{\partial f}{\partial x}(0), \quad and \quad g_i(x) = \left[\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right] x$$

The function $g_i(x)$ satisfies

$$|g_i(x)| \leq \left\| \frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0) \right\| \|x\|$$

- Page 132, Line 18: change 'in unstable' to 'is unstable'.
- Page 138, Line 11: change $x \in R^n$ to $x \in D$.
- Page 138, Example 3.16, second equation: change $\dot{x}_1$ to $\dot{x}_2$.
- Page 149, Line 3 of the footnote 18: change 'it continuous' to 'it is continuous'.
- Page 153, Exercise 3.6, part (a): change $x \in R^2$ to $x \in R^2$, $x \neq 0$.
- Page 154, Exercise 3.10: change 'a function $V_1(x)$' to 'a continuously differentiable function $V_1(x)$'.
- Page 158, Line 4: change 'differentiable' to 'continuously differentiable'.
- Page 159, Exercise 3.23: the set G is a simply connected domain containing a neighborhood of the origin.
- Page 173, Lines 1 and 2: remove 'piecewise'.
- Page 177, Line 2 from the bottom: change 'a Lyapunov' to 'Lyapunov'.
- Page 178: In line 10 replace L by $\frac{L}{\sqrt{n}}$. In lines 10, 12, 13, 15, and 16, replace f and z by f_i and z_i, respectively. In line 14, replace $f(t,x)$ by $f_i(t,x)$. Replace lines 17 to 22 by
  Hence
  $$f(t,x) = A(t)x + g(t,x)$$
  where
  $$A(t) = \frac{\partial f}{\partial x}(t,0) \quad and \quad g_i(t,x) = \left[\frac{\partial f_i}{\partial x}(t,z_i) - \frac{\partial f_i}{\partial x}(t,0)\right] x$$
  The function $g(t,x)$ satisfies
  $$\begin{aligned} \|g(t,x)\|_2 &\leq \left(\sum_{i=1}^{n} \left\| \frac{\partial f_i}{\partial x}(t,z_i) - \frac{\partial f_i}{\partial x}(t,0) \right\|_2^2 \right)^{1/2} \|x\|_2 \\ &\leq L\|x\|_2 \end{aligned}$$
- Page 195, line 16: change 3.06 to 3.026.

- Page 195, Line 9 from the bottom: change 'a closed' to 'closed'.
- Page 199, Line 8: change $\leq$ to $=$.
- Page 202, Definition 4.4: b, c, and T are positive constants.
- Page 203, Line 13: remove '$t_0(....)$ or '.
- Page 210, Line 4 from the bottom: change ΔF and f to ΔF_i and f_i.
- Page 218, Line 7: change $1 < k < n$ to $1 \leq k < n$.
- Page 222: change the footnote to 'The function $\nu(y, w)$ is continuously differentiable every where around the origin, but not at the origin itself. It can be easily seen that the statement of Theorem 3.1 is still valid'.
- Page 235, Line 8: change the forth term of $V(x)$ from $x_1 x_2$ to $x_1 x_3$.
- Page 282, Line 16: change 'form a' to 'is a'.
- Page 286, Line 3: change $\frac{3}{2}\|z\|_2^2$ to $\frac{\sqrt{10}}{2}\|z\|_2$.
- Page 286, Line 3: change $\|z\|_2^2$ to $\frac{\sqrt{5}}{2}\|z\|_2$.
- Page 292, Line 18: change $-a \leq x_2 \leq a$ to $-a < x_2 < a$.
- Page 294, Line 12: change the first term on the right-hand side from $MAM^{-1}x$ to $MAM^{-1}z$.
- Page 295, First equation: change the last element of the right-hand side column from $\alpha(T(y))/\beta(T(y))$ to $-\alpha(T(y))/\beta(T(y))$.
- Page 295, the stack of equations in the middle: change the last equation from
$$\frac{\partial T_n}{\partial y} f(y) = \alpha/\beta \quad to \quad \frac{\partial T_n}{\partial y} f(y) = -\alpha/\beta$$
- Page 296, Equation (5.61): change the numerator of the expression for α from $(\partial T_n/\partial y)f(y)$ to $-(\partial T_n/\partial y)f(y)$.
- Page 298, Line 8, the equation for α: multiply the right-hand side on the first line by a minus sign and remove the minus sign from the denominator of the second line.
- Page 315: The Kalman-Yakubovich-Popov lemma is proved in the book only for minimal realizations. In the adaptive control problem, the realization (A_m, b_m, c_m^T) could be nonminimal. An extension due to Miller (see [83, Section 2.6]) shows that the last two equation on page 315 are valid for nonminimal realizations.
- Page 326, Lines 1–2: change Theorem 4.8 to Theorem 5.8.
- Page 327, Exercise 5.1: change $Z^T(-s)$ to $Z^T(s^*)$.
- Page 332, Exercise 5.16: change the inequality satisfied by V to
$$c_{i1}\|x_i\|^2 \leq V_i(t, x_i) \leq c_{i2}\|x_i\|^2$$
- Page 332, Exercise 5.19: change the lower limit of the summation from $j = 1$ to $j = 1; j \neq i$.

- Page 338, Line 3, Part (b): change $\rho_1 = \sqrt{2}$ to $\rho_1 = 2\sqrt{2}$.
- Page 338, Exercise 5.38: change $\psi(x)$ to $\psi(t, x)$.
- Page 356, Figure 6.7: change $\Psi(\sin\theta)$ to $\Psi(a\sin\theta)$.
- Page 388, Line 2 of the 2nd paragraph: change 'Tayolr' to 'Taylor'.
- Page 395, Line 15: change $e_0(t, \epsilon)$ to $e_0(t)$.
- Page 404, Line 5: change $r_0 < r$ to $r_0 \leq r$.
- Page 429, Line 4 from the bottom: change $O(\alpha(\epsilon))$ to $O(\alpha(\epsilon))$ close.
- Page 446, Line 14: remove the sentence 'The origin $y = 0$ is also an equilibrium point for the nonautonomous system (8.12).'
- Page 447, Equation (8.16): change $-c$ to $-c < 0$.
- Page 449, Lines 7, 9, 12, and 14: change $\eta_0 - \zeta_0$ to $\eta_0 + \zeta_0$.
- Page 450, Line 12: $\gamma = \frac{\tan^{-1}\rho}{\rho}$.
- Page 452, Line 4 from the bottom: change 'the the' to 'the'.
- Page 462, equation (4.38): change $\zeta_1(y)$ and $\zeta_2(y)$ to $\zeta_1(\|y\|)$ and $\zeta_2(\|y\|)$, respectively.
- Page 463, Line 4 from the bottom (the line preceding 'where'): change $\leq$ to $=$.
- Page 471, Line 8: change)) to).
- Page 472, Lines 9 and 12: change $x(t, \epsilon) - \bar{x}(t)$ to $\|x(t, \epsilon) - \bar{x}(t)\|$.
- Page 474, Line 17: change 'a given' to 'given'.
- Page 474, Line 6 from the bottom: change y_p to y_m.
- Page 475, Line 20: change $k_p(z - h(t, x))$ to $K(z - h(t, x))$ where $K = [k_p, 0, 0]^T$.
- Page 481, Exercise 8.19: change $\psi(x)$ and $\phi(y)$ to $\psi_1(x)$ and $\psi_2(y)$.
- Page 483, Line 3: change 'current' to 'the current'.
- Page 488, Figure A.2: change c_0 to c_0^*.
- Page 490, Line 10: change $(sI - A_m)$ to $(sI - A_m)^{-1}$.
- Page 491, Line 11: change 'sequences $\{\tau_i\}$' to 'sequence $\{\tau_i\}$'.
- Page 495, Lines 3 and 5: change $\geq$ to $>$.
- Page 498: replace lines 2 to 4 by

  $$H(s) = \frac{\exp[-\eta^{-1}(s)]}{h(\eta^{-1}(s))}, \quad s \geq 0$$

  Since η^{-1} is continuous and h is positive, $H(s)$ is continuous on $0 < s < \infty$, while $\eta^{-1}(s) \to \infty$ as $s \to 0^+$. Hence, $H(s)$ defines a class $\mathcal{K}$ function on $[0, \infty)$.
- Page 507: change Line 6 to

  $$\pi_i(t, y, \eta) = \pi_i(t, y, 0) + \pi_{iy}(t, \zeta)y = y_i + \pi_{iy}(t, \zeta)y$$

- Page 508, Line 5 from the bottom: change g_1 to g_{1i}.
- Page 508, Line 2 from the bottom: change $(k_1 + k_2k_3)$ to k_4.
- Page 524: change equation (B.29) to

$$|\delta\Psi| = \|P_1 g\psi y_1 - P_1 g\psi(y_1 + y_h)\| / \sqrt{2}a|G(j\omega)|$$

- Page 527, Line 4 from the bottom: change $d[\eta(\cdot,\mu), p, D]$ to $d[\eta(\cdot,\mu), D, p]$.
- Page 533, Lines 1 and 5: change $\leq$ to $<$.
- Page 533: change the second term on the right-hand side of the line to $[\theta_1\epsilon(t_1 - t_0) + \frac{\theta_2\epsilon\mu}{\alpha}]$.
- Page 535, Line 5: change G to G_i.
- Page 536, Line 6 from the bottom: change the left-hand side to $\|v(\tau)\|$.
- Page 542, Line 3 from the bottom: change 'Coreless' to 'Corless'.
- Page 544, Line 1: change 'Coreless' to 'Corless'.
- Page 558, Line 4 from the bottom: change 'end' to 'end of'.

Hassan Khalil
May, 1992

Session 1

Examples. Implicit Function Theorems. Manifolds. Tangent Vectors.

The course will focus on *analysis* of nonlinear control systems. The first half of the course concentrates on the geometric approach. This is the hardest part of the course and presents several new concepts. To be able to read modern research articles in nonlinear control it is essential to understand the mathematical vocabulary, such as manifolds, vector fields, Lie brackets etc. It is not so difficult if you work with examples at the same time as you read the material. It will help if the material from Spanne's nonlinear systems course is fresh in mind.

Reading Assignment

Nijmeijer pp. 1-43 (until 2.2). Concentrate on the manifold and tangent vector concepts in Ch. 2.

Chapter 1.

The geometric theory has mainly been developed for systems of the form

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \tag{1}$$

where f, g, h are smooth functions. Many nonlinear systems, but of course not all, are of this form. Read Ch. 1 very briefly, don't get stuck in the derivations of the equations. You can skip examples 1.3 and 1.5. TS^1 means 'the tangent space to the circle', something we will describe in Ch. 2. (1.12)-(1.19) follows from mechanics, you can read more in e.g. Craig's Robotics book. Fig 1.6 illustrates the circle divided into two overlapping sets. Important in Ch. 1 is only the structure of the resulting equations (1.6), (1.14), (1.19) and (1.35).The author's goal in the chapter is to motivate

- Models affine in control, see (1).
- More general state spaces than R^n (so called 'manifolds').

You should be familiar with the existence and uniqueness theory of ODEs on p.11-13, see Spanne's course otherwise. Stop reading after equation (1.49), start near (1.50) again. I don't know why the authors have to scare away readers in the introduction by this fiber bundle stuff. We will understand this only much later. The rest of Ch. 1 is probably known to you.

Chapter 2.

This is a tough chapter but all work here is well spent and will pay off later. There are many new concepts and it is important to get started with exercises quickly. Don't wait until you have read the entire chapter. If you find the survey section 2.0 hard, go directly to 2.1.

The idea with manifolds (svenska: mångfalder) is to give a generalization to the concept of 'state space'. This is done by piecing together small pieces of R^n. Your friends in finding your way around the manifold M are the charts φ_i. These will help us define 'smooth' state-transformations, derivatives of states (so called tangent vectors) etc. Draw a lot of pictures when reading this material. Note that manifolds will be finite dimensional in this course, a generalization is needed if one wants to study systems with time delays.

To define a topology on a set M means to decide which subsets will be called 'open'. This indirectly defines e.g. which functions are called continuous. Different topologies can be useful at different times. Often the open sets are defined using 'distance functions' $d(x, y)$ (metrics). The open sets are then those sets U that are such that 'given any point x in U there is an $\varepsilon > 0$ so that all points y with $d(x, y) < \varepsilon$ are inside U'. Not all topologies are 'metrizable'. That is why one needs the more general definition.

When you have understood defs. 2.1 and 2.4 the first main step is over. Theorems 2.11 and 2.12 are very important and you should get experience in using them. Theorem 2.13 shows the way manifolds are constructed most oftenly, see Ex. 2.14-5. The proof of 2.13 is a good test that you have understood Th. 2.12.

Read the discussion on functions between manifolds coordinate charts and submanifolds carefully, but dont get stuck in the details in immersions, submersions.

The next thing is to define state derivatives $\dot{x}(t)$ on the manifold M. This is nontrivial since the only thing we know about M is through the looking-glasses of the charts $\varphi_i(x)$. One should think of a tangent vector as a pair (x, v) or v_x where x is a point on the manifold and v is a tangent to the manifold at that point. To get a coordinate free formulation we define the state derivatives, tangent vectors, as abstract objects. This can be done in several ways. Nijmeijer uses one of the most common methods: A tangent vector (x, v) is an object that derivates real-valued functions $f(x)$ on the manifold, i.e. at the point x. One should think of $v_x(f)$ as the directional derivative of f in the direction of v at the point x. Note that the remark on p.38 really don't make sense unless M is an embedded manifold since $\dot{c}(0)$ is not defined. Draw a picture for (2.41). F_* is sometimes called the push-forward operator. In coordinates it's a Jacobian: $F_* = (\,\partial f_i / \partial x_j\,)$. Equations (2.54)-(55) are the important coordinate-versions of the tangent vector.

Exercises Exercises marked with a "*" are either difficult or not so central.

Exercise 1.1 Show that the sphere $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ has the structure of a manifold with atlas, for example, consisting of two charts $(U_i, \phi_i, i = 1, 2)$ in stereographic projection.

Exercise 1.2 Consider a car with N trailers. The front-wheels of the car can be controlled, and the car can drive forwards and backwards. Describe a manifold that can be used as state-space. Show that its dimension is $N + 4$.

Exercise 1.3 * = Nij 2.1.

Exercise 1.4 Use the implicit function theorem to show that under a certain condition, determine which, the root-locus, i.e. $\{s : a(s) + kb(s) = 0\}$, locally is a function of k. Determine a differential equation the branches $s_i(k)$ of the root locus satisfy.

Exercise 1.5 = Nij. 2.2

Exercise 1.6 = Nij. 2.3

Exercise 1.7 = Nij. 2.4a. Find counter-examples to 2.4b,c. Try to reformulate 2.4b,c so that they become correct.

Exercise 1.8 Let $f_1(x), \ldots, f_n(x)$ be n vector fields that are smooth and linearly independent around $x_0 \in R^n$. Let $\Phi^t_{f_i}(x)$ be the corresponding transformation groups (defined at least for small t). Show that the transformation

$$X(t_1, \ldots, t_n) = \Phi^{t_n}_{f_n} \circ \ldots \circ \Phi^{t_1}_{f_1}(x_0)$$
$$\widetilde{X}(t_1, \ldots, t_n) = \Phi^1_{t_n f_n + \ldots + t_1 f_1}(x_0)$$

both defines smooth bijections between a neighborhood U of 0 and a neighborhood of x_0 (i.e. they are diffeormorphisms). These are nice ways to change coordinates (from x_i:s to t_i:s).

Exercise 1.9 Transform $X_p = x_2 \frac{\partial}{\partial x_1}$ to polar coordinates (r, φ). Calculate $X_p(f)$ when $p = (x_1, x_2) = (1, 1)$ and $f(x_1, x_2) = x_1^2 + x_2^2$.

Exercise 1.10 Consider the set $M = (x_1, x_2) = (t, t^2),\ t \in R$. Show that M is a one-dimensional submanifold of R^2. Show that a basis for TM is

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial x_1} + 2x_1 \frac{\partial}{\partial x_2}.$$

Exercise 1.11 Let

$$X_1 = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + x_3 \frac{\partial}{\partial x_3}$$

$$X_2 = \frac{\partial}{\partial x_3}$$

Change coordinates to $z_1 = \ln(x_1/x_2)$; $z_2 = x_2$; $z_3 = x_3$, where $x_i > 0$, and show that span $(X_1, X_2) =$ span $(\frac{\partial}{\partial z_2}, \frac{\partial}{\partial z_3},)$

Session 2

Vector Fields. Lie Brackets. Distributions. Frobenius Theorem.

Reading Assignment

pp. 43-60 + Prop. 3.6.

Chapter 2, continued

A vector field on a manifold is the counterpart to an ODE in R^n. To every point x on M is given a tangent vector X_x that starts in that point. Now we of course want to go from the equation $\dot{x} = X(x)$, or what is the same in our terminology, the vector field $X = \sum X_i(x)\frac{\partial}{\partial x_i}$, to solution curves $x(t)$ via (2.74). A little thought gives that equation (2.80) follows. Eq. (2.81) is the point of everything developed so far. Don't bother if you don't understand the discussion about fiber bundles. One example where a separation between state space and control space, $M \times U$, is impossible was given in Example 1.7 on page 17. Equation (2.84) is the starting point for the rest of the geometric theory.

In the good old days Theorem 2.26 was shown in Spanne's course. The reason it's not enough for us is that we have control signals. In what follows we will generalize vector fields to something called 'distributions' to cover this.

The concept of Lie-brackets is central. The reason, vaguely speaking, is that if it's possible to control a system in the direction of both $g_1(x)$ and $g_2(x)$, then one can also drive it in the direction of $[g_1, g_2]$ (approximately). This will be explained more exactly later. Equations (2.94), (2.96) say how you compute Lie-brackets in practice. Do a couple of such calculations now. In parallel with page 49-51 you should read Prop. 3.6, which illuminates Theorem 2.3.3, draw a picture. Go through Ex. 2.35 carefully.

Dont confuse the distributions in 2.2.2 with the unrelated Schwartz' distribution theory. A distribution will here be a generalization of a vector field. It models "all the directions you can control in by varying the control signal u". Mathematically it is the linear span of a number of vector fields.

$$X(q) = \text{span}\{X_i(q)\} = \sum \alpha_i(q)X_i(q)$$

For example span$\{\frac{\partial}{\partial x_1}, \sin(x_1)\frac{\partial}{\partial x_2}\}$ This distribution does not have constant dimension.

Involutivity of a distribution means that we can not generate more vector fields by forming Lie-brackets between two vector fields in the distribution. From a control point of view it means that the distribution already describes 'all directions you can control in'.

An integral submanifold P to a distribution D is a generalization of the concept of solution to an ODE, just like distributions were generalizations of vector fields. In every point of the integral submanifold the tangent space should equal the distribution $TP(q) = D(q)$.

It is usually impossible to find an integral submanifold to a given distribution, it can be done if and only if the distribution is involutive. This is Frobenius theorem, which is classical in the theory of first order PDEs. Prop. 2.41 is the easy half of the implication *involutivity* $\Leftrightarrow \exists$ *integral submanifold.* This equivalence is used extensively in geometric nonlinear control theory. The integral submanifolds generalize the notion of invariant subspaces in linear systems like the controllability subspace and nonobservability subspace (as we will see in Ch. 3).

Skip the remark on page 58. Skip Theorem 2.45. Start reading again at p. 60 bottom, where $D_1 + D_2$ and $D_1 \cup D_2$ are defined.

Hand In Problem 1: Hand in solutions to 10 of the following exercises.

Exercise 2.1 Let $X(x) = x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}$. Write the vector field as a set of differential equations. Find $X^t(x_1, x_2)$. Calculate $X^t_*\frac{\partial}{\partial x_1}|_{(1,0)}$.

Exercise 2.2 = Nij. 2.6.

Exercise 2.3 The SISO system $\dot{x} = f(x) + g(x)u; \quad y = h(x)$ is said to have relative degree r at x_0 if

$$\begin{aligned} L_gL_f^ih(x) &\equiv 0 \quad \text{in a neighborhood of } x_0 \quad i = 0, \ldots, r-2 \\ L_gL_f^{r-1}h(x_0) &\neq 0 \end{aligned}$$

Calculate the relative degree for

$$\dot{x} = \begin{pmatrix} x_2 \\ -x_2 - x_1^3 - x_1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$
$$y = x_1 \quad \text{or} \quad y = x_2$$

Show also that this definition is consistent with the usual definition of relative degree for linear systems (as being the excess of poles over zeros). Can relative degree be changed by feedback $u = \alpha(x) + \beta(x)v$?

Exercise 2.4 * Show that the following conditions are equivalent:

$$\begin{bmatrix} L_gL_f^kh(x) \equiv 0 \\ 0 \leq k \leq \mu \ \forall x \in U \end{bmatrix} \Longleftrightarrow \begin{bmatrix} L_{ad_f^kg}h(x) \equiv 0 \\ 0 \leq k \leq \mu \ \forall x \in U \end{bmatrix}$$

Here $ad_fg = [f, g]$ and $ad_f^kg := [f, ad_f^{k-1}g]$ for $k \geq 2$.

Exercise 2.5 Consider the system

$$\dot{x} = \begin{pmatrix} x_2 - 2x_2x_3 + x_3^2 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix} u$$

Change coordinates using $z = S(x) = \begin{pmatrix} x_1 + x_2^2 \\ x_2 + x_3^2 \\ x_3 \end{pmatrix}$.

Exercise 2.6 * Let S^{2n-1} be the submanifold of R^{2n} generated by

$$x_1^2 + x_2^2 + \ldots + x_{2n}^2 = 1$$

(Comment for the mathematician: the topology and differentiable structure are of course the ones generated in the natural way by that in R^{2n}). Show that the following defines a nowhere vanishing vector field on S^{2n-1}:

$$\sum_{i=1}^{n} x_{2i}\frac{\partial}{\partial x_{2i-1}} - x_{2i-1}\frac{\partial}{\partial x_{2i}}$$

("It is possible to comb the hair on the S^{2n-1}-spheres.")

Exercise 2.7 Find a coordinate chart around $x = [0\ \ 0]^T$ so that $\frac{\partial}{\partial z_1} + 2z_1\frac{\partial}{\partial z_2}$ is transformed as in Theorem 2.26 to $\frac{\partial}{\partial x_1}$.

Exercise 2.8 = Nij 2.7.

Exercise 2.9 = Nij. 2.9.

Exercise 2.10 * Can you find matrices A, B so that $[A, B] = I$?

Exercise 2.11 Show that $sl(n)$, the $n \times n$ matrices of trace 0, form a Lie subalgebra.

Exercise 2.12 = Nij. 2.10. Show also that all such (symplectic) matrices are of the form

$$A = \begin{pmatrix} X & R \\ Q & -X^T \end{pmatrix}$$

where R and Q are symmetric.

Exercise 2.13 = Nij. 2.11.

Exercise 2.14 = Nij 2.8.

Exercise 2.15 = Nij. 3.1ab.

Exercise 2.16 Check that the spheres $x_1^2 + x_2^2 + x_3^2 = a^2$ are integral manifolds of the two-dimensional distribution D spanned by

$$
\begin{aligned}
X_1 &= x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \\
X_2 &= x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1} \\
X_3 &= x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2}
\end{aligned}
$$

Check also that D is involutive.

Exercise 2.17 Assume $z = f(x, y)$ is an integral manifold to the distribution spanned by

$$
\begin{aligned}
X &= \frac{\partial}{\partial x} + g(x, y, z) \frac{\partial}{\partial z} \\
Y &= \frac{\partial}{\partial y} + h(x, y, z) \frac{\partial}{\partial z}
\end{aligned}
$$

Show that $[X, Y] = 0$.

Exercise 2.18 * Discuss the possibility to write a numerical routine that computes the coordinate charts mentioned in Frobenius theorem given an involutive distribution. Can it be done using ODEs or does one need PDE-software?

Session 3

Local Controllability.

Reading assignment

pp. 73-93 (to Sect. 3.2) + 101-104.

Chapter 3

Now we start working with control theory. We first discuss nonlinear controllability. The theory is here quite satisfactory but much remains to be done from a practical point of view. One point is that the present theory only gives yes/no answers to controllability questions. A personal remark is that as an engineer one would like to have more quantitative measures of controllability. Remember that in the linear case the rank of the controllability matrix answers the controllability question but that it is really the singular values of the controllability Gramian P that are more interesting. This is important in i.e. model reduction. Many mathematical linear system concepts have been generalized to nonlinear control but the engineering side, like numerical algorithms, is lagging behind.

Note first that in the nonlinear case there are several different controllability concepts. In the linear (continuous time) case all of the following were equivalent

- controllable from zero to arbitrary end point= controllable from arbitrary point to zero
- local controllability = global controllability
- can reach open set = can reach all of R^n
- x reachable in some time $t \Rightarrow x$ reachable in arbitrary time t

These are not equivalent in the nonlinear case. Another problem is that different nonlinear authors use different definitions.

In the linear case we also had a duality between controllability and observability. Keep your eyes open for any corresponding duality in the nonlinear case.

Most of the results in Ch. 3 generalize in some way to the $\dot{x} = f(x, u)$ case. This is mentioned in the end of the chapter. The text, however, concentrates on the affine case.

Assumption 3.1a allows for restrictions such as $u_i \geq 0$. Dont worry too much about Assumption 3.1b, we can often approximate u with such functions. But note e.g. that $u = x$ formally is not allowed since u is not piecewise constant then.

In Def. 3.2 the time T can be chosen freely. Prop. 3.3 says that controllability of the linearized system implies local nonlinear controllability. Note also that $T > 0$ can be chosen arbitrarily small then. The linearized system is unfortunately often not controllable, so the proposition is of limited value. Do exercises 3.1 and 3.2 now.

Note that Prop 3.6 only says that we can control approximately in the direction of $[g_1, g_2]$, not exactly in that direction. Remember that even for linear systems it is almost never possible to make $x(t)$ an arbitrary function of time just because the system is controllable.

One usually describes C by drawing a Lie-bracket tree; bracketing with $[f, \ldots$ in the left branch and with $[g_i, \ldots,$ in the right branches. You might wonder if this covers all possibilities. The answer is yes, since symmetry and the Jacobi identity give that all other bracket-combinations are linear combinations of these. This is Prop. 3.8.

Note that x_0 does not necessarily belong to $R_T^v(x_0)$. Accessibility is hence a weak concept. Th. 3.9 is a (partial) success for the concepts developed so far. It is enough that you get the idea of the proof.

Prop. 3.12 describes what happens if the accessibility condition 3.10 is not satisfied. S_{x_0} is a nonlinear version of a controllable subspace. Cor. 3.13 is a converse of Th. 3.9. (A set S is dense in M if every point $x \in M$ can be approximated arbitrarily well by points $y \in S$. For instance the rational numbers are dense in the set of real numbers.)

Prop 3.15 gives a better controllability concept. If there is no drift term $f = 0$, the accessibility algebra if often called the controllability algebra instead.

For systems with drift term the situation is more complicated. The search for stronger rank conditions implying controllability, instead of just accessibility, is still an active research area. Most famous is probably the *Sussman sufficient condition* of so called "odd" systems: If dim $= n$ and all Lie-brackets with an even (≥ 2) number of g:s are linear combinations of brackets with an odd, smaller, number of g:s then the system is locally controllable. For instance, Ex. 3.14 is not an odd system since $[g, [f, g]] \notin \text{span}(g, [f, g], [f, [f, g]], [f, [f, [f, g]]], \ldots)$.

Page 84 is important, make sure you understand how the linear case follows from the nonlinear.

The difference between 3.18 and Def. 3.10 is that $R^v(x_0, T)$ is used instead of $R_T^v(x_0)$ i.e. you reach the states at a fixed time T.

Skip the proof of Th. 3.21, except (3.42). Prop. 3.22 and especially (3.45) are important. Example 3.24 is fun now. Also read the bilinear system stuff on page 91 carefully.

Exercise 3.1 When is the linearization of $\dot{x} = \sum g_i(x)u_i$ controllable at $u_i = 0$?

Exercise 3.2 Is Prop. 3.3 true for the more general case $\dot{x} = f(x,u)$?

Exercise 3.3 = Nij. 3.2.

Exercise 3.4 = Nij. 3.4ab.

Exercise 3.5 = Nij. 3.5.

Exercise 3.6 Write $[[f,g],[f,[f,g]]]$ as a linear combination of two brackets in the Lie-bracket tree.

Exercise 3.7 Compute the accessibility distribution and its span at $x = 0$ for the system

$$\begin{aligned}\dot{x}_1 &= u_1\\ \dot{x}_2 &= u_2\\ \dot{x}_3 &= x_1u_2 - x_2u_1\end{aligned}$$

Is the system controllable? [Find a control that takes the system from $(0,0,0)$ to $(0,0,1)$. Hint: Try using u_i = sum of some sinusoids.]*.

Exercise 3.8 Consider the system on the cylinder $R \times S^1$ described by

$$\begin{aligned}\dot{x} &= u\\ \dot{\theta} &= 1\end{aligned}$$

where $\theta \in S^1$ is only consider modulo 2π. Is the system locally accessible, locally strongly accessible, controllable?

Exercise 3.9 Show that the rolling penny is locally strongly accessible:

$$\frac{d}{dt}\begin{pmatrix}\phi\\ \psi\\ x\\ y\end{pmatrix} = \begin{pmatrix}1\\ 0\\ 0\\ 0\end{pmatrix}u_1 + \begin{pmatrix}0\\ 1\\ \cos\phi\\ \sin\phi\end{pmatrix}u_2$$

Is the system controllable?

Exercise 3.10 Check controllability for

$$\begin{aligned}\dot{x}_1 &= u_1x_3 + u_2\\ \dot{x}_2 &= u_1x_1\\ \dot{x}_3 &= u_1x_2\end{aligned}$$

Exercise 3.11 * A pendulum with variable length d is described by

$$\ddot{\theta} + 2\dot{d}\dot{\theta}/d + (1/d)\sin\theta = 0$$

What can be said about controllability and accessibility if $\dot{d}$ is the input?

Exercise 3.12 Check that the following system is controllable by showing that it is an odd system:

$$\begin{aligned}\dot{x}_1 &= x_2^3 \\ \dot{x}_2 &= u\end{aligned}$$

Exercise 3.13 The following bilinear model for a fermentor is taken from [Axelsson and Hagander, CDC90].

$$\begin{aligned}\frac{dS}{dt} &= -S + (1 - S)u \\ \frac{dE}{dt} &= S - Eu\end{aligned}$$

Here S, E, u are sugar, ethanol and flow rate through the reactor respectively. Check local accessibility/controllability around an arbitrary point (S_0, E_0).

Exercise 3.14 Construct an example that illustrates Prop. 3.22., including (3.44) and (3.45).

Session 4

One-forms. Codistributions. Local Observability. Nonlinear Kalman Decomposition

Reading assignment

pp. 61-66, 93-116.

Section 2.2.3

To discuss observability we need the dual concepts to tangent vectors, vector fields, and distributions. These are cotangent vectors, differential one-forms and codistributions. This is discussed in Section 2.2.3 that we skipped on the first reading. A cotangent vector has the form

$$\sigma = \sum_{i=1}^{n} \sigma_i dx_i$$

It acts on a tangent vector $X = \sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial x_i}$ by

$$\sigma(X) = \sum_{i=1}^{n} \sigma_i \alpha_i$$

The change of coordinate formula (2.149) can also be written

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = \frac{\partial x}{\partial z} \begin{pmatrix} dz_1 \\ \vdots \\ dz_n \end{pmatrix}$$

which perhaps is easier to remember. Do exercises 4.1-3 now.

$F^*\sigma$ is often called the "pull-back" of σ. In coordinates it is just a multiplication with the Jacobian $\frac{\partial F}{\partial x}$ and change of evaluation point. (2.150) and (2.152) are important, make sure you understand them. It is a good check that you have understood the material to verify (2.169). The discussion on page 65-66 describes the basis for the duality between controllability and observability. Make sure you understand the statement of Frobenius' theorem with one-forms.

Sections 3.2-3.3

We now introduce measurements. Definition 3.27 is a natural generalization of linear non-observability (now everything depends on x_0 and u however). Local observability means that all states close to x_0 can be separated using control signals that keep the states close to x_0. To form the rank condition we calculate repeated Lie derivatives (3.70) and then form $d\mathcal{O}$ by (3.74). Read props. 3.34, 3.38 but skip the proofs.

Section 3.3 presents the nonlinear Kalman decomposition. In Prop. 3.42 k is the dimension of D. Unfortunately it is often nontrivial to find the transformations. Prop. 3.46 relates some dual concepts. Remember that (A, B) is controllable iff (A^T, B^T) is observable. Note now, however, that local controllability is a condition that depends only on f, g whereas local observability depends on f, g, h. So the duality can not be drawn as far as in the linear case. To enjoy 3.47 you must recap the geometric theory from linear systems, you can skip 3.47 if you dont have time to do this. Theorems 3.49 and 3.51 are the goals of Sec. 3.3. Skip the proof of Prop. 3.50.

Exercise 4.1 Calculate df when $f = x_1^2 + e^{x_2}$ and "check" that $df(X) = X(f)$ on $X = \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}$.

Exercise 4.2 Compute $\sigma(X)$ for $X = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}$, $\sigma = dr + r d\phi$ where r, ϕ are polar coordinates.

Exercise 4.3 Transform $x_1 dx_1 - x_2 dx_2$ to polar coordinates (r, ϕ).

Exercise 4.4 Calculate $L_X \sigma$ when $\sigma = x_1 dx_1 + x_2 dx_2$ and $X = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2}$. Note then that $\sigma = df$ for $f = r^2/2$ and check that $L_X df = d(L_X f)$.

Exercise 4.5 Show that $(F \circ G)^* = G^* \circ F^*$.

Exercise 4.6 = Nij. 2.13. Verify also(2.169).

Exercise 4.7 = Nij. 3.4c.

Exercise 4.8 * = Nij. 3.5c.

Exercise 4.9 * = Nij 3.6.

Exercise 4.10 = Nij. 3.7.

Exercise 4.11 * = Nij. 3.11.

Exercise 4.12 Consider

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x + u \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} x \qquad u \in R$$

$$y = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} x.$$

Compute $\mathcal{O}$ for the two cases $u \equiv 0$ and $u(t)$ arbitrary? Interpretation?

Exercise 4.13 Construct an example that illustrates Theorem 3.51.

Session 5

State Space Transformations. Feedback. Normal Form. Exact Linearization.

Reading assignment

pp. 148-175 + part of Sastry's manuscript.

Chapter 5

The goal of chapters 5 and 6 is to present some different versions of "feedback linearization". The most famous variant is to change inputs from u to v by $u = \alpha(x) + \beta(x)v$ and coordinates from x to z by $z = S(x)$ in such a way that the new system is a linear system from v to z. The idea is then to use linear design methods on this new system and afterwards transform back to the original coordinates. Unfortunately, not all systems are feedback linearizable. Chapters 5 and 6 shows exactly what the conditions are.

Peoples opinion about feedback linearization as a design method differ. Even if the system is feedback linearizable, the transformations needed can be quite involved and hide any physical insight. See also the discussion on p. 177. Anyway, it is widely discussed and one should know the results. It is also a nice application of what we have learned so far.

Ch. 5 describes some preliminary results on linearization by coordinate transformations $z = S(x)$ alone (no feedback). Theorem 5.3 gives a clean answer to when a system is "coordinate linearizable". The story is the following: Linear systems have a very simple Lie-bracket tree, everything except the linear subtree, $g, [f, g], \ldots, \mathrm{ad}_f^j g, \ldots$ is zero. *This structure is not changed by coordinate transformations.* f and g are changed to S_*f and S_*g and Lie-brackets are changed in the same way: $S_*[f, g] = [S_*f, S_*g]$. The result of Theorem 5.3 (see also (5.39)) is therefore very natural. Corollary 5.6 is a stronger version of Theorem 5.3, but is only needed in the MIMO case, you can skip it.

We will not need all of Ch. 5, in fact you can stop reading after (5.40) and skip Problem 5.8. Start reading at page 158 after (5.49) again. Th. 5.13 gives the dual result of Th. 5.3. If you find the MIMO-notation cumbersome, rewrite the equations to the SISO-case. The nonlinear transformation to observable form (5.59) is of course important. If we allow the *:s to depend on z the transformation to (5.59) is always possible if (5.46) is satisfied, which is equivalent to that the linearized system is observable. You can skip pages 161-165.

Section 5.2 is just a discussion and some definitions introducing feedback. Ex. 5.20 is the famous "computed torque" technique in robotics. Note that everything in the controller (5.97) depends on x. The only restriction is that v enters affinely. Dynamic output feedback is more realistic in practice.

Sastry's Notes

Sastry's presentation is now a good complement. Th. 8.7 in Sastry gives an implicit answer to the exact linearization problem defined by f and g. A system is feedback linearizable if there is an output function such that the new system has relative degree n. This might be helpful for the intuition, but does not answer the problem completely since one often does not know how to find such an output function. The anwer is given later: check involutivity of D_{n-1}. The normal form mentioned in Sastry is sometimes called *"the controller form"* when $\gamma = n$ (so η disappears).

Exercise 5.1 Check Theorem 5.3 on Example 5.1.

Exercise 5.2 What does (5.52) mean for a linear system?

Exercise 5.3 Is the following system linearizable by $z = S(x)$ (Maple)?

$$\begin{aligned}
\dot{x}_1 &= \frac{1}{1+x_1}u \\
\dot{x}_2 &= \frac{x_2^2}{2} + x_3 - \frac{x_1}{1+x_1}u \\
\dot{x}_3 &= x_1 + \frac{x_1^2}{2} - \frac{x_2^3}{2} - x_2x_3 + \frac{x_1x_2}{1+x_1}u
\end{aligned}$$

Exercise 5.4 * What is a controllability counterpart to the observability canonical form (5.59)?

Exercise 5.5 = Nij. 5.5a

Exercise 5.6 = Nij. 5.6

Exercise 5.7 = Nij. 5.7

Exercise 5.8 * = Nij. 5.8

Exercise 5.9 * = Nij. 5.9

Exercise 5.10 * = Nij. 5.10 (wrong).

Exercise 5.11 Transform

$$\begin{aligned}
\dot{x}_1 &= \sin x_2 \\
\dot{x}_2 &= \sin x_3 \\
\dot{x}_3 &= u
\end{aligned}$$

into normal form. In what region of the state space is the transformation well defined? Compute a state feedback $u = \alpha(x) + \beta(x)v$ giving linear dynamics with poles in $-1, -2, -3$. Also compute a feedback $u = -Lx$ by first linearizing the system (using $\sin x \approx x$ etc) and then placing the poles in $-1, -2, -3$. Compare the gains of the two feedback laws for $|x_i| < 0.5, i = 1, 2, 3$. Do they differ much?

Exercise 5.12 Transform the following system into normal form

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

Session 6

Exact Linearization (MIMO). Disturbance Decoupling. Input Output Decoupling. Dynamic Extension.

Reading assignment

(Nij. pp. 176-298.)

This lecture will be a very condensed summary of the results in chapters 6-9 in Nijmeijer. You can prepare by looking up the problem statements in each chapter. It will also help if you have a look on the geometric theory in the linear systems course, especially the algorithm for calculating maximal A, B-invariant subspace in ker C.

We will not go into any details in this lecture. These chapters should be read only by those interested in continuing with nonlinear research. It could then also be a good idea to take a look on the presentation in Isidori in parallel.

Exercises

There will be no exercises on this material.

Session 7

Zero Dynamics. Nonlinear Minimum Phase. Center Manifold Theorem

Reading assignment

pp. 310-317 + 323-338 (stop after Ex 11.18).

Chapter 10.3-11

This lecture will give more information on *zero dynamics* with applications. The main point will be a proof of stability using center manifold theory.

Exercises on Ch. 6-11

Exercise 7.1 = Nij. 6.3

Exercise 7.2 = Nij. 6.4

Exercise 7.3 = Nij. 6.5

Exercise 7.4 = Nij. 6.8

Exercise 7.5 = Nij. 6.9

Exercise 7.6 * = Nij. 6.10

Exercise 7.7 Vechicle Dynamics (Sastry). The dynamics (extremely simplified) of a wheeled vehicle on a flat road from enginge force input F to the position fo the vehicle center of mass x are described by

$$m\ddot{x} = F - \rho(\dot{x}^2)\mathrm{sgn}(\dot{x}) - d_m$$

In this equation ρ stands for the coefficient of wind drag, d_m the mechanical drag and m the mass of the vehicle. Further the engine dynamics are modeled by the first order system

$$\dot{\xi} = -\frac{\xi}{\tau(\dot{x})} + \frac{u}{m\tau(\dot{x})}$$

Here ξ is a state variable modeling all of the complex engine dynamics, u is the throttle input and $\tau(\dot{x})$ is a velocity dependent time constant. Also $F = m\xi$. Write these equations in state space form and examine the feedback linearizability of the engine from input u to the output x. What are the relative degree, zero dynamics of the system?

Exercise 7.8 Consider the SISO system

$$\begin{aligned}\dot{x} &= f(x) + g(x)u + p(x)w\\ y &= h(x),\end{aligned}$$

where w is a disturbance. Show that there is a feedback law $u = \alpha(x) + \beta(x)v$ that makes y independent of w if

$$L_p L_f^k h(x) \equiv 0 \quad \text{for } 0 \leq k \leq r-1.$$

where r is the relative degree of the system. (The condition is in fact also necessary.) Note that there is no guarantee for stability, this has to be achieved separately.

Exercise 7.9 Consider the same system as above, but assume w is measurable. Show that there exists a feedback/feedforward law $u = \alpha(x) + \beta(x)v + \delta(x)w$ so that y is independent of w if

$$L_p L_f^k h(x) \equiv 0 \quad \text{for } 0 \leq k \leq r-2$$

What does this mean in words?

Exercise 7.10 Prove that if $z = S(x)$; $v = \phi(x,u) \in R$; $\phi_u \neq 0$ transforms $\dot{x} = f(x) + g(x)u$ to $\dot{z} = Az + bv$ then $\phi(x,u) = \alpha(x) + \beta(x)u$, for some $\alpha(x), \beta(x)$. Comments?

Exercise 7.11 Is the following system feedback linearizable with $z = S(x), u = \alpha(x) + \beta(x)v$?

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

Exercise 7.12 * What is the connection between controlled invariant distributions and controlled invariant submanifolds?

Exercise 7.13 Calculate some center manifolds.

Session 8

Nonlinear Mechanical Control Systems

Reading Assignment

Take a quick look on Nijmeijer Ch. 12. RolfJ will hand out his own material.

Session 9

Volterra Systems

Reading assignment

T. Glad Manuscript.

Volterra Systems

Guest Lecture with Sven Spanne.

Session 10

Lyapunov functions, regions of attraction, center manifolds

Reading assignment

Khalil pp. 97-179 + 186-225. The first two chapters are excluded, since most of their content is known from other courses.

Chapter 3

This chapter is devoted to the study of equilibrium points of nonlinear autonomous systems. The main issues are the following.

- The use of Lyapunov functions and invariant sets for proofs of asymptotic stability (LaSalle's theorem). Consider in particular its application to the pendulum, Example 3.4.
- The use of Lyapunov functions for proofs of instability (Chetaev's theorem).
- Stability analysis by linearization.
- Two ways to estimate a region of attraction, by Lyapunov functions (LaSalle's theorem) and by computer iteration. Consider in particular Examples 3.22 and 3.23.

Chapter 4

Here the main new ingredient is time-variations. The terms *uniform asymptotic stability* and *exponential stability* are introduced to specify time dependence in the stability behaviour. Main results:

- Uniform asymptotic stability can be proved from time-invariant bounds on the Lyapunov function. This is Theorem 4.1.
- For linear systems, uniform asymptotic stability is equivalent to exponential stability (Theorem 4.2).
- Exponential stability implies input-output stability (Theorem 4.13).
- The center manifold theorem is a powerful complement to stability analysis by linearization. Section 4.7 should therefore be read in connection to Section 3.3.

Exercises on Chapter 3-4

Exercise 10.1 = Kha. 3.7

Exercise 10.2 What is the region of attraction for the origin in the previous example?

Exercise 10.3 = Kha. 3.26

Exercise 10.4 = Kha. 4.34

Exercise 10.5 = Kha. 3.29

Exercise 10.6 = Kha. 4.5

Exercise 10.7 = Kha 4.7

Exercise 10.8 = Kha. 4.11

Exercise 10.9 = Kha. 4.17

Exercise 10.10 = Kha. 4.20

Hand in problem number 3

Select a reasearch article in geometric nonlinear control theory and make an 8 min presentation of the content on Wednesday, May 4.

Session 11

Absolute stability, Kalman-Yakubovich Lemma, The Circle and Popov criteria

Reading assignment

Khalil pp. 237 - 268. Extra material on the K-Y-P Lemma and the Matlab toolbox LMI-lab is provided.

Comments on the text

This section of the book will get the most detailed coverage of all. The results have played a central role in control theory for a long time and have recently been vitalized by new progress, both in theory and in computational methods.

The concept ***absolute stability*** is introduced for nonlinear systems consisting of two parts, one linear time-invariant and one nonlinear. Detailed knowledge about the nonlinear part is not used, only inequality constraints.

The Kalman-Yakubovich-Popov Lemma shows that a transfer function inequality is equivalent to a condition on solvability of a linear matrix inequality (LMI) defined by the state space matrices. In the proof of the circle and Popov criteria, the LMI appears naturally in the attempt to construct a Lyapunov function. The K-Y-P Lemma therefore connects the existence of a certain Lyapunov function to a transfer function condition on the linear part. Khalil does not provide a complete proof, instead we refer to separate notes which are distributed this week.

Recently, the same lemma has often been used in the opposite direction, as frequency conditions on multivariable transfer functions are verified by translating them into an LMI condition, which can be solved by convex optimization. Some of the exercises below will illustrate this and the MATLAB Toolbox LMI-lab will be useful for the calculations.

Soon after the appearance of the Popov criterion, for example in the textbook by Aiserman and Gantmacher from 1965, it was pointed out that the Popov criterion holds also with negative slope $1/\eta$ on the Popov line. However, this fact is ignored by Khalil and several other western textbooks. Can you see why it must be true?

Exercise 11.1 = Kha. 5.5

Exercise 11.2 = Kha. 5.6

Exercise 11.3 = Kha. 5.4

Exercise 11.4 Solve the previous exercise with the circle criterion replaced by the Popov criterion.

Exercise 11.5 **a.** Find a quadratic simultaneous Lyapunov function (for example using LMI-lab) for the linear time-varying system

$$\dot{x} = \begin{bmatrix} 1 & -8 \\ 6 & -13 \end{bmatrix} x + \begin{bmatrix} 8 & 1 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} \delta_1(t) & 0 \\ 0 & \delta_2(t) \end{bmatrix} y$$
$$y = \begin{bmatrix} 1 & -1 \\ -11 & 16 \end{bmatrix} x$$

where $|\delta_k(t)| \leq 1$, $k = 1, 2$.

b. Does Theorem 5.1 prove stability if the δ-matrix is replaced by a memoryless nonlinearity satisfying the sector condition

$$[\psi(t, y) + y]^T [\psi(t, y) - y] \leq 0, \quad \forall t \geq 0, y \in R^2$$

c.* If it is also assumed that ψ is decentralized (Khalil, page 239), then the matrix $G(s)$ in Theorem 5.1 can be replaced by $DG(s)D^{-1}$ for any diagonal invertible matrix D. How can the optimization over D be included in an LMI formulation? Does it give stability in this case?

Exercise 11.6 Prove Lemma 5.1 in Khalil using Theorem 1 in "A Note on the Kalman-Yakubovich-Popov Lemma". What role is played by the controllability and observability assumptions? How about the positivity of P?

Exercise 11.7 Consider a $p \times p$ matrix function $M(s)$, which is analytic for Re $s > 0$ and satisfies $M(s) = \overline{M(\bar{s})}$. The matrix function is called

output strictly passive (OSP) if $\exists \epsilon > 0$ such that for Re $s > 0$

$$M(s) + M(s)^* \geq \epsilon M(s)^* M(s)$$

input strictly passive (ISP) if $\exists \epsilon > 0$ such that for Re $s > 0$

$$M(s) + M(s)^* \geq \epsilon$$

positive real (PR) if for Re $s > 0$

$$M(s) + M(s)^* \geq 0$$

strictly positive real (SPR) if $\exists \epsilon > 0$ such that $M(s - \epsilon)$ is PR.

a. For the scalar transfer function $M(s) = C(sI - A)^{-1}B + D$ with A Hurwitz, show that

$$\text{ISP} \Rightarrow \text{SPR} \Rightarrow \text{OSP} \Rightarrow \text{PR}$$

b. Prove a counterpart to Lemma 5.1 with SPR replaced by OSP and (5.10-12) replaced by a convex LMI.

c.* Modify the proof of Theorem 5.1 to replace SPR by OSP.

Moral: OSP is a nicer concept than SPR.

Session 12

Dissipativity, Multipliers, IQC's, Slow variations, Interconnections

Reading assignment

Khalil pp. 268 - 286. Extra material: An article on dissipativity and a preprint on integral quadratic constraints.

Comments on the text

In this lecture, the results of absolute stability will be further developed towards a general theory for analysis of systems with nonlinearities, time-variations and uncertain elements. This is still research area, so old and well established results will mixed with recent developments.

The memoryless nonlinearities from the previous lecture are replaced by more general nonlinear operators. The concepts dissipativity and storage functions, introduced by Willems in 1972, describe time-domain properties of such operators. In the frequency domain, we will work with so called Integral Quadratic Constraints (IQC's). As it turns out, a large variety of results in the litterature can be interpreted and unified based on these few concepts.

Two more sections on Lyapunov functions in Khalil are also covered. One demonstrates how Lyapunov functions for several subsystems can be combined in the analysis of their interconnections. The other shows how the effects of slow time variations can be analysed with Lyapunov functions. Such analysis is for example appropriate in the context of gain-scheduling, where controllers are designed for constant model parameters, but are expected to work also for a certain amount of time-variations.

Exercise 12.1 State and solve your own exercise on this session.

Exercise 12.2 = Kha. 5.18

Exercise 12.3 Is the center manifold theorem applicable to Khalil 5.18? Compare the results.

Exercise 12.4 Kha. 5.22-24

Exercise 12.5 = Kha. 5.9

Exercise 12.6 Derive an LMI condition for asymptotic stability in Khalil 5.9. Apply it to 5.9(c) using LMI-lab. How do you take into account the diagonal structure of the nonlinearity?

Exercise 12.7 a. Let $\int_0^\infty |h(t)|dt \leq 1$. For bounded $v(t)$ and $w(t) =$ sign $v(t)$, verify the inequality

$$v(t)\left[w(t) - \int_0^t h(t-s)w(s)ds\right] \geq 0.$$

Use this to derive an IQC for the sign -operator.

b*.It is believed by the author of this excercise that also saturations satisfy the same IQC. Prove or disprove that statement.

Exercise 12.8 Suppose that the operators Δ_1 and Δ_2 satisfy the IQC's defined by $\Pi_1(j\omega)$ and Π_2 respectively. Can you define IQC's that are satisfied by

a. $\begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$ **b.** $\begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix}$ **c.** $\begin{bmatrix} \Delta_1 & \Delta_2 \end{bmatrix}$

Session 13

Oscillations, Describing functions, Stability analysis, Perturbation theory

Reading assignment

Khalil pp. 339 - 407.

Comments on the text

Three methods to detect periodic orbits are treated in this course, the Poincare'-Bendixon theorem (for second order systems only), describing functions and averaging. This lecture will cover the first two and also introduce some perturbation theoretic results for later use.

Second order systems Periodic orbits in the plane are special in that they divide the plane in two parts, the region inside the orbit and the region outside. No trajectories can pass from one region to the other. This leads to the fact that the only limit sets that can exist in the plane are

- periodic orbits
- equilibrium points
- trajectories connecting equilibrium points

Describing functions The describing function method applies to nonlinear systems that can be represented as a feedback connection of a nonlinear component with a linear system of low-pass character. The basic idea is that the low-pass filter removes high frequency signal components, so the input to the nonlinearity is essentially sinusoidal. The *describing function* is therefore introduced as the map from amplitude of a sinusoidal input to the nonlinearity to the amplitude and phase shift of the corresponding frequency component in the output.

The results demonstrate how bounds on the low-pass character of the linear component can be turned into existence and non-existence proofs for periodic solutions.

Exercise 13.1 Kha. 6.7

Exercise 13.2 Kha. 6.10

Exercise 13.3 Kha. 6.13 (1),(3),(5)

Exercise 13.4 * Outline how the nonlinearity in Khalil 6.13 (5) can be viewed as a combination of simpler nonlinear components and suggest how to analyze it with a multivariable describing function.

Exercise 13.5 Kha. 6.14

Exercise 13.6 Kha. 6.19

Exercise 13.7 Kha. 7.3

Session 14

Averaging, Singular perturbations

Reading assignment

Khalil pp. 408 - 476.

Comments on the text

This last session is devoted to system dynamics that operate on two time-scales. In the analysis of the slow dynamics, the influence of the fast variables is essentially governed by their average values. On the other hand, the slow variables may be considered constant in the analysis of the fast dynamics. The theory is developed in two steps as described below.

Averaging First, we consider systems where the fast dynamics appear as explicit time-variations in the system equations. The equations can be written on the form

$$\frac{dx}{d(\epsilon t)} = f(t, x, \epsilon)$$

in order to emphasize the two different time scales ϵt and t. The main idea of Chapter 7 is to approximate x with the solution x_{av} of the equation

$$\begin{aligned} \dot{x}_{\mathrm{av}} &= \epsilon f_{\mathrm{av}}(x_{\mathrm{av}}) \\ f_{\mathrm{av}}(x) &= \lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} f(\tau, x, 0)d\tau \end{aligned}$$

Singular Perturbations In models from first principles there are often neglected high frequency dynamics, that are caused by "parasitic" masses, resistors, inductances, etc. System perturbations of this kind, that change the order of the model are called *singular perturbations.*

Here the fast dynamics are determined by a differential equation that includes the slow variables, so the slow and fast dynamics are intertwined. The general system description takes the form

$$\begin{aligned} \dot{x} &= f(t, x, z, \epsilon) \\ \epsilon\dot{z} &= g(t, x, z, \epsilon) \end{aligned}$$

A potentially approximative solution is generated by the equations

$$\begin{aligned} \dot{\bar{x}} &= f(t, \bar{x}, h(t, \bar{x}), 0) \\ 0 &= g(t, \bar{x}, h(t, \bar{x}), 0) \end{aligned}$$

The second equation is solved for h, and the first equation is used for $\bar{x}$.

Exercise 14.1 Kha. 7.8 (1),(2)

Exercise 14.2 Kha. 7.10

Exercise 14.3 Kha. 7.13 (1),(2)

Exercise 14.4 Kha. 7.14

Exercise 14.5 Kha. 7.17

Exercise 14.6 Kha. 8.3

Exercise 14.7 Kha. 8.4

Exercise 14.8 Kha. 8.7

Exercise 14.9 Kha. 8.11

Nonlinear Geometric Control

Why $\dot{x} = f(x) + g(x)u$?

Good models sometimes. Exciting theory has been built last 10-20 years. Can control in all $\dot{x} \in$ linear subspace.

Why so much math?

Could have done away with less. Not so much math, mostly notation. Must learn vocabulary to be able to read modern research articles.

Why "manifolds"?

Angles. Needed for analysis of global properties. Extensively used.

Why no noise?

Theory is hard and scattered. Must understand this to have practical theory though. Work needed.

Why is Nijmeijer's book the best alternative?

Personal. Must *work* with the material. Has hundreds of exercises. Isidori has three (3) exercises. Better than Isidori in some proofs. A little more recent. Several drawbacks also.

What is missing?

MUCH !

- Almost no synthesis
- Basic ODE etc elsewhere.
- Differential Algebra Approach
- Optimal Control
- Identification
- Global results
- Disturbance Models/stochastics
- Algorithms, Numerics

No good material or lower priority.

Summary of Geometric Approach

Mathematical Handwork

$$(M, \varphi_i(x))$$

$$X = \sum_{i=1}^{n} a_i(x) \frac{\partial}{\partial x_i}$$

$$L_X(f) = \sum a_i(x) \frac{\partial f}{\partial x_i}$$

$$[f, g] = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g$$

$$\sigma = \sum a_i(x) dx_i$$

Etc

$$\text{Steer} = \frac{\partial}{\partial \theta}$$

$$\text{Drive} = \cos(\phi+\theta)\frac{\partial}{\partial x} + \sin(\phi+\theta)\frac{\partial}{\partial y} + \sin\theta\frac{\partial}{\partial \phi}$$

$$[\text{Steer}, \text{Drive}] = \text{Wriggle}$$

$$[\text{Steer}, \text{Wriggle}] = \text{Slide}$$

Local Controllability:

- *A nonlinear system is controllable if the linearized system is controllable.*
- $\dot{x} = f(x) + g(x)u$ *is "accessible" iff*

$$\dim\ (f, g, [f,g], [f,[f,g]], \ldots) = n$$

- Controllable submanifold.

If $f = 0$ then accessibility = controllability.

Local Observability. Depends on x_0 and u.

$y_j = h_j(x)$

$\mathcal{O} = \text{span}\, L_{X_1} \ldots L_{X_k} h_j(x)$

$d\mathcal{O} = \text{span}\ (dH \mid H \in \mathcal{O})$

The system is locally observable if

$$\dim\ (d\mathcal{O}) = n$$

Duality between observability and controllability

Nonlinear Kalman Decomposition

Can find coordinates (x_1, x_2, x_3, x_4) so that

$$\begin{aligned}
\dot{x}_1 &= f^1(x_1, x_3) + g(x_1, x_3)u \\
\dot{x}_2 &= f^2(x_1, x_2, x_3, x_4) + g(x_1, x_2, x_3, x_4)u \\
\dot{x}_3 &= f^3(x_3) \\
\dot{x}_4 &= f^4(x_1, x_3) \\
y &= h(x_1, x_3)
\end{aligned}$$

Relative Degree

Smallest r such that $L_g L_f^{r-1} h(x_0) \neq 0$

Exact Linearization by Feedback

$$\begin{aligned}
&\dot{x} = f(x) + g(x)u \\
&u = \alpha(x) + \beta(x)v \text{ and } z = Z(x) \Longrightarrow \dot{z} = Az + Bv
\end{aligned}$$

The system is feedback linearizable around x_0 *if one can find* $y = \lambda(x)$ *so the system has relative degree* n*. Can be checked with some Lie-brackets.*

Zero Dynamics

Nonlinear Minimum Phase

Disturbance Decoupling

Normal Forms

Stabilization

ETC, ETC

Examples

Exercises !!

Nonlinear Control Theory 94

Lecture 1

- Examples
- Manifolds
- Inverse and Implicit Function Theorems
- Tangent Vectors

Must Specialize

$$\begin{aligned} \dot{x} &= f(x,u) \\ y &= h(x,u) \end{aligned}$$

No noise

No algebraic equations $f(x,\dot{x},u) = 0$.

Important special affine case:

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned}$$

f : drift term

g : input term

Example Pendulum

$$\ddot{\theta} = \sin(\theta) + u$$

Natural state space: $R \times S^1$ = cylinder

$S^1 =$ unit circle

Example Robotics

$$M(\theta)\ddot{\theta} + C(\theta,\dot{\theta}) + K(\theta) = u$$

Use structure and physical insight.

$$\begin{aligned} x_1 &= \theta \\ x_2 &= \dot{\theta} \end{aligned}$$

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ M^{-1}(x_1)(-C(x_1,x_2) - K(x_1)) \end{pmatrix} + \begin{pmatrix} 0 \\ M^{-1}(x_1) \end{pmatrix} u$$

Structure is lost, eg sparseness in M, C, K matrices

Rigid Bodies

Natural State Space

$$R = \begin{pmatrix} r_1 & r_2 & r_3 \end{pmatrix} \in SO(3)$$

$$\text{ie } RR^T = I \text{ and } \det(R) = 1$$

$$\dot{r} = -r \times w \Leftrightarrow \dot{r} = -rS(w)$$

$$S(w) = \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}$$

Gas Jet Actuators

$$\dot{R} = -RS(w)$$
$$J\dot{w} = S(w)Jw + Bu$$

J inertia matrix

Momentum Wheel Actuators

$$\dot{R} = -RS(w)$$
$$J\dot{w} = -RS(w)h + Bu$$
$$\dot{h} = 0$$

h total momentum

Linearized dynamics **not** controllable $\dot{R} = 0$.

Will see that it is nonlinear controllable

Bilinear Models

$$\dot{x} = Ax + \sum_{i=1}^{k} u_i D_i x + Bu$$

Typically x_i concentrations

Car With N trailer

Hard to back. (Must use many Lie-brackets).

Keep your eyes on the last trailer

Rolling Penny

$$\dot{\varphi} = u_1$$
$$\dot{\theta} = u_2$$
$$\dot{x} = u_1 \cos(\theta)$$
$$\dot{y} = u_1 \sin(\theta)$$

Can it be moved sideways in small time (keeping the head up)?

Holonomic constraints $h(x) = 0 \Longrightarrow h_x \dot{x} = 0$.

Non-holonomic constraints $a(x)\dot{x} = 0$

Manifolds

What are natural mathematical models for state spaces?

Piece together "bent" pieces of R^n.

Same local properties as R^n.

Different globally

Gauss, Riemann, Poincare, Weyl, Whitney

Definition A C^∞ (=smooth) manifold is a topological space M together with an atlas $\{U_\alpha, \varphi_\alpha\}$ of pairwise C^∞-compatible coordinate charts that covers M.

Topological space ?

Atlas $\{U_\alpha, \varphi_\alpha\}$?

Pairwise C^∞-compatible coordinate charts ?

Topology

A topology on a set M is a collection T of subsets of M.

O is called "open" if $O \in T$.

The collection T must be such that

- $\emptyset, M \in T$
- $O_1, O_2 \in T \Longrightarrow O_1 \cap O_2 \in T$
- $\{O_i\} \in T \Longrightarrow \cup O_i \in T$

A set C is called *closed* if $M - C$ is open.

Example $M = R^n$, T = open subsets of R^n, i.e. $x \in O \Longrightarrow B_\epsilon(x) \subset O$ some $\epsilon > 0$.

$$x^2 + y^2 - 1 < 0$$

More Topology

$f : M_1 \to M_2$ is called *continuous* if

$$f^{-1}(\text{open}) = \text{open}$$

Metric: $d(x, y)$

Norm: $||x - y||$

Sufficient for continuity:

$$||f(x) - f(y)|| \leq C||x - y||$$

Example f = differential operator d/dx

$M_1 = C^1$ with norm $\sup |f'(x)|$

$M_2 = C^0$ with norm $\sup |f(x)|$

Compatible Coordinate Charts

$$\psi \circ \varphi^{-1}(x) \in C^\infty$$

f smooth if $f(\varphi^{-1}(x)) \in C^\infty$

$$f \circ \varphi^{-1}(x) = f \circ \psi^{-1} \circ \psi \circ \varphi^{-1} \in C^\infty$$

Independent on coordinate charts

Coordinate free descriptions

Example: Cylinder

$\psi \circ \varphi^{-1}$ smooth on $U \cap V = (x_2 \neq 0, z_2 \neq 0)$

$z = \psi(\varphi^{-1}(x))$ is given by

$$z_1 = x_1$$
$$z_2 = 4/x_2$$

C^∞-manifold

Examples

Example $(x, m\dot{x}) \in R^{2n}$

Example

Example

Example Sphere S^2

Global Differences

Smooth velocity field v on S^2 must have (at least) 1 equilibrium, $v(x) = 0$ (degree theory).

"You can't comb the hair of a tennis ball"

Many manifolds are given implicitly by

$$f_1(x_1,\ldots,x_n)=0$$
$$\vdots$$
$$f_k(x_1,\ldots,x_n)=0$$

Does this describe an $n-k$-dimensional manifold?

Differentials

$f: A \to B$ is called differentiable at $x \in A$ iff there is a continuous linear map $DF_x(h)$ such that

$$\|f(x+h)-f(x)-DF_x(h)\| \to 0, \quad h \to 0$$

DF_x = differential

$$DF = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \ddots & \\ \vdots & & \end{pmatrix}$$

Definition Rank of f at $x :=$ rank(DF_x).

If f smooth then

$$\text{Rank } (DF_{x_0}) = k \quad \Longrightarrow \quad \text{Rank } (DF_x) \geq k$$

for all x close to x_0.

Proof: $D_k(x) = k \times k$ submatrix of DF_x with $\det(D_k(x_0)) \neq 0 \Longrightarrow \det(D_k(x)) \neq 0$, for x close to x_0.

Inverse Function Theorem

Theorem Let X be open in U and $f \in C^1(X,V)$, $f(x_0)=y_0$. For the existence of $g \in C^1(Y,U)$ where Y is a neighborhood of y_0 such that

a) $f \circ g =$ identity near y_0
b) $g \circ f =$ identity near x_0
c) a) and b)

it is necessary and sufficient that there is a linear map A such that respectively

a') $f'(x_0)A = I_V$
b') $Af'(x_0) = I_U$
c') a') and b')

Condition c' implies that g is uniquely determined near y_0.

Proof idea: To solve $y=f(x)$ use

$$x_k = x_{k-1} + f'(x_0)^{-1}(y - f(x_{k-1}))$$

Prove $\sum(x_k - x_{k-1})$ converges for y near y_0.

Implicit Function Theorem

$$h(x,y)=0$$

$$\frac{\partial h}{\partial x} \text{ full rank} \quad \Longrightarrow \quad x=x(y) \text{ uniquely}$$

Example

$$h(x,y)=x^2+y^2-1, \quad h'_x=2x$$

So $x=x(y)$ uniquely except near $(0,\pm 1)$.

In fact $x=\sqrt{1-y^2}$, $x_0>0$ and $x=-\sqrt{1-y^2}$, $x_0<0$.

Discussion

Implicit F. T. $\Longrightarrow$ Inverse F. T. c).

$h(x,y)=y-f(x)$; $h'_x=f'_x \Longrightarrow x=x(y)$ uniquely

Inverse F. T. c) $\Longrightarrow$ Implicit F. T.

$$f(x,y)=(h(x,y),y)$$

$$f'=\begin{pmatrix} h'_x & h'_y \\ 0 & I \end{pmatrix} \quad \text{full rank}$$

So (x,y) locally determined by $(h,y)=(0,y)$

$$\Longrightarrow \quad x=x(y) \text{ uniquely locally}$$

Note

$$f(x,y)=0$$
$$f'_x\frac{\partial x}{\partial y}+f'_y=0$$
$$\frac{\partial x}{\partial y}=-(f'_x)^{-1}f'_y$$

Example

$$\begin{aligned} x_1^3-e^{x_2}+x_3^3-1&=0\\ x_1^2+x_2-x_3^2&=0 \end{aligned}$$

Are x_1,x_2 functions of x_3 around $(1,0,1)$?

Functions Between Manifolds

Definition

$$f\in C^\infty \Longleftrightarrow \psi\circ f\circ\varphi^{-1}\in C^\infty$$

Submanifolds

$$f_1(x_1, \ldots, x_n) = 0$$
$$\vdots$$
$$f_k(x_1, \ldots, x_n) = 0$$

determines an $n-k$ dimensional manifold near $\bar{x}$ if

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \ddots & \\ \vdots & & \end{pmatrix} \quad \text{has full rank } (= k) \text{ at} \bar{x}$$

$f_1, \ldots, f_k$ linearly independent at $\bar{x}$.

Tangent Vectors

Different definitions

- Define it only for manifolds embedded in R^n:
  $$\dot{x} = \lim_{h \to 0} \frac{\varphi(t) - \varphi(0)}{t}$$
  Velocity vectors in R^n.
- Coordinate free version.
  $$\varphi(t) \sim \psi(t) \text{ if } \varphi(0) = \psi(0) = x \text{ and}$$
  $$\lim_{h \to 0} \frac{\varphi(t) - \psi(t)}{t} \quad \text{in some chart}$$
  Tangent vectors at x = equivalance classes of curves with $\varphi(0) = x$

Our Definition

Derivative operator $X(f) : (f : M \mapsto R) \mapsto R$

$$X(\alpha f + \beta g) = \alpha X(f) + \beta X(g)$$
$$X(fg) = fX(g) + gX(f)$$

Example: Take any coordinate chart (U, φ) with coordinates x. Then

$$X_a = \sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial x_i}$$

is a tangent vector, where

$$X_a(f) = \sum_{i=1}^{n} \alpha_i \frac{\partial f(a)}{\partial x_i}$$

Theorem All tangent vectors are of this form

Proof: Taylors formula

$$f(x) = \underbrace{f(a)}_{X(\cdot)=0} + \underbrace{\sum (x_i - a_i) \frac{\partial f(a)}{\partial x_i}}_{\text{linear term}} + \underbrace{\text{higher terms}}_{X(\cdot)=0}$$

$$\Longrightarrow \quad X(f) = \sum_{i=1}^{n} \underbrace{X(x_i - a_i)}_{\alpha_i} \frac{\partial f(a)}{\partial x_i}$$

Different notation

$L_X(f) = X(f)$ Lie-derivative = fishermans derivative

Examples

$$\frac{\partial}{\partial \theta}; \quad \frac{\partial}{\partial z}; \quad z\frac{\partial}{\partial \theta} + \sin(\theta)\frac{\partial}{\partial z}$$

$$X = \begin{pmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Change coordinates $\beta = \frac{\partial z}{\partial x}\alpha$ or

$$\begin{pmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial z_1} & \cdots & \frac{\partial}{\partial z_n} \end{pmatrix} \frac{\partial z}{\partial x}$$

Example

$$z_1 = x_1$$
$$z_2 = x_1 + x_2$$
$$\frac{\partial}{\partial x_1} = \frac{\partial z_1}{\partial x_1}\frac{\partial}{\partial z_1} + \frac{\partial z_2}{\partial x_1}\frac{\partial}{\partial z_2} = \frac{\partial}{\partial z_1} + \frac{\partial}{\partial z_2}$$

Note that $x_1 = z_1$ does not imply $\frac{\partial}{\partial x_1} = \frac{\partial}{\partial z_1}$

Tangent Space

$$\bigcup_{p \in M} X_p$$

X_p close to Y_q if (p close to q) AND (α_{X_p} close to β_{Y_q})

Push Forward Operator. Derivative.

$$[f_* X](g) := X(g \circ f)$$

Alternative

$$f_* X = \frac{d}{dt} f(\varphi(t))|_{t=0}$$

$$(g \circ f)_* = g_* \circ f_*$$

Example

Next Week

$X(x)$ vector field.

$\sigma(t)$ solution curve to $\dot{x} = X(x)$ if

$$\frac{d}{dt} f(\sigma(t)) = X(f)|_{x=\sigma(t)}, \qquad \forall t, f$$

Lie-bracket. New vector field

$$[X, Y] = \frac{\partial Y}{\partial x} X - \frac{\partial X}{\partial x} Y$$

Frobenius Theorem

Given k vector fields. Find a k dimensional submanifold P such that

$$TP = \mathrm{span}(X_1, \ldots, X_k)$$

i.e. find lin. independent $f_1, \ldots, f_{n-k}$ such that

$$X_i(f_j) = 0 \quad \forall i, j$$

Theorem This can be done if and only if $X_1, \ldots, X_k$ involutive, i.e. $\exists \alpha_{ijl}$ such that

$$[X_i, X_j] = \sum_{l=1}^{k} \alpha_{ijl} X_l \quad \forall i, j = 1, \ldots, k$$

Controllable submanifold.

To think on: Construct a pair of vector fields in R^3 that does not have an integral manifold P.

Nonlinear Control Theory 94

Lecture 2

- Vector Fields
- Lie Brackets
- Distributions
- Frobenius Theorem

pp. 43-60 + Prop. 3.6 (skip Sect. 2.2.3)

Last Week

Manifold

IFTs

$\exists A: \ f'A = I \ \Leftrightarrow \ \exists g \in C^1: \ f \circ g = \text{id}$

$f(x,y) = 0: \ x = x(y) \ \Leftrightarrow \ f'_x$ full rank

Submanifolds $\{f_1 = 0, \ldots, f_k = 0\}$

Tangent Vector

$$X_p = \sum \alpha_i \frac{\partial}{\partial x_i}\Big|_p$$

$$\begin{pmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial z_1} & \cdots & \frac{\partial}{\partial z_n} \end{pmatrix} \frac{\partial z}{\partial x}$$

$$F_* X_p(g) = X_p(g \circ F)$$

Vector Fields

$$p \mapsto X_p$$

$$X = \sum_{i=1}^{n} X_i(p) \frac{\partial}{\partial x_i}$$

$X_i(p)$ smooth functions of p.

Notation:

$$X \sim \begin{pmatrix} X_1(x_1, \ldots, x_n) \\ \vdots \\ X_n(x_1, \ldots, x_n) \end{pmatrix}$$

Integral Curve

$\sigma(t)$ is an *integral curve* to X if

$$\sigma_* \left(\frac{\partial}{\partial t} \right) = X(\sigma(t)) \quad \forall t \in (t_0, t_1)$$

In local coordinates

$$\sigma(t) = \begin{pmatrix} \sigma_1(t) \\ \vdots \\ \sigma_n(t) \end{pmatrix}$$

$$\frac{\partial}{\partial t}(g(\sigma(t)) = X(\sigma(t))(g)$$

$$\sum \frac{\partial g}{\partial x_i} \frac{d\sigma_i}{dt} = \sum X_i(\sigma(t)) \frac{\partial g}{\partial x_i}$$

i.e.

$$\dot\sigma_1 = X_1(\sigma(t))$$
$$\vdots$$
$$\dot\sigma_n = X_n(\sigma(t))$$

A set of ODEs

Transformation Group, Flow

$$X^t(p) = \text{ solution to } \dot{x} = X(x), x(0) = p$$

X^t is smooth. $X^0 =$ id

$$L_X(g) = X(g) = \sum_{i=1}^{n} X_i \frac{\partial g}{\partial x_i} = \lim_{h \to 0} \frac{g(X^h(p)) - g(p)}{h}$$

$$L_{\alpha X + \beta Y} = \alpha L_X + \beta L_Y, \quad \alpha, \beta \in R$$

$$\dot{x} = f(x, u) \qquad f : M \times U \mapsto TM$$

Example

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

$$\dot{y} = \frac{\partial h}{\partial x}\dot{x} = \frac{\partial h}{\partial x}(f + gu) = L_{f+gu}h$$
$$= L_f h + u L_g h$$
$$y^{(k)} = (L_{f+gu})^k h$$

Change of Coordinates

As before

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix} = \frac{\partial z}{\partial x} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

Push Forward

As before

$$(F_* X)\big|_{F(p)}(g) = X_p(g \circ F)$$

$$F_* X = X(g \circ F) \circ F^{-1}$$

The Flow-Box Theorem

Theorem Let X be a vector field with $X(p) \neq 0$. Then there exists a coordinate chart $(U, x_1, \ldots, x_n)$ around p such that

$$X = \frac{\partial}{\partial x_1} \qquad \text{in } U$$

Integral curves : $x_i(q) =$ const., $i = 2, \ldots, n$.

Proof idea

$$T(a_1, a_2, \ldots, a_n) = X^{a_1}(0, a_2, \ldots, a_n)$$
$$T_{*0} = \text{ full rank}$$
$$T_*\left(\frac{\partial}{\partial r_1}\right) = X$$

Example

$$X = z_1\frac{\partial}{\partial z_2} - z_2\frac{\partial}{\partial z_1} = \frac{\partial}{\partial \varphi}$$

$$x_1 = \arctan(z_2/z_1)$$
$$x_2 = \sqrt{z_1^2 + z_2^2}$$

$$\frac{\partial}{\partial x_1} = \frac{\partial z_1}{\partial x_1}\frac{\partial}{\partial z_1} + \frac{\partial z_2}{\partial x_1}\frac{\partial}{\partial z_2} = -z_2\frac{\partial}{\partial z_1} + z_1\frac{\partial}{\partial z_2} = X$$

Lie-Brackets

$$[X,Y]_p(f) = X_p(Y(f)) - Y_p(X(f))$$

$$X \sim \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}; \qquad Y \sim \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$$

$$[X,Y] = \frac{\partial Y}{\partial x}X - \frac{\partial X}{\partial x}Y$$

Example

$$X = \cos\phi\frac{\partial}{\partial r} + r\frac{\partial}{\partial \phi} \sim \begin{pmatrix} \cos\phi \\ r \end{pmatrix}$$
$$Y = r\frac{\partial}{\partial r} + \frac{\partial}{\partial \phi} \sim \begin{pmatrix} r \\ 1 \end{pmatrix}$$
$$[X,Y] = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \cos\phi \\ r \end{pmatrix} - \begin{pmatrix} 0 & -\sin\phi \\ 1 & 0 \end{pmatrix}\begin{pmatrix} r \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} \cos\phi - \sin\phi \\ -r \end{pmatrix} \sim (\cos\phi - \sin\phi)\frac{\partial}{\partial r} - r\frac{\partial}{\partial \phi}$$

Lie-Brackets

Prop. 3.6 $\quad \dot{x} = g_1 u_1 + g_2 u_2$

$$(u_1(t), u_2(t)) = \begin{cases} (1,0) & t \in [0,h) \\ (0,1) & t \in [h,2h) \\ (-1,0) & t \in [2h,3h) \\ (0,-1) & t \in [3h,4h) \end{cases}$$
$$x(4h) = x_0 + h^2[g_1,g_2] + O(h^3)$$

Trotters Product Formula

$$\Phi^t_{[X,Y]} = \lim_{n\to\infty}\left(\Phi_{-Y}^{\sqrt{\frac{t}{n}}}\Phi_{-X}^{\sqrt{\frac{t}{n}}}\Phi_{Y}^{\sqrt{\frac{t}{n}}}\Phi_{X}^{\sqrt{\frac{t}{n}}}\right)^n$$

Proof sketch

$$\left(1 + \frac{tf}{n} + o\left(\frac{tf}{n}\right)\right)^n \to e^{tf}$$

Lie-Bracket Formulas

$$[fX, gY] = fg[X,Y] + fX(g)Y - gY(f)X$$
$$[X,Y] = -[Y,X]$$
$$[X_1,[X_2,X_3]] + [X_2,[X_3,X_1]] + [X_3,[X_1,X_2]] = 0$$
$$F_*[X,Y] = [F_*X, F_*Y]$$
$$L_XY = [X,Y] = \lim_{h\to 0}\frac{1}{h}[X_*^{-h}Y - Y]$$

Campbell-Baker-Hausdorff Formula

$$X_*^{-h}Y = \sum_{n=0}^{\infty}\mathrm{ad}_X^n Y\frac{h^n}{n!} = Y + h[X,Y] + \frac{h^2}{2}[X,[X,Y]]\ldots$$

related to

$$e^Ae^B = e^C; \qquad C = A + B + \frac{1}{2}[A,B] + \ldots$$

Vector Fields, Summary

A vector field X is associated with

a) A system of differential equations

$$\frac{dx}{dt} = X(x)$$

b) A flow $\Phi^t : M \mapsto M, t \in [t_0, t_1]$, where $\sigma(t) = \Phi^t(x)$ is the solution to

$$\frac{d\sigma}{dt} = X(\sigma), \qquad \sigma(0) = x$$

c) A directional derivative

$$X_x f = \frac{d}{dt} f(\Phi^t(x))\Big|_{t=0}$$

d) A derivation of the algebra $C^\infty(M)$.

e) A partial differential operator

$$X = \sum X_j \frac{\partial}{\partial x_j}$$

$a \longrightarrow b$ solution to differential equations

$b \longrightarrow c$ direct

$c \longrightarrow d$ direct

$d \longrightarrow e$ proposition

$e \longrightarrow a$ direct

Park Your Car Using Lie-Brackets!

$$\begin{aligned} (x,y) &: \text{ position} \\ \phi &: \text{ direction of car} \\ \theta &: \text{ direction of wheels} \\ (x, y, \phi, \theta) &\in R^2 \times S^1 \times [\theta_{\text{min}}, \theta_{\text{max}}] \end{aligned}$$

$$\text{Steer} := \frac{\partial}{\partial \theta}$$

$$\text{Drive} := \cos(\phi+\theta)\frac{\partial}{\partial x} + \sin(\phi+\theta)\frac{\partial}{\partial y} + \sin(\theta)\frac{\partial}{\partial \phi}$$

$$\begin{aligned} &[\text{Steer}, \text{Drive}] = \\ &[\frac{\partial}{\partial \theta}, \cos(\phi+\theta)\frac{\partial}{\partial x} + \sin(\phi+\theta)\frac{\partial}{\partial y} + \sin(\theta)\frac{\partial}{\partial \theta}] \\ &= -\sin(\phi+\theta)\frac{\partial}{\partial x} + \cos(\phi+\theta)\frac{\partial}{\partial y} + \cos(\theta)\frac{\partial}{\partial \phi} \\ &:= \text{Wriggle} \end{aligned}$$

Define Slide $:= -\sin(\phi)\frac{\partial}{\partial x} + \cos(\phi)\frac{\partial}{\partial y}$.

$$\text{Slide}^t(x, y, \phi, \theta) = (x - t\sin(\phi), x + t\cos(\phi), \phi, \theta)$$

An easy calculation (exercise) shows that

$$[\text{Wriggle}, \text{Drive}] = \text{Slide}$$

Fundamental Parking Theorem You can get out of a parking lot that is larger than the car. Use the following control: Wriggle, Drive, –Wriggle (this requires a cool head), –Drive (repeat).

Proof: Trotters Product Formula

Linear Systems

$$\dot{x} = Ax + Bu = f(x) + g(x)u$$

$$[f, g] = [Ax, B] = 0 - AB$$
$$[g, [f, g]] = 0$$
$$[f, [f, g]] = [Ax, -AB] = A^2B$$
$$\vdots$$
$$\text{Ad}_f^k g = \underbrace{[f, [f, \ldots, [f, g]]]}_{k \text{ Lie-brackets}} = (-1)^k A^k B$$

σ related to controllability indices

Nonholonomic Systems

$$\dot{x} = g_1u_1 + \ldots + g_mu_m$$
$$F_1 = \text{span}\{g_1, \ldots, g_m\}$$
$$F_2 = F_1 + [F_1, F_1] = \text{span } \{g_i, [g_i, g_j]\}$$
$$\vdots$$
$$F_i = F_{i-1} + [F_1, F_{i-1}] \quad \text{brackets to level } i-1$$

If $\dim F_i$ constant on neighborhood $\forall i$, define

$$m_i = \dim F_i \quad \text{growth vector}$$
$$\sigma_1 = m_1$$
$$\sigma_i = m_i - m_{i-1} \quad \text{relative growth vector}$$

$$\underbrace{F_1}_{\dim = m} \subset F_2 \subset \ldots \subset \underbrace{F_p}_{\dim = n}$$
$p =$ degree of nonholonomy
$\sigma = (\sigma_1, \ldots, \sigma_p) =$ relative growth vector

Example Car with N-trailers has

$$p = N + 3; \quad \sigma = (2, 1, \ldots, 1)$$

Lie Algebra

Vector space V with operation $V \times V \mapsto V$, denoted $[\cdot, \cdot]$, satisfying:

i) $[\alpha_1v_1 + \alpha_2v_2, w] = \alpha_1[v_1, w] + \alpha_2[v_2, w]$
ii) $[v, w] = -[w, v]$
iii) $[v_1, [v_2, v_3]] + [v_2, [v_3, v_1]] + [v_3, [v_1, v_2]] = 0$

Example The $n \times n$ matrices form a Lie-algebra, with $[A, B] := AB - BA$ (check).

Jacobi identity: 12 terms, each occurring twice with different signs

Example $Gl(n)$, the invertible $n \times n$ matrices, do not form a Lie-Algebra (why?)

Distributions

$p \mapsto$ linear subspace of T_pM (smoothly)

$$D = \text{span}\{X_1, \ldots, X_k\}$$
$$D_1 + D_2 \qquad D_1 \cap D_2$$

D is *involutive* if

$$X, Y \in D \implies [X, Y] \in D$$

Enough to check on basis.

$$[X_i, X_j] = \sum_{l=1}^{k} c_{ij}^l X_l \quad \forall i, j$$

Examples

Example A single vector field $\{X_1\}$ is always involutive, since $[X_1, X_1] = 0$.

Example An n-dimensional distribution $\{X_1, \ldots, X_n\}$ (n=dimension of manifold) is always involutive.

Example

$$X_1 = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + x_3 \frac{\partial}{\partial x_3}$$
$$X_2 = \frac{\partial}{\partial x_3}$$

is involutive since $[X_1, X_2] = -X_2$

Integral Manifold

$$T_qP = D(q)$$

$\exists$ integral manifold $\Longrightarrow [X_i, X_j] \in D$ i.e. D involutive

Frobenius Th.: Reverse implication also true

Frobenius Theorem

Let D be an involutive, constant dimensional, distribution

Version 1 For each p there is a coordinate chart (U, ϕ) so that $\phi(p) = 0$ and

$$P_a = \{q \in U \mid \phi_{k+1}(q) = a_{k+1}, \ldots, \phi_n(q) = a_n\}$$

are integral manifolds to D.

Version 2 For each p there is a coordinate chart $(U, x_1, \ldots, x_n)$ such that

$$D = \text{span}\left\{\frac{\partial}{\partial x_1}, \ldots \frac{\partial}{\partial x_k}\right\} \quad \text{in } U$$

D is then called a *flat* distribution.

Classical PDE result

Proof Idea for Version 2

Step 1 : If $[X_i, X_j] = 0, \ \forall i, j$ then

$$T(a_1, \dots, a_n) = X_1^{a_1} X_2^{a_2} \cdots X_k^{a_k}(0, \dots, 0, a_{k+1}, \dots, a_n)$$

defines a coordinate transformation (Ex. 1.8) such that

$$T_* \frac{\partial}{\partial a_i} = X_i \qquad i = 1, \dots, k$$

Step 2: Reduce to this case by projecting away everything outside the distribution, see book

$$\pi : R^n \to R^k$$

Example

$$X_1 = x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + x_3 \frac{\partial}{\partial x_3} = r \frac{\partial}{\partial r}$$
$$X_2 = \frac{\partial}{\partial x_3}$$

D involutive (checked before).

$$z_1 = x_2 / x_1$$
$$z_2 = x_2$$
$$z_3 = x_3$$

P is given by $z_1 =$ constant, i.e. $x_2 = c x_1$.

Software

- Condens (Macsyma), Maryland.
- Maple
- Omsim (Simnon, Gnans)
- Matlab. Nonlinear Toolbox

Much left to do

Homework Problems

Problem Set 1 Hand in solutions to 10 of the exercises 2.1-2.18.

Due: Monday Feb 7, 13.15

Problem Set 2 Find an (open loop) control that steers the cars out from a tight parking lot (e.g. 1.25 times the car length). Simulate.

Alternative (harder): Find a control that backs the car with $N = 1$ trailer into a parking lot (sideways).

Hint can be obtained on request.

Due: Monday Feb 21, 13.15

Next Week

Controllability

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

$C =$ smallest Lie subalg. containing $\{f, g_1, \ldots, g_m\}$
$C_0 =$ smallest Lie subalg. containing $\{g_1, \ldots, g_m\}$
and satisfies $[f, X] \in C_0, \ \forall X \in C_0$

Accessibility

$$\dim C = n \implies \text{can reach open set}$$

If $f = 0$ then equivalent to "controllability", i.e. that one can steer from x_0 to every x_1 in a neighborhood.

Controllable submanifold: Can find local coordinates $x_1, \ldots, x_n$ such that

$$x(t) \in S_{x_0} \quad \forall u(t); \qquad S_{x_0} \text{ integral manifold to } C$$

Nonlinear Control Theory 94

Lecture 3

- Local Controllability
  - Linear Controllability
  - Linearization
  - Nonlinear Rank Conditions
  - Control Submanifolds
  - System Decomposition

pp. 73-93 + 101-104.

Last Week

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

Flow $X^t(x)$ or $e^{tf}(x)$

Lie brackets $[X, Y]$

Distribution $\{X_1, \dots, X_k\}$

Frobenius theorem $[X_i, X_j] \in D$. Change coordinates so

$$D = \text{span}\left\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_k}\right\}$$

$$S_a = \left\{x \mid x_{k+1} = a_{k+1}, \dots, x_n = a_n\right\}$$

Review of Linear Controllability

$$\dot{x} = Ax + Bu$$

Controllability

$$0 \to x(T)$$
$$x(0) \to 0$$
$$x(0) \to x(T)$$
$$T > 0 \text{ arbitrary}$$

Rank Condition

$$W_n = \begin{pmatrix} B & AB & \dots & A^{n-1}B \end{pmatrix} \quad \text{full rank}$$

$$x(T) \in e^{AT}x(0) + \text{Im}(W_n)$$

Linear Controllability

Controllability indices. Crate II, row by row

	b_1	$\cdots$	b_m
I	$\times$	$\cdots$	$\times$
A	$\times$		
A^2	$\times$		
$\vdots$			
A^{n-1}			

$\kappa = \kappa_1 \geq \dots \geq \kappa_m$ length of columns

W_κ full rank, κ controllability index

$$\dot{x} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} x + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} u$$

where (A_{11}, B_1) controllable.

$$E\dot{x} = Ax + Bu$$

Linearization at (x_0, u_0)

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

Prop. 3.3 Suppose $f(x_0) + g(x_0)u_0 = 0$ and let U contain a neighborhood of u_0. If

$$\begin{aligned}\dot{z} &= Az + Bv\\ A &= \frac{\partial f}{\partial x}(x_0) + \frac{\partial g}{\partial x}(x_0)u_0\\ B &= g(x_0)\end{aligned}$$

is controllable, then for all $T, \epsilon > 0$

$$X_{T,\epsilon} = \{x(T);\ \ |u - u_0| \le \epsilon\}$$

contains a neighborhood of x_0.

Proof Can assume $x_0 = 0$, $u_0 = 0$. Will use inverse function theorem on

$$\xi \mapsto x(T, \xi); \quad u(t, \xi_1, \ldots, \xi_n) = \xi_1 v^1(t) + \ldots + \xi_n v^n(t)$$

If $Z(T) = \frac{\partial x}{\partial \xi}(t, \xi)|_{\xi=0,t=T}$ has full rank the theorem follows. But $Z(t)$ satisfies

$$\dot{Z}(t) = AZ(t) + B\begin{pmatrix} v^1(t) & \ldots & v^m(t)\end{pmatrix}, \quad Z(0) = 0$$

and by controllability the columns of $Z(T)$ can be made independent by choice of $v^i(t)$.

Exercise Is the theorem true also for

$$\dot{x} = f(x, u)?$$

Example

$$\dot{x} = \begin{pmatrix} x_2 \\ \sin x_1 \end{pmatrix} + \begin{pmatrix} 0 \\ \cos x_1 \end{pmatrix} u$$

$$\dot{z} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} z + \begin{pmatrix} 0 \\ 1 \end{pmatrix} v$$

Linearization is controllable at 0.

Example Rolling penny

$$\frac{d}{dt}\begin{pmatrix} \phi \\ \psi \\ x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} 0 \\ 1 \\ \cos\phi \\ \sin\phi \end{pmatrix} u_2$$

$$\dot{z} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} v_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} v_2$$

Linearization is not controllable at 0.

Linearization around $(x_0(t), u_0(t))$

Assume $\dot{x}_0(t) = f(x_0(t), u_0(t))$. The linearization of $\dot{x} = f(x, u)$ is then given by

$$\begin{aligned}\dot{z} &= A(t)z + B(t)v(t)\\ A(t) &= \frac{\partial f}{\partial x}(x_0(t), u_0(t))\\ B(t) &= \frac{\partial f}{\partial u}(x_0(t), u_0(t))\end{aligned}$$

With $u(t) = u_0(t) + \epsilon v(t)$ we get

$$x(t) = x_0(t) + \epsilon z(t) + O(\epsilon^2)$$

Equivalent with the so called variational equations.

Nonlinear Controllability

$$\dot{x} = g_1 u_1 + g_2 u_2$$

What about ϵ^3-terms etc?

$$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x) u_i = \sum_{i=0}^{m} g_i(x) u_i$$

where $g_0 = f$ and $u_0 \equiv 1$.

Can one control in e.g. $[f, g]$-direction?

Assumptions

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

where $u(t) \in U$ for some set U and define

$$\mathcal{F} = \left\{ f(x) + g(x)u \mid u \in U \right\}$$

Assumption 3.1 Linear span of $\mathcal{F}$ contains $\{f, g_1, \ldots, g_m\}$.

Example Two inputs u_1, u_2 where

$$U = \{(u_1, u_2) \mid u_1 \geq 0, u_2 \in \{0, 1\}\}$$

satisfies Ass. 3.1.

Piecewise constant controls.

Local Accessibility

Definition

$$R^V(x_0, T) = \{x(T) \;:\; x(t) \in V,\; 0 \leq t \leq T\}$$
$$R^V_T(x_0) = \bigcup_{0 \leq \tau \leq T} R^V(x_0, \tau)$$

Definition A system is *locally accessible* at x_0 if $R^V_T(x_0)$ contains an open set for any $T > 0$ and any $V \ni x_0$.

The Accessibility Algebra C

Definition The smallest *subalgebra* C with

$$\{f, g_1, \ldots, g_m\} \in C$$

(subalgebra = linear subspace such that $v, w \in C \Rightarrow [v, w] \in C$). "Involutive closure".

Lie-bracket tree

Example

$$[[f, g_1], [g_1, g_2]] + [g_1, [g_2, [f, g_1]]] + [g_2, [[f, g_1], g_1]] = 0$$

Example

$$\dot{x} = Ax + Bu$$

Example, bilinear systems

$$\dot{x} = Ax + \sum_{i=1}^{m} u_j B_j x$$

Direct calculation shows that $C = \{Mx\}$ where

$$M = [D_k, [\ldots [D_2, D_1]]]$$

for some $D_i \in \{A, B_1, \ldots, B_m\}$.

Finite dimensional Lie algebra.

Example

$$\dot{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \qquad 0 \leq u \leq 1$$

Dim $C = 2$. LS.

Example The car.

$$C = \text{span}\{g_1, g_2, [g_1, g_2], [g_2, [g_1, g_2]]\}$$

Dim $= 4$ so locally accessible.

Rank Condition

Theorem 3.9 If dim $C(x_0) = n$, then the system is locally accessible at x_0 for all $T > 0$ and all $V \ni x_0$.

Proof

$$(t_n, \ldots, t_1) \to X_n^{t_n} \circ \ldots \circ X_1^{t_1}(x_0)$$

full rank in $0 \leq \sigma_i < t_i < \epsilon_i$. Here $X_i = f + gu_i$. Then use the inverse function theorem.

Necessity of Rank Condition

Local accessibility at all $x_0 \Rightarrow$ dim$(C) = n$ in open and dense set.

Example

$$\dot{x}_1 = e^{-1/|x_2|}$$
$$\dot{x}_2 = u$$
$$g(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$f(0) = [f, g](0) = \ldots = 0$$

Locally accessible everywhere. Satisfies accessibility rank condition except at line $x_2 = 0$.

Control Manifold

Theorem Assume that dim$(C) = k < n$ in a neighborhood of x_0 and that

$$S_{x_0} = \{q \mid x_{k+1}(q) = x_{k+1}(x_0), \ldots, x_n(q) = x_n(x_0)$$

are integral manifolds of C. Then

$$R_T^V(x_0) \subset S_{x_0}.$$

Moreover $R_T^V(x_0)$ is open as a subset of S_{x_0}.

Local Strong Accessibility

Definition A system is *locally strongly accessible* at x_0 if for all neighborhoods $V \ni x_0$ and all sufficiently small $T > 0$ the set $R^V(x_0, T)$ contains an open set.

Can reach open set *for fixed time* T.

Definition The strong accessibility algebra C_0 is the smallest subalgebra containing $\{g_1, \ldots, g_m\}$ which is closed under bracketing with f.

$$\begin{aligned}
&g_1, \ldots, g_m \in C_0 \\
&X \in C_0 \;\Rightarrow\; [f, X] \in C_0 \\
&X, Y \in C_0 \;\Rightarrow\; [X, Y] \in C_0
\end{aligned}$$

Same tree as before but without drift field f.

$C = \text{span}\{f, C_0\}$.

Local Strong Accessibility

Theorem If dim $C_0(x_0) = n$, then the system is locally strongly accessible at x_0 for all $T > 0$ and all $V \ni x_0$.

Proof Add $\dot{t} = 1$ as equation and reduce to previous case. See book.

Control Manifold C_0

Theorem Assume dim $C_0 = k < n$ around x_0, then one can change coordinates such that

$$\begin{aligned}
\dot{x}_1 &= f_1(x) + g_1(x)u \\
&\vdots \\
\dot{x}_k &= f_k(x) + g_k(x)u \\
\dot{x}_{k+1} &= 0 \text{ or } 1 \\
\dot{x}_{k+2} &= 0 \\
&\vdots \\
\dot{x}_n &= 0
\end{aligned}$$

Here $\dot{x}_{k+1} = 0$ if $f \in C_0$ and $\dot{x}_{k+1} = 1$ if $f \notin C_0$.

Moreover, the system restricted to $x_1, \ldots, x_k$ is locally strongly accessible.

Example, Satellite

$$J\dot{w} = S(w)Jw + Bu$$

- Lin. is controllable only if $\dim(B) = 3$.
- $\dim(B) = 2$: LSA iff

$$\dim \operatorname{span} \{b_1, b_2, S(w)J^{-1}w;\ \ w \in \operatorname{span}(b_1, b_2)\} = 3$$

which is often true.

- can be LSA even if $\dim(B) = 1$.

Local Controllability at x_0

Definition $\dot{x} = f + gu$ is *small time locally controllable* (STLC) qat x_0 if for every neighborhood V of x_0 there is a neighborhood W to x_0 so that for every sufficiently small time $T > 0$ every state $x_1 \in W$ can be reached from x_0 in exactly time T.

$$x_0 \in W \subset R^V(x_0, T)$$

Example

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

NOT locally controllable at 0.

Controllability Theorems, 1

General U:

Theorem If $\mathcal{F} = -\mathcal{F}$ + LSA, then $R_T^V(x_0)$ contains open neighborhood of x_0, $\forall T > 0, V \ni x_0$.

Proof

Controllability Theorems, 2

If U contains a neighborhood of $u = 0$, then STLC if

- $f = 0$ + locally strongly accessible at x_0.
- $f \in \operatorname{span}\{g_1, \ldots, g_m\}$ + LSA at x_0.
- $\dim \operatorname{ad}_f^k g = n$ and $f(x_0) = 0$ (controllable linearization).
- "odd system" + $f(x_0) = 0$ + LSA at x_0.

A system is "odd" if all brackets with an even number of g:s are linear combination of brackets with a smaller number of g:s (Sussman).

"even brackets are evil brackets"

Example

$$\dot{x}_1 = x_2^2$$
$$\dot{x}_2 = u$$

is not an odd system, because $[g,[f,g]]$ is not a linear combination of f, g, $[f,g]$, $[f,[f,g]]$, . . . at $x = 0$.

Example

$$\dot{x}_1 = \sin(x_2) - x_2$$
$$\dot{x}_2 = u$$

is an odd system; $[g,[f,g]] = 0$ and g and $[g,[g,[f,g]]]$ span R^2.

Decompositions

V invariant subspace to A if $AV \subset V$.

$$\dot{x} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} x + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} u$$

If $AV \subset V$ and $\mathrm{Im} B \subset V$ then

$$\dot{x} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} x + \begin{pmatrix} B_1 \\ 0 \end{pmatrix} u$$

Local Nonlinear Decomposition

Let D be an involutive distribution which contains g_j and is invariant under $\dot{x} = f + gu$

$$g_j \in D$$
$$[f, D] \subset D$$
$$[g_j, D] \subset D, \quad j = 1, \ldots, m$$

Then by Frobenius

$$\dot{x}^1 = f^1(x^1, x^2) + g(x^1, x^2)u$$
$$\dot{x}^2 = f^2(x^2)$$

For every fixed x^2 the system in x^1 is locally strongly accessible.

Proof $[f, \frac{\partial}{\partial x_i}] \in D$, $i = 1, \ldots, k$ etc. See book

Research Topics

- Quantitative controllability. Relate to model reduction.
-
$$E\dot{x} = f(x) + g(x)u$$
Includes $f(x, \dot{x}) + g(x, \dot{x})u = 0$.
- Numerical algorithms

Next Week

One-forms. Linear functions of tangent vectors.

$$\sigma = \sum \sigma_i dx_i$$

Co-distributions

$$\{\sigma_1, \ldots, \sigma_k\}$$

Local observability

$$\mathcal{O} = \{L_{X_1} \ldots L_{X_i} h\}$$

$$\dim(d\mathcal{O}) = n$$

Nonlinear Kalman Decomposition

Nonlinear Control Theory 94

Lecture 4

- Cotangent Vectors
- One-forms
- Codistributions
- Local Observability
- Nonlinear Kalman Decomposition

pp. 61-66, 93-116.

Last Week

$$C = \{f, g_i, [f, g_i], \ldots\}$$
$$C_0 = \{g_i, [f, g_i], \ldots\}$$

dim $C = n \;\Rightarrow\;$ Can reach open set in $t \le T$

dim $C_0 = n \;\Rightarrow\;$ Can reach open set in $t = T$

$$\dot{x}^1 = f^1(x^1, x^2) + g(x^1, x^2)u$$
$$\dot{x}^2 = f^2(x^2)$$

For fixed x_2 the system in x_1 is LSA.

$$S_{x_0} = \{q \mid x_{k+1}(q) = x_{k+1}(x_0), \ldots, x_n(q) = x_n(x_0)$$

integral manifolds of C. *Control Submanifold*

$$R_T^V(x_0) \subset S_{x_0}.$$

Cotangent Vectors, $T_p^* M$

Linear form

$$\sigma \;:\; T_pM \mapsto R$$

Determined by $\sigma_i := \sigma\left(\frac{\partial}{\partial x_i}\right)$ so

$$\sigma\Big(\sum X_i \frac{\partial}{\partial x_i}\Big) = \sum \sigma_i X_i$$

$$\sigma = \sum_{i=1}^{n} \sigma_i \, dx_i = \begin{pmatrix} \sigma_1 & \ldots \sigma_n \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}$$

Here $\{dx_i\}$ is the dual base defined by

$$dx_i\Big(\frac{\partial}{\partial x_j}\Big) = [i = j]$$

Cotangent Vectors

Notation

$\sigma(X) =$

$$\begin{pmatrix} \sigma_1 & \ldots & \sigma_n \end{pmatrix} \underbrace{\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \end{pmatrix}}_{=I} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$= \sum \sigma_i X_i$

Often $\sigma(X)$ is denoted $< \sigma, X >$.

Example

$$\sigma = x_2^2 dx_1 + x_2 dx_2$$
$$X = x_1 \frac{\partial}{\partial x_1} + \sin(x_2) \frac{\partial}{\partial x_2}$$
$$\sigma(X) = x_1 x_2^2 + x_2 \sin(x_2)$$

Differential df

Definition If $f : M \mapsto R$ then the *differential* df_p is a cotangent vector defined by

$$df_p(X_p) = X_p(f)$$

By evaluating df_p on $X_p = \frac{\partial}{\partial x_i}|_p$ we see that

$$df_p = \frac{\partial f}{\partial x_1}(p)dx_1 + \dots \frac{\partial f}{\partial x_n}(p)dx_n$$

Coordinate Change

$$\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = \frac{\partial x}{\partial z} \begin{pmatrix} dz_1 \\ \vdots \\ dz_n \end{pmatrix}$$

Example $z_1 = \sin(x_1 + x_2); \quad z_2 = x_2$

$$\begin{aligned} dz_1 &= \cos(x_1 + x_2)dx_1 + \cos(x_1 + x_2)dx_2 \\ dz_2 &= dx_2 \end{aligned}$$

Note that $z_i = x_i$ gives $dz_i = dx_i$.

Pull-back F^*

Remember that

$$F_*X|_{F(p)}(f) = X_p(f \circ F)$$

We now define $F^*\sigma$ by

$$(F^*\sigma)X := \sigma(F_*X)$$

Transformed by *Jacobian* $\frac{\partial F}{\partial x}$:

$$F^*\sigma = \sigma_{F(p)} \frac{\partial F}{\partial x}$$

Example $z = F(x_1, x_2) = x_1^2 x_2$ and $\sigma = zdz$

$$F^*\sigma = \begin{pmatrix} z \end{pmatrix} \begin{pmatrix} 2x_1x_2 & x_1^2 \end{pmatrix} = 2x_1^3x_2^2dx_1 + x_1^4x_2dx_2$$

One-forms

$p \mapsto T_p^*M$ smoothly, ie $\sigma_i(p)$ smooth

Dual to vector fields

Example $\sigma = df$ defines a one-form by

$$df(X) = X(f)$$

Such a σ is called *exact*

$$F^*df = d(f \circ F)$$

Lie-Derivative of One-forms

$$L_X\sigma := \lim_{h\to 0} \frac{(X^h)^*\sigma - \sigma}{h}$$

Good exercise to show that the row vector representing $L_X\sigma$ is given by

$$L_X\sigma = \left(\frac{\partial\sigma^T}{\partial x} X\right)^T + \sigma\frac{\partial X}{\partial x}$$

(Hint: Calculate the Jacobian of X^h).

More Formulas

$$L_{\alpha X+\beta Y}\sigma = \alpha L_X\sigma + \beta L_Y\sigma$$
$$L_X df = d(L_X f)$$
$$L_X < \sigma, Y > = < L_X\sigma, Y > + < \sigma, L_X Y >$$

Proof of the last equality:

$$L_X < \sigma, Y > = \sum_{i,j} X_j \frac{\partial}{\partial x_j}(\sigma_i Y_i)$$
$$< L_X\sigma, Y > = \sum_{i,j} X_j \frac{\partial\sigma_i}{\partial x_j} Y_i + \sigma_i \frac{\partial X_i}{\partial x_j} Y_j$$
$$< \sigma, L_X Y > = \sum \sigma_i (\frac{\partial Y_i}{\partial x_j} X_j - \frac{\partial X_i}{\partial x_j} Y_j)$$

Co-distributions

$$P(q) = \text{span}\{\sigma_1(q), \ldots, \sigma_l(q)\}$$

Dual to distributions

Definitions

$$\text{ker } P = \text{span}\{X \mid \sigma(X) = 0;\ \forall\sigma \in P\}$$
$$\text{ann } D = \text{span}\{\sigma \mid \sigma(X) = 0;\ \forall X \in D\}$$

$$D \subset \text{ker}(\text{ann}(D))$$
$$P \subset \text{ann}(\text{ker}(P))$$

Equality if D and P are constant dimensional.

Involutivity and Frobenius

Definition A codistribution P is *involutive* if ker P is involutive.

Frobenius: P is involutive iff $\exists f_1, \ldots, f_l$ (coordinates) such that

$$P(q) = \text{span}\{df_1(q), \ldots, df_l(q)\}$$

In words: P is spanned by exact one-forms.

$P_1 + P_2$, $P_1 \cap P_2$, F^*P

If you ever see something like

$$f(x) dx_1 \wedge dx_2 \ldots \wedge dx_k$$

it is a k-form.

Differential forms. Tensors.

End of Differential Geometry

Observability

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

Definition: x_1 and x_2 are called *indistinguishable* ($x_1 \; I \; x_2$) if

$$y(t, x_1, u) = y(t, x_2, u) \quad \forall t \; \forall u$$

(for those t where y is well-defined). The system is called *observable* if

$$x_1 \; I \; x_2 \implies x_1 = x_2$$

Depends on u. Depends on $\{f, g, h\}$.

Local Observability

$x_1 I^V x_2$ if $x(t) \in V$ for all t and give same y

Local observability at x_0: For every sufficiently small neighborhood $V \ni x_0$ one has that $x_1 \; I^V \; x_2$ implies $x_1 = x_2$.

Example

$$\dot{x}_1 = x_2 u$$
$$\dot{x}_2 = 0$$
$$y = x_1$$

Observable, e.g with $u \equiv 1$. Not observable with $u \equiv 0$.

Example

$$\dot{x}_1 = \varphi(x_2)$$
$$\dot{x}_2 = u$$
$$y = x_1$$

Observable, not locally observable at $x_2 < 0$.

The Observation Space $\mathcal{O}$

Definition $\mathcal{O}$ is the smallest linear space containing $h_1, \ldots, h_p$ and all repeated Lie derivatives:

$$L_{X_1} L_{X_2} \ldots L_{X_k} h_j, \qquad k = 1, 2, \ldots$$

with X_i in the set $\{f, g_1, \ldots, g_m\}$.

Remember $y_i^{(k)} = L^k_{f+gu} h_i$

$\mathcal{O}$ contains the output functions and all derivatives of the output functions *along all possible system trajectories*, i.e. as above but with $Z_i = f + gu^i$.

Can also define $\mathcal{O}$ using all $X_i \in C$, the accessibility algebra.

Same tree as before.

Observability Codistribution $d\mathcal{O}$

$$d\mathcal{O}(q) = \mathrm{span}\{dH(q) \mid H \in \mathcal{O}\}$$

Main Theorem Assume

$$\dim \, d\mathcal{O}(x_0) = n$$

then the system is locally observable at x_0.

Proof: Assume $x_1 I^V x_2$ then

$$h_i(Z_k^{t_k} \circ \ldots \circ Z_1^{t_1})(x_1) = h_i(Z_k^{t_k} \circ \ldots \circ Z_1^{t_1})(x_2)$$

with $Z_i = f + gu^i$. Differentiation w.r.t. $t_k, t_{k-1}, \ldots, t_1$ at 0 gives

$$L_{Z_1} L_{Z_2} \ldots L_{Z_k} h_i(x_1) = L_{Z_1} L_{Z_2} \ldots L_{Z_k} h_i(x_2)$$

This means that $H(x_1) = H(x_2)$ for all $H \in \mathcal{O}$. But this gives that $x_1 = x_2$ by the linear independence of $dH_1(x_0), \ldots, dH_n(x_0)$ and the inverse function theorem.

Example

$$\dot{x}_1 = x_2 u$$
$$\dot{x}_2 = 0$$
$$y = x_1$$

Here $h = x_1$ and $L_g h = x_2$ span $\mathcal{O}$ and

$$\dim d\mathcal{O} = \dim \{dx_1, dx_2\} = 2$$

so the system is locally observable.

With $u \neq 0$ we in fact have

$$x_1 = y; \qquad x_2 = \dot{y}/u$$

Nonobservable Submanifold

Theorem If dim $d\mathcal{O} = k < n$ around x_0, then by Frobenius we can find coordinates so that

$$S_{x_0} = \{q \in U \mid x_i(q) = x_i(x_0), i = n-k+1, \ldots, n\}$$

is an $n-k$ dimensional integral manifold of ker $d\mathcal{O}$, *an unobservable manifold.* Locally

$$\{x \mid x\, I^V x_0\} = S_{x_0} \cap V$$

Proof: Change coordinates using $H_1, \ldots, H_k$. See book.

A Converse

Corollary If the system is locally observable near x_0 then dim $d\mathcal{O}(x) = n$ for all x in an open and dense subset near x_0.

Proof: Use previous theorem.

Remark: If the system is LA and analytic then $d\mathcal{O}$ is constant dimensional. Hence the rank condition at x_0 is both necessary and sufficient for LO at x_0.

Linear Case

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

$$\mathcal{O} = \{c_i x, c_i A x, \ldots, c_i A^{n-1} x\} + \text{ constant functions}$$

Hence rank condition becomes

$$\text{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

Invariant Distributions

A distribution D is invariant for $\dot{x} = f(x)$ if

$$[f, X] \in D, \quad \forall X \in D$$

Theorem Let D be an involutive distribution of constant dimension k which is invariant for $\dot{x} = f(x)$. Then we can change coordinates $x = (x_1, \ldots, x_n)$ so that with $x^1 = (x_1, \ldots, x_k)$ and $x^2 = (x_{k+1}, \ldots, x_n)$

$$\dot{x}^1 = f^1(x^1, x^2)$$
$$\dot{x}^2 = f^2(x^2)$$

Proof: $D = \text{span}\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_k}\}$.

$$[f, \frac{\partial}{\partial x_i}] = -\sum_{j=1}^{n} \frac{\partial f_j}{\partial x_i} \frac{\partial}{\partial x_j} \in D \Rightarrow$$
$$\frac{\partial f_j}{\partial x_i} = 0, \;\; i = 1, \ldots, k \;\; j = k+1, \ldots, n$$

Invariant Distributions II

A distribution D is invariant for the system $\dot{x} = f(x) + g(x)u$ if

$$[f, D] \subset D$$
$$[g_j, D] \subset D$$

Invariant Codistributions

A codistribution P is invariant for the system $\dot{x} = f(x) + g(x)u$ if

$$L_f P \subset P$$
$$L_{g_j} P \subset P$$

If D and P are constant dimensional then

ker P invariant $\Leftrightarrow$ P invariant

ann D invariant $\Leftrightarrow$ D invariant

Proof: Follows from (see book)

$$L_f(\sigma(X)) = (L_f\sigma)(X) + \sigma([f, X])$$

Linear Case

$$W_c = \text{Im}\begin{pmatrix} B & AB & \ldots & A^{n-1}B \end{pmatrix}$$

W_c is the smallest A-invariant subspace containing B.

$$W_o = \text{Ker}\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

W_o is the largest A-invariant subspace contained in ker C

Nonlinear Invariance

Proposition

- C_0 is the smallest $\{f, g\}$-invariant distribution containing $\{g_1, \ldots, g_m\}$.
- ker dO is the largest $\{f, g\}$-invariant distribution contained in ker dh.

Proof: Direct from definitions, see book.

Nonlinear Decomposition

Theorem

a) Choose Frobenius-coordinates so $C_0 = \text{span}\{\frac{\partial}{\partial x^1}\}$, then

$$\begin{aligned} \dot{x}^1 &= f^1(x^1, x^2) + g(x^1, x^2)u \\ \dot{x}^2 &= f^2(x^2) \end{aligned}$$

The system in x^1 is LSA.

b) Choose Frobenius-coordinates so ker $d\mathcal{O} = \text{span}\{\frac{\partial}{\partial x^2}\}$, then

$$\begin{aligned} \dot{x}^1 &= f^1(x^1) + g(x^1)u \\ \dot{x}^2 &= f^2(x^1, x^2) + g^2(x^1, x^2)u \\ y &= h(x^1) \end{aligned}$$

The system in x^1 is locally observable.

Kalman Decomposition

$$\begin{aligned} X^2 &:= X_{\bar{o}} \cap X_s \\ X_s &= X^1 \oplus X^2 \quad \text{controllable} \\ X_{\bar{o}} &= X^2 \oplus X^4 \quad \text{nonobservable} \\ X &= X^1 \oplus X^2 \oplus X^3 \oplus X^4 \end{aligned}$$

$$A = \begin{pmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{pmatrix} \quad B = \begin{pmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{pmatrix}$$

$$C = \begin{pmatrix} C_1 & 0 & C_3 & 0 \end{pmatrix}$$

Generalized Frobenius' Theorem

Theorem If D_a, D_b and $D_a + D_b$ are constant dimensional involutive distributions then one can change coordinates to $x = (x^1, x^2, x^3, x^4)$ so that

$$\begin{aligned} D_a &= \text{span}\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\} \\ D_b &= \text{span}\{\frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^4}\} \end{aligned}$$

Proof: Redo the last part of the old proof of Frobenius theorem more carefully. See book.

Nonlinear Kalman Decomposition

Assume C_0, ker $d\mathcal{O}$ and C_0 + ker $d\mathcal{O}$ all have constant dimension. Then we can find coordinates $x = (x^1, x^2, x^3, x^4)$ so that

$$\begin{aligned} \dot{x}^1 &= f^1(x^1, x^3) + g^1(x^1, x^3)u \\ \dot{x}^2 &= f^2(x^1, x^2, x^3, x^4) + g^2(x^1, x^2, x^3, x^4)u \\ \dot{x}^3 &= f^3(x^3) \\ \dot{x}^4 &= f^4(x^3, x^4) \\ y &= h(x^1, x^3) \end{aligned}$$

Here

$$\begin{aligned} C_0 &= \text{span}\ \{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}\} \\ \ker\ d\mathcal{O} &= \text{span}\ \{\frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^4}\} \end{aligned}$$

Proof: Follows from the generalized Frobenius' theorem.

Next Week

- State Space Transformations
- Feedback
- Normal Form
- Exact Linearization I
- Zero Dynamics I

Chapter 5 + Sastry's notes

Nonlinear Control Theory 94

Lecture 5

- Feedback, Definitions
- State Space Transformations (SST)
- Linearization by SST
- Canonical Forms
  - Observability Form
  - Controller/Normal Form
- Exact Linearization
- Zero Dynamics

Nij. pp. 148-175 + Sastry

Last Week

1-forms: $\sigma = \sum \sigma_i(x) dx_i$.

Exact 1-form: $df = \sum \frac{\partial f}{\partial x_i} dx_i$

Co-distribution: $P = \text{span}\{\sigma_1, \ldots, \sigma_l\}$.

$\ker P = \{X \ \mid <\sigma, X> = 0, \ \ \forall \sigma \in P\}$

Frobenius:

$$\ker\ P \text{ involutive} \quad \Longleftrightarrow \quad P = \text{span}\{dx_{k+1}, \ldots, dx_n\}$$

Observation Space

$$\mathcal{O} = \text{span}\{L_{X_1} L_{X_2} \ldots L_{X_k} h_j\}; \ \ X_i \in \{f, g_1, \ldots, g_m\}$$

Local Observability $\Longleftrightarrow$ dim $d\mathcal{O} = n$.

Nonlinear Kalman Decomposition

State Feedback

Strict Static State Feedback

$$u = \alpha(x)$$

Regular Static State Feedback

$$u = \alpha(x) + \beta(x)v$$

$\beta(x)$ non-singular. New input v.

Need full state x

Example Computed Torque in Robotics

$$M(\theta)\ddot{\theta} + C(\theta, \dot{\theta}) + K(\theta) = u$$
$$u = M(\theta)v + C(\theta, \dot{\theta}) + K(\theta)$$
$$\Longrightarrow \quad \ddot{\theta} = v$$

Output Feedback

Strict Static Output Feedback

$$u = \alpha(y)$$

Regular Static Output Feedback

$$u = \alpha(y) + \beta(y)v$$

Dynamic Feedback

Dynamic State Feedback

$$\dot{z} = \gamma(z,x) + \delta(z,x)v$$
$$u = \alpha(z,x) + \beta(z,x)v$$

z are controller states and v new input.

Dynamic Output Feedback

$$\dot{z} = \gamma(z,y) + \delta(z,y)v$$
$$u = \alpha(z,y) + \beta(z,y)v$$

Can also study more general feedback structures, e.g. $u = \alpha(x,v)$.

Another Useful Idea:

$$\dot{x} = f(x,u)$$
$$\dot{u} = v$$

States x, u, new control v. Affine system.

Transformations, Invariants

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \tag{1}$$

When is (1) equivalent to a linear system?

$$(\tilde{x}, \tilde{u}, \tilde{y}) = S(x,u,y)$$

Invariants under transformations?

(SST): $z = S(x)$ state space trans.

(IT): $u = \alpha(x) + \beta(x)v$ input trans.

State Space Transformations

$$\begin{aligned} z &= S(x); \quad x = S^{-1}(z) \\ \dot{z} &= \frac{\partial S}{\partial x}(f(x) + g(x)u) \\ &= \frac{\partial S}{\partial x}(S^{-1}(z))\Big(f(S^{-1}(z)) + g(S^{-1}(z))\Big)u \\ &= S_* f + S_* g u \end{aligned}$$

So

$$\begin{aligned} f &\mapsto S_* f \\ g_i &\mapsto S_* g_i \\ h &\mapsto h \circ S^{-1} \\ x(t,0,x_0,u) &\mapsto z(t,0,S(x_0),u) \end{aligned}$$

Example

Transform

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

using $z = S(x)$ given by

$$z_1 = e^{2x_1} - 1$$
$$z_2 = e^{x_2} - 1$$

Solution:

$$S_* \sim \frac{\partial S}{\partial x} = \begin{pmatrix} 2e^{2x_1} & 0 \\ 0 & e^{x_2} \end{pmatrix}$$

$$S_* f \sim \begin{pmatrix} 2e^{2x_1} & 0 \\ 0 & e^{x_2} \end{pmatrix} \begin{pmatrix} -e^{x_2 - 2x_1} \\ e^{2x_1 - x_2} \end{pmatrix} = \begin{pmatrix} -2z_2 - 2 \\ z_1 + 1 \end{pmatrix}$$

$$S_* g \sim \begin{pmatrix} 2e^{2x_1} & 0 \\ 0 & e^{x_2} \end{pmatrix} \begin{pmatrix} e^{-2x_1} \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$\dot{z} = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} z + \begin{pmatrix} 2 \\ 0 \end{pmatrix} u + \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

The Linear Subtree

SST Linearization of (f, g)

Theorem 5.3'

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

can be transformed using $z = S(x)$ locally around x_0 to a controllable linear system

$$\dot{z} = Az + Bu + v?$$

if and only if

(i) $\dim(\mathrm{span}\{ad_f^k g_j(x_0), k \leq n-1\}) = n$

(ii) $[g_i, \mathrm{ad}_f^k g_j](x) = 0, \ \ \forall i, j, k$ around x_0.

Remark (i) $\Longleftrightarrow$ linearization is controllable.

Remark (ii) Everything in C_0 except linear subtree is zero.

Remark $v = 0$ if $f(x_0) = 0$.

Remark By the Jacobi identity:

$$[\mathrm{ad}_f^l g, \mathrm{ad}_f^k g] = -[\mathrm{ad}_f^{l-1} g, \mathrm{ad}_f^{k+1} g] - [[\mathrm{ad}_f^{l-1} g, \mathrm{ad}_f^k g], f]$$

one can see that (ii) is equivalent to

(ii') $[\mathrm{ad}_f^l g_i, \mathrm{ad}_f^k g_j](x) = 0, \ \forall i, j, k, l$ around x_0.

which is the condition in Nij.

Proof of Theorem 5.3'

• Necessity of (i) and (ii): They are satisfied for a linear system. The conditions are invariant under S, for instance:

$$S_*[g_j, [f, g_i]] = [S_* g_j, [S_* f, S_* g_i]] = [B, [Az + v, B]] = 0$$

• Sufficiency of (i) and (ii): Choose linearly independent vector fields $X_1, \ldots, X_n$ from (i) so that by (ii'):

$$[X_i, X_j] = 0 \qquad \forall i, j$$

Choose Frobenius-coordinates so that

$$S_* X_i = \frac{\partial}{\partial z_i}$$

$$[\frac{\partial}{\partial z_i}, S_* g_j] = S_*[X_i, g_j] = 0 \ \Rightarrow \ S_* g_j = b_j$$

$$[\frac{\partial}{\partial z_j}, [\frac{\partial}{\partial z_i}, S_* f]] = S_*[X_j, [X_i, f]] = 0 \Rightarrow S_* f = Az + v$$

$$v = \frac{\partial S}{\partial x} f(x_0)$$

Example

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 + x_2x_3 - x_3^2 \\ -x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix} u$$

$$g(x), \mathrm{ad}_f g(x), \mathrm{ad}_f^2 g(x) = \begin{pmatrix} 4x_2x_3 \\ -2x_3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2x_2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

dim=3 for all x so (i) is satisfied.

Also easy to check that (ii) is satisfied:

$$[g, \mathrm{ad}_f g] = 0$$
$$[g, \mathrm{ad}_f^2 g] = 0$$

Example, Continued

Need to find S such that

$$S_* \begin{pmatrix} 4x_2x_3 & -2x_2 & 1 \\ -2x_3 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$S_* = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2x_3 \\ 1 & 2x_2 & 0 \end{pmatrix}$$

With $S(0) = 0$ we get

$$z = S(x) = \begin{pmatrix} x_3 \\ x_2 + x_3^2 \\ x_1 + x_2^2 \end{pmatrix}$$

This gives

$$\dot z_1 = u$$
$$\dot z_2 = z_1$$
$$\dot z_3 = z_2$$

Improved Condition (ii)

(ii) involves infinitely many conditions

However from (i)

$$\mathrm{ad}_f^l g_i(x) = \sum_j \sum_{k=0}^{n-1} \alpha_{l,k,j}(x) \mathrm{ad}_f^k g_j(x)$$

From (ii') it follows (after some work) that $\alpha_{l,k,j}(x)$ is independent of x.

Using this, (ii) can be changed to

$$[g_j, \mathrm{ad}_f^k g_i] = 0, \;\; 0 \le k \le 2n-2(?)$$

Details left as exercise.

In book: $k = 1, 3, \ldots, 2n-1$. Mistake ??

Canonical Forms, Linear Case

Observability Form (SISO)

$$z = Sx = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} x$$

S invertible gives

$$\dot z = \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & \ddots & \\ & & & 1 \\ -a_1 & -a_2 & \ldots & -a_n \end{pmatrix} z + \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} u$$

$$y = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} z$$

Nice observability matrix: $W_o = I$.

Nonlinear Observability Form

$$\begin{aligned}\dot{x} &= f(x) + g(x)u \\ y &= h(x)\end{aligned}$$

$$z = S(x) = \begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{pmatrix}$$

$S(x)$ full rank gives

$$\begin{aligned}\dot{z}_1 &= \frac{\partial h}{\partial x}\dot{x} = L_f h + L_g h u = z_2 + b_1(z)u \\ \dot{z}_2 &= z_3 + b_2(z)u \\ &\vdots \\ \dot{z}_n &= L_f^n h + b_n(z)u =: -\sum_{j=1}^{n} a_j(z) z_j + b_n(z)u\end{aligned}$$

$$\dot{z} = \begin{pmatrix} 0 & 1 & & \\ 0 & 0 & \ddots & \\ & & & 1 \\ -a_1(z) & -a_2(z) & \dots & -a_n(z) \end{pmatrix} z + \begin{pmatrix} b_1(z) \\ b_2(z) \\ \vdots \\ b_n(z) \end{pmatrix} u$$

$$y = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} z$$

SST Linearization of (f, h).

Theorem 5.13 Consider the nonlinear system

$$\begin{aligned}\dot{x} &= f(x) \\ y &= h(x)\end{aligned}$$

Assume $f(x_0) = 0$ and $h(x_0) = 0$. Then the system can be transformed using $z = S(x)$ locally around x_0 to

$$\begin{aligned}\dot{z} &= Az \\ y &= Cz\end{aligned}$$

if and only if

(i) Linearization is observable: $S(x_0)$ full rank.

(ii)

$$L_f^n h(x) = -\sum_{j=0}^{n-1} a_j L_f^j h(x)$$

where a_j are constants.

Proof of Theorem 5.13

Necessity of (i) and (ii): Follows from linear case.

Sufficiency of (i) and (ii): Immediate from the observability form.

SST Linearization of (f, g, h)

Theorem 5.9 The nonlinear system

$$\begin{aligned}\dot{x} &= f(x) + g(x)u \\ y &= h(x)\end{aligned}$$

with $f(x_0) = 0, h(x_0) = 0$ can be transformed using $z = S(x)$ locally around x_0 to

$$\begin{aligned}\dot{z} &= Az + Bu \\ y &= Cz\end{aligned}$$

if and only if

(i) Linearization is controllable

(ii) Linearization is observable

(iii) $L_g L_f^j dh(x) = 0, \quad j = 0, \dots, n-1, \quad \forall x$

Proof: Skip.

More Linear Canonical Forms

Controller/Normal Form

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

Introduce an auxiliary output $\tilde{y} = \phi(x)$. Remember that the *relative degree* γ satisfies

$$L_g L_f^i \phi(x) = 0 \quad \forall x, \quad i = 0, \dots, \gamma - 2$$
$$L_g L_f^{\gamma-1} \phi(x_0) \neq 0$$

Now use the same first coordinates as in the transformation to observability form

$$S_\gamma(x) = \begin{pmatrix} \phi(x) \\ L_f \phi(x) \\ \vdots \\ L_f^{\gamma-1}\phi(x) \end{pmatrix}$$

Controller Form

If $\gamma = n$ (full relative degree) for the system with output $\phi(x)$ then $z = S_\gamma(x)$ gives the nonlinear controller form:

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_n &= a(z) + b(z)u \\ y &= h(S_\gamma^{-1}(z)) \end{aligned}$$

Where $a(z) = L_f^\gamma \phi(x)$ and $b(z) = L_g L_f^{\gamma-1}\phi(x)$.

Exact Linearization (SISO)

Definition A system f, g is *exact linearizable* if there is an IT $u = \alpha(x) + \beta(x)v$ and a SST $z = S(x)$ so that the new system is linear

$$\dot{z} = Az + Bv$$

Also called *state space linearizable.*

Theorem S8.7 A system is exact linearizable if and only if one can find an output function $\phi(x)$ so that the system has relative degree n.

Proof:

Sufficiency: Put $u = b^{-1}(z)(-a(z) + v)$ in the controller form.

Necessity: Relative degree is invariant under SST and IT. A linear system in controller form with $\phi(x) = x_1$ has relative degree n.

Exact Linearization (SISO)

Given $f(x), g(x)$. When is there a $\phi(x)$ that gives relative degree n?

Theorem Nij. 6.17 A system is exact linearizable if and only if

(i) Linearization is controllable, i.e.

$$D_n := \text{span } \{g(x), \text{ad}_f g(x), \ldots, \text{ad} f^{n-1} g(x)\}$$

has dimension n around x_0.

(ii) The distribution

$$D_{n-1} := \text{span } \{g(x), \text{ad}_f g(x), \ldots, \text{ad} f^{n-2} g(x)\}$$

is *involutive*.

Proof of Nij. 6.17

Relative degree $= n$ if and only if

$$\frac{\partial \phi}{\partial x} \begin{pmatrix} g(x) & \text{ad}_f g(x) & \ldots & \text{ad}_f^{n-2} g(x) \end{pmatrix} = 0$$

This means that the one-dimensional co-distribution $\text{ann}(D_{n-1})$ is spanned by an exact one-form (given by $\sigma = d\phi$). Frobenius theorem then gives the result.

See Sastry for some more details.

Normal Form, Zero Dynamics

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned}$$

What if relative degree $< n$?

$$S_\gamma(x) = \begin{pmatrix} \phi(x) \\ L_f \phi(x) \\ \vdots \\ L_f^{\gamma-1} \phi(x) \end{pmatrix}$$

Complement with $\eta_1(x), \ldots, \eta_{n-\gamma}(x)$ such that $L_g \eta_i(x) = 0, \ \forall i$ (by Flow-box Theorem):

$$z = S(x) = \begin{pmatrix} \phi(x) \\ L_f \phi(x) \\ \vdots \\ L_f^{\gamma-1} \phi(x) \\ \eta_1(x) \\ \vdots \\ \eta_{n-\gamma}(x) \end{pmatrix}$$

Normal Form, Zero Dynamics

In the new coordinates

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_\gamma &= a(z, \eta) + b(z, \eta)u \\ \dot{\eta} &= q(z, \eta) \\ y &= z_1 \end{aligned}$$

The last dynamical equations with $z = 0$

$$\dot{\eta} = q(0, \eta) \tag{2}$$

are called the *zero dynamics* of (f, g, h).

The system is said to be *locally asymptotically minimum phase* at x_0 if (2) is locally asymptotically stable.

Example

Nonlinear Ball and Beam

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 x_4^2 - G \sin x_3 \\ x_4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u$$

$(x_1, x_2, x_3, x_4)^T := (x, \dot{x}, \theta, \dot{\theta})^T.$

$$D_4 = \begin{pmatrix} 0 & 0 & 2x_1x_4 & * \\ 0 & -2x_1x_4 & -2x_2x_4 - G\cos x_3 & * \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

has full rank since $\det D_4 = G^2 \neq 0$.

However

$$[g, \mathrm{ad}_f^2 g] = \begin{pmatrix} 2x_1 \\ -2x_2 \\ 0 \\ 0 \end{pmatrix}$$

does not lie in the span of the first 3 columns.

D_3 not involutive.

Not exact linearizable.

Elastic Robot Arm

$$\ddot{q}_1 + \dot{q}_1 + q_2 - q_1 = u$$
$$\ddot{q}_2 + \dot{q}_2 + q_2 - q_1 + \cos q_2 = 0$$

$x = (q_1, q_2, \dot{q}_1, \dot{q}_2)^T$ gives

$$f(x) = \begin{pmatrix} x_3 \\ x_4 \\ x_1 - x_2 - x_3 \\ x_1 - x_2 - \cos x_2 - x_4 \end{pmatrix} \qquad g(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Elastic Robot Arm

State Space Linearizable?

Using $y = x_2$ gives

$$\begin{aligned} &h = x_2 \\ &L_f h = x_4 \\ &L_f^2 h = f_4(x_1, x_2, x_4) \\ &L_g h = L_g L_f h = L_g L_f^2 h = 0 \\ &L_g L_f^3 h \neq 0 \end{aligned}$$

so relative degree is 4.

Linearizing coordinates: $q_2, \dot{q}_2, \ddot{q}_2, q_2^{(3)}$.

Linearizing feedback:

$$u = \frac{-L_f^4 h(x) + v}{L_g L_f^3 h(x)}$$

Nonlinear Control Theory 94

Lecture 6

6 Feedback Linearization (MIMO)
7 Disturbance Decoupling
8 Input/Output Decoupling I
9 Input/Output Decoupling II

Nij. pp. 176-298

NOTE: This is a very short summary of parts of the book that we will skip. No details. No exercises.

Last Week

Feedback, definitions

State Space Transformation $z = S(x)$

Linearization by SST, $[g, \text{ad}_f^k g] = 0$.

Observability Form $S(x) = [h, L_f h, \ldots, L_f^{n-1} h]$

Controller/Normal Form. $h(x) = \phi(x)$.

Linearization by SST and IT: $u = \alpha(x) + \beta(x)v$

SISO: Possible if relative degree $= n$ (for some $\phi(x)$)

$$\Leftrightarrow D_{n-1} = \text{span}\{g, \text{ad}_f g, \ldots, \text{ad}_f^{n-2} g\} \text{ involutive}$$

Zero Dynamics: $\dot{\eta} = q(0, \eta)$

6. Feedback Linearization (MIMO)

$$\begin{aligned}
\dot{x} &= f(x) + g(x)u \\
u &= \alpha(x) + \beta(x)v \\
z &= S(x) \\
\dot{z} &= Az + Bv \\
A &= S_*(f + gu) \\
B &= S_*(g\beta)
\end{aligned}$$

Introduce

$$D_k(x) = \text{span}\{\text{ad}_f^r g_1, \ldots, \text{ad}_f^r g_m\} \quad r \le k-1$$

Linear Subtrees

6. Feedback Linearization

Theorem 6.3 Assume f, g LSA in x_0 then

$$\begin{gathered}
\text{Feedback Linearizable to } \dot{z} = Az + bv + f(x_0) \\
\Longleftrightarrow \\
D_1, \ldots, D_n \text{ involutive}
\end{gathered}$$

Proof: Technical MIMO-fication of our previous SISO proof + super-duper Frobenius' theorem.

Remark The conditions are "generically" NOT satisfied. Transformation is not easy to find.

Remark From Sastry we know that in the single input case we only need to check

$$\dim D_n(x_0) = n \quad \text{and} \quad D_{n-1} \text{ involutive}$$

6. Feedback Linearization

$$\dot{x} = f(x,u); \qquad f(x_0,u_0) = 0$$

Definition Feedback Linearizable if

- $\exists u = \alpha(x,v)$ with $\alpha(x_0,0) = u_0$ and $\frac{\partial\alpha}{\partial v}$ nonsingular
- $\exists S(x)$ such that $S(x_0) = 0$.

$$S_* f(x,\alpha(x,v)) = AS(x) + Bv$$

Idea Extend to affine system

$$\begin{aligned} \dot{x} &= f(x,u) \\ \dot{u} &= w \end{aligned}$$

New state: (x,u). New input: w.

Theorem 6.12 Assume LSA in (x_0,u_0) then

$$\dot{x} = f(x,u) \text{ FL} \iff \text{Extended system FL}$$

6. Example Rocket

$$\begin{aligned} \dot{x}_1 &= x_3 \\ \dot{x}_2 &= x_4 \\ \dot{x}_3 &= -gR^2/x_1^2 + T/m \cos u + x_1 x_4^2 \\ \dot{x}_4 &= -2x_3x_4/x_1 + T/mx_1 \sin u \end{aligned}$$

Extended system:

$$f = \begin{pmatrix} x_3 \\ x_4 \\ -gR^2/x_1^2 + T/m\cos u + x_1x_4^2 \\ -2x_3x_4/x_1 + T/mx_1 \sin u \\ 0 \end{pmatrix} \quad g = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Easy to see that $D_2 = \text{span}\{g, [f,g]\}$ not involutive since $[g,[f,g]] \notin D_2$

So the rocket is NOT feedback linearizable

7. Disturbance Decoupling

$$\begin{aligned} \dot{x} &= f(x) + g(x)u + e(x)d \\ y &= h(x) \end{aligned}$$

y is decoupled from d if d does not influence y (for any u.)

Disturbance Decoupling Problem When is there $u = \alpha(x) + \beta(x)v$ such that y becomes decoupled from d?

7. Invariant Distributions Revisited

D invariant under f, g if

$$\begin{aligned} [f,D] &\subset D \\ [g_i,D] &\subset D \end{aligned}$$

Prop 4.23 y is decoupled from d if there exists D such that

$$\begin{aligned} [f,D] &\subset D \\ [g_i,D] &\subset D \quad \forall i \\ e_j &\in D \quad \forall i \\ D &\subset \ker dh \end{aligned}$$

7. Locally Controlled Invariant D

Controlled invariant if $\exists \alpha, \beta$ such that $u = \alpha + \beta v$ makes $\tilde{f} = f + g\alpha$, $\tilde{g} = g\beta$ invariant.

Example Linear system: V controlled invariant, or (A, B)-invariant, if

$$\exists F : (A + BF)V \subset V \quad \Leftrightarrow \quad AV \subset V + \mathrm{Im} B$$

Theorem 7.5 D is locally controlled invariant if $G, D, D \cap G$ are constant dimensional and D involutive and

$$[f, D] \subset D + G$$
$$[g_i, D] \subset D + G \quad \forall i \in \underline{m}$$

7. Linear DDP

$$\dot{x} = Ax + Bu + Ed$$
$$y = Cx$$

$V^* = \max A, B$-invariant subspace in ker C

$$AV^* \subset V^* + B$$
$$V^* \subset \ker C$$

Theorem DDP solvable if and only if

$$\mathrm{Im}\ E \subset V^*$$

7. Nonlinear DDP

Maximally Controlled Invariant Involutive Distribution D^* in ker dh:

$$D^*(f; g; \ker\ dh)$$

Theorem 7.14 Assume G, $D^*, D^* \cap G$ constant dimensional. Then nonlinear DDP solvable if and only if

$$\mathrm{span}(e_1, \ldots, e_l) \subset D^*$$

D^* is closely related to the zero dynamics.

7. Algorithm for D^*

In the linear case:

$$V^0 := R^n$$
$$V^{\mu+1} = \ker\ C \cap A^{-1}(V^\mu + \mathrm{Im}\ B)$$

In the nonlinear case:

$$D^0 := TM$$
$$D^{\mu+1} := \ker\ dh \cap \{X : [f, X] \in D^\mu + G, [g_i, X] \in D^\mu + G\}$$

Prop. 7.16 Under some assumptions on constant dimensions

- $D^0 \supset D^1 \supset \ldots D^n = D^*$
- D^μ is involutive

Note that the algorithm converges in at most n steps.

7. Dual Algorithm

$$
\begin{aligned}
P^0 &:= 0 \\
P^1 &:= \mathrm{span}\{dh_1, \ldots, dh_p\} \\
P^{\mu+1} &:= P^\mu + L_f(P^\mu \cap \mathrm{ann}\ G) + \sum L_{g_i}(P^\mu \cap \mathrm{ann}\ G)
\end{aligned}
$$

$$D^\mu = \ker\ P^\mu$$

$$D^* = \ker\ P^*$$

7. Explicit Formula for D^* in SISO

Theorem 7.21 Let r be relative degree

$$L_g L_f^{r-1} h \neq 0$$

then

$$
\begin{aligned}
D^* &= \ker\ (\mathrm{span}\ \{dh, \ldots, dL_f^{r-1}h\}) \\
P^* &= \mathrm{span}\ \{dh, \ldots, dL_f^{r-1}h\}
\end{aligned}
$$

Compare Normal Form and Zero Dynamics

7. Modified DDP

Feedforward: d also measurable.

Theorem 7.24 The MDDP = Feedforward/Feedback problem is solvable iff

$$\mathrm{span}\ (e_1, \ldots, e_l) \subset D^* + G$$

Remark The system $\dot{x} = \tilde{f}(x)$ is not necessarily stable. Stabilization can be treated as a separate problem.

Output Dynamic DDP? More realistic.

8. Input/Output Decoupling I

$$
\begin{aligned}
\dot{x} &= f(x) + g(x)u \\
y &= h(x)
\end{aligned}
$$

Definition 8.1 *I/O-decoupled* if, after possible relabeling of inputs,

(i) $j \neq i:\quad y_i$ invariant under u_j.
(ii) y_i not invariant under u_i.

Prop. 4.14 (i) is equivalent to

$$j \neq i : L_{g_j} L_{X_1} \ldots L_{X_k} h_i = 0, \qquad X_k \in \{f, g_1, \ldots, g_m\}$$

8. Static State Feedback

When does there exist $u = \alpha(x) + \beta(x)v$ such that the new system is IOD?

Definition The *Characteristic Numbers* $\rho_1, \dots, \rho_p$ of the outputs are the smallest integers such that

$$\begin{aligned} &L_g L_f^k h_j(x) = 0 \;\; k = 0, \dots, \rho_j - 1 \\ &L_g L_f^{\rho_j} h_j(x) \neq 0 \end{aligned}$$

$\approx$ Relative Orders ($r_j = \rho_j + 1$.)

8. I/O-Decoupling Theorem

Theorem 8.9 The IOD-problem is solvable if and only if

$$A(x) = \begin{pmatrix} L_{g_1} L_f^{\rho_1} h_1(x) & \dots & L_{g_m} L_f^{\rho_1} h_1(x) \\ & \vdots & \\ L_{g_1} L_f^{\rho_m} h_m(x) & \dots & L_{g_m} L_f^{\rho_m} h_m(x) \end{pmatrix}$$

has full rank m.

In fact

$$\begin{pmatrix} y_1^{(\rho_1+1)} \\ \vdots \\ y_m^{(\rho_m+1)} \end{pmatrix} = \begin{pmatrix} L_f^{\rho_1+1} h_1(x) \\ \vdots \\ L_f^{\rho_m+1} h_m(x) \end{pmatrix} + A(x)u$$

8. Static IOD

So the static state feedback

$$u = -(A(x))^{-1} \begin{pmatrix} L_f^{\rho_1+1} h_1(x) \\ \vdots \\ L_f^{\rho_m+1} h_m(x) \end{pmatrix} + (A(x))^{-1} v$$

gives a new system:

$$\begin{pmatrix} y_1^{(\rho_1+1)} \\ \vdots \\ y_m^{(\rho_m+1)} \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

Automatically "I/O-linearized".

Not state space linearized, since there might be nonlinear zero dynamics.

8. Normal Form (MIMO)

If we define

$$z^i = \begin{pmatrix} h_i(x) \\ L_f h_i(x) \\ \vdots \\ L_f^{\rho_i} h_i(x) \end{pmatrix}$$

and choose $\bar{z}$ so that $(\bar{z}, z^1, \dots, z^m) = S(x)$ forms a local coordinate system, then

$$\begin{aligned} \dot{z}^i &= A_i z^i + b_i v, \quad i \in \underline{m} \\ \dot{\bar{z}} &= \bar{f}(\bar{z}, z^1, \dots, z^m) + \bar{g}(\bar{z}, z^1, \dots, z^m) v \\ y_i &= z_{i1} \end{aligned}$$

where A_i, b_i are in Brunovsky canonical form.

Still nonlinear, except when

$$\sum_{i=1}^{m} (\rho_i + 1) = n$$

(full relative degree)

8. Dynamic State Feedback IOD

What if $A(x)$ is not invertible?

Dynamic State Feedback:

$$\begin{aligned}\dot{z} &= \gamma(z,x) + \delta(z,x)v\\ u &= \alpha(z,x) + \beta(z,x)v\end{aligned}$$

Still requires full state x.

Problem 8.16 Consider a square system. When is there a dynamic state feedback such that the modified dynamics are IOD?

Linear Case If and only if det $G(s) \not\equiv 0$.

8. Dynamic State Feedback IOD

Nonlinear Case?

Dynamic Extension Algorithm If $A(x)$ is not invertible then add integrators at the inputs (in a systematic way) so that the new system gets invertible $\tilde{A}(x)$.

Treat the integrators as part of the controller.

Theorem 8.19 Dynamic State Feedback IOD if and only if the "rank" of the nonlinear system is full:

$$q^* = n.$$

For details see book.

Still nonlinear unless full relative degree.

Robotics Example (Ola Dahl)

Read Ola's Report

9. I/O Decoupling II

Geometric Approach

$$\begin{aligned}\dot{x} &= f(x) + g(x)u\\ y_1 &= h_1(x)\\ &\vdots\\ y_p &= h_p(x)\end{aligned}$$

where now $h_i \in R^{p_i}$ (vector valued).

Not only square systems

Assume LSA, MIMO

Necessary conditions for block IOD:

$$\begin{gathered}L_{g_j} L_{X_1} \ldots L_{X_k} h_{il}(x) = 0, \quad l \in \underline{p_i}, j \neq i\\ X_i \in \{f, g_1, \ldots, g_m\},\end{gathered}$$

"No interaction from input j to output block i." (then automatically by LSA: input i does influence output block i, so IOD-problem is solved).

9. Static IOD II

When is there $u = \alpha(x) + \beta(x)v$ so that IOD?

Linear Case: $u = Fx + Gv$. Answer: Iff

$$\text{Im } B = \text{Im } B \cap V_1^* + \dots \text{Im } B \cap V_m^*$$

where V_i^* is the maximal controlled invariant subspace contained in $\cap_{j \neq i} \ker C_j$.

9. Nonlinear Static IOD II

Introduce

$$D_i^* = D^*(f; g; \cap_{j \neq i} \ker dh_j)$$

Theorem 9.7 Under some technical assumptions the static IOD problem is solvable if and only if

$$G = D_1^* \cap G + \dots D_m^* \cap G$$

For square systems this can be shown to be equivalent to the previous condition on $A(x)$.

Nonlinear Control Theory 94

Lecture ~~8~~ 7

10.3 Center Manifold Theory
11 More on Zero Dynamics

Nij. pp. 310-317 + 323-338

What Have We Covered?

1 Introduction
2 Mathematics
3 Controllability; Observability
5 State Space Transformations
6 Feedback Linearization
7 Disturbance Decoupling
8 I/O Decoupling I
9 I/O Decoupling II

Future

Volterra Systems (Ch 4): Sven Spanne

Mechanical Systems (Ch 12): Rolf J

Will skip Ch 13-14

2nd half of the course: Khalil

Recaption; I/O Decoupling

When does there exist $u = \alpha(x) + \beta(x)v$ such that the new system is IOD, i.e. diagonal?

$p = m$

Definition The *Characteristic Numbers* $\rho_1, \ldots, \rho_p$ of the outputs are the smallest integers such that

$$L_g L_f^{\rho_j} h_j(x) \neq 0$$

Relative Orders $r_j = \rho_j + 1$.

I/O-Decoupling Theorem

Put

$$A(x) = \begin{pmatrix} L_{g_1} L_f^{\rho_1} h_1(x) & \dots & L_{g_m} L_f^{\rho_1} h_1(x) \\ \vdots & & \\ L_{g_1} L_f^{\rho_m} h_m(x) & \dots & L_{g_m} L_f^{\rho_m} h_m(x) \end{pmatrix}$$

$$B(x) = \begin{pmatrix} L_f^{\rho_1+1} h_1(x) \\ \vdots \\ L_f^{\rho_m+1} h_m(x) \end{pmatrix}$$

Theorem 8.9 The IOD-problem is solvable if and only if $A(x)$ is invertible.

Static IOD

$$\begin{pmatrix} y_1^{(\rho_1+1)} \\ \vdots \\ y_m^{(\rho_m+1)} \end{pmatrix} = B(x) + A(x)u$$

So the static state feedback

$$u = -(A(x))^{-1}(B(x) + v)$$

gives a new system:

$$\begin{pmatrix} y_1^{(\rho_1+1)} \\ \vdots \\ y_m^{(\rho_m+1)} \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$$

Automatically "I/O-linearized" $v \mapsto y$.

Not state space linearized $v \mapsto x$.

Not Stable !

Normal Form

If we define

$$z^i = \begin{pmatrix} h_i(x) \\ L_f h_i(x) \\ \vdots \\ L_f^{\rho_i} h_i(x) \end{pmatrix}$$

and choose $\bar{z}$ so that $(\bar{z}, z^1, \dots, z^m) = S(x)$ forms a local coordinate system, then

$$\begin{aligned} \dot{z}^i &= A_i z^i + b_i v, \quad i \in \underline{m} \\ \dot{\bar{z}} &= \bar{f}(\bar{z}, z^1, \dots, z^m) + \bar{g}(\bar{z}, z^1, \dots, z^m) v \\ y_i &= z_{i1} \end{aligned}$$

where A_i, b_i are in Brunovsky canonical form.

Zero Dynamics

In the case $p = m$ and $A(x)$ invertible the *Zero Dynamics* are given by

$$\dot{\bar{z}} = f(\bar{z}, 0, \dots, 0)$$

In Ch. 11.1-11.2 the zero dynamics are defined in a more general situation.

The system is called *asymptotically nonlinear minimum phase* at $z = 0$ if the zero dynamics is asymptotically stable.

Zero Dynamics for Linear SISO

$$G(s) = K\frac{s^{n-r} + \ldots + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_0}$$

$$\dot{x} = \begin{pmatrix} 0 & & I \\ -a_0 & \ldots & -a_{n-1} \end{pmatrix} x + \begin{pmatrix} 0 \\ K \end{pmatrix} u$$

$$y = \begin{pmatrix} b_0 & \ldots b_{n-r-1} & 1 & 0 & \ldots & 0 \end{pmatrix} x$$

Transform to normal form using

$$\bar{z} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-r} \end{pmatrix}$$

$$z = \begin{pmatrix} Cx \\ \vdots \\ CA^{r-1}x \end{pmatrix}$$

Zero Dynamics for Linear SISO

$$\dot{z}_1 = z_2$$

$$\vdots$$

$$\dot{z}_r = P\bar{z} + Qz + Ku$$

$$\dot{\bar{z}} = \begin{pmatrix} 0 & & I \\ -b_0 & \ldots & -b_{n-r-1} \end{pmatrix} \bar{z} + \begin{pmatrix} 0 & \ldots & 0 \\ 1 & \ldots & 0 \end{pmatrix} z$$

The zero dynamics are given by

$$\dot{\bar{z}} = R\bar{z}$$

where the eigenvalues of R=zeros of $G(s)$.

Stability Theorem

Theorem 11.16 Consider a square system with invertible $A(x)$. Assume that the zero dynamics are locally asymptotically stable around $\bar{z}_0$. Then there exist a *decoupling* feedback $u(x) = \alpha(x) + \beta(x)v$ such that the closed loop system for $v = 0$ is locally asymptotically *stable* around x_0.

To prove this we will use the *center manifold theorem.*

10.3 Center Manifold Theorem

Assume

$$\dot{z}^1 = A^0 z^1 + f^0(z^1, z^2)$$
$$\dot{z}^2 = A^- z^2 + f^-(z^1, z^2)$$

A^-: asymptotically stable

A^0: eigenvalues on imaginary axis

f^0 and f^- second order and higher terms.

Center Manifold Theorem Assume $z = 0$ is an equilibrium point. For every $k \geq 2$ there exists a C^k mapping φ such that $\phi(0) = 0$ and $d\phi(0) = 0$ and the surface

$$z^2 = \phi(z^1)$$

is invariant under the dynamics above.

Proof Idea: Construct a contraction with the center manifold as fixpoint.

Example

Example

Usage

1) Determine $z_2 = \phi(z_1)$, at least approximately

2) The local stability for the entire system can be proved to be the same as for the dynamics restricted to a center manifold:

$$\dot{z}_1 = A^0 z_1 + f^0(z_1, \phi(z_1))$$

Center Manifold Theorem

Example

$$\begin{aligned}\dot z^1 &= z_1 z_2 + z_1^3 + z_1 z_2^2\\ \dot z^2 &= -z_2 - 2z_1^2 + z_1^2 z_2\end{aligned}$$

Here $A^0 = 0$ and $A^- = -1$. $z_2 = \phi(z_1)$ gives

$$-\phi - 2z_1^2 + z_1^2\phi - \phi'[z_1\phi + z_1^3 + z_1\phi^2] = 0$$

hence

$$\phi(z_1) = -2z_1^2 + O(|z_1|^3)$$

Substituting into the dynamics we get

$$\dot z_1 = -2z_1^3 + z_1^3 + O(z_1^4)$$

so $z = (0,0)$ is asymptotically stable.

Nonuniqueness

The center manifold need not be unique

Example

$$\begin{aligned}\dot z^1 &= -z_1^3\\ \dot z^2 &= -z_2\end{aligned}$$

$z_2 = \phi(z_1)$ gives

$$\phi' z_1^3 = z_2 = \phi(z_1)$$

which has the solutions $\phi = Ce^{-1/z^2}$.

IOD with Stability

Theorem 11.16 $A(x)$ invertible + Zero Dynamics locally asymptotically stable

$$\Longrightarrow \quad u = (A(x))^{-1}(-B(x) + v)$$

gives a stable, I/O-decoupled system.

Proof First transform to normal form. Then stabilize A_i by additional feedback $v = \bar v + k_i z^i$. Setting $\bar v = 0$ gives

$$\begin{aligned}\dot z^1 &= \bar A_1 z^1\\ &\vdots\\ \dot z^m &= \bar A_m z^m\\ \dot{\bar z} &= \bar f(\bar z, z^1, \ldots, z^m) + \sum \bar g_j k_j z^j\end{aligned}$$

Proof continued

The eigenvalues of the linearization are the eigenvalues of $\bar A_1 \ldots, \bar A_m$ together with linearization of the zero dynamics. Rewrite the zero dynamics into (with $\bar z = (s^1, s^2)$):

$$\begin{aligned}\dot s^1 &= A^0 s^1 + f^0(s_1, s_2)\\ \dot s^2 &= A^- s^2 + f^-(s_1, s_2)\end{aligned}$$

From center manifold theorem there is an invariant submanifold $s^2 = \phi(s^1)$. This is also a center manifold for the entire system. The theorem follows from the stability of

$$\dot s^1 = A^0 s^1 + f^0(s^1, \phi(s^1))$$

Nonlinear Control Theory 94

Lecture ~~7~~ 8

Guest Lecture

Nonlinear Mechanical Control Systems

Nonlinear Control Theory 94

Lecture ~~8~~ 9

Guest Lecture, Sven Spanne

Volterra Systems

Nonlinear Control Theory 94

Lecture10

- Lyapunov Theorems
- Linearization and Center manifolds
- Regions of Attraction
- Nonautonomous systems

pp. 97-179 + 186-225

Autonomous systems

For $f : R^n \to R^n$ locally Lipschitz, consider

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

Stability of autonomous systems

The equilibrium point $x = 0$ is

stable if $\forall \epsilon > 0 : \exists \delta > 0$ such that

$$\|x(0)\| < \delta \Rightarrow \|x(t)\| < \epsilon \quad \forall t \geq 0$$

unstable if not stable

asymptotically stable if it is stable and there $\exists \delta > 0$ such that

$$\|x(0)\| < \delta \Rightarrow \lim_{t\to\infty} x(t) = 0$$

globally asymptotically stable if

$$\lim_{t\to\infty} x(t) = 0$$

independently of $x(0) \in R^n$.

Example: Pendulum

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -(g/l)\sin x_1, \quad |x_1| \leq \pi \end{aligned}$$

Stable equilibrium: $(0, 0)$

Unstable equilibria: $(\pm\pi, 0)$

No asymptotic stability

Lyapunov's stability theorem

Let $f(0) = 0$ and let $V : D \to R$ be a continuously differentiable function on a neighbourhood D of $x = 0$, such that

$$\begin{aligned} V(0) &= 0 \\ V(x) &> 0 \text{ for } x \neq 0 \\ \dot{V} &\leq 0 \end{aligned}$$

then $x = 0$ is stable and V is called a Lyapunov function for (1). Furthermore, asymptotic stability holds if in addition

$$\dot{V} < 0 \text{ for } x \neq 0$$

and global asymptotic stability holds if also

$$V(x) \to \infty \text{ as } ||x|| \to \infty$$

Level curves for Lyapunov functions

LaSalle's stability theorem

Let D be a compact set, which is invariant for (1). Let $V : D \to R$ be a continuously differentiable function on D such that

$$\dot{V} \leq 0 \text{ in } D$$

Let M be the largest invariant set in D such that

$$\dot{V} = 0 \text{ in } M.$$

Then every solution starting in D approaches M as $t \to \infty$.

Proof

Let $\phi(t, x_0) = x(t)$, where

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

Introduce

$$\begin{aligned} L^+(x_0) &= \cap_{T>0}\overline{\{\phi(t, x_0) : t \geq T\}} \\ L &= \cup_{x_0 \in D} L^+(x_0) \end{aligned}$$

For any x_0, $L^+(x_0)$ is invariant because for $y \in L^+(x_0)$ there is a sequence $\{t_i\}$ such that

$$\begin{aligned} y &= \lim_{i\to\infty} \phi(t_i, x_0) \\ \phi(t, y) &= \lim_{i\to\infty} \phi(t, \phi(t_i, x_0)) \\ &= \lim_{i\to\infty} \phi(t + t_i, x_0) \in L^+(x_0) \end{aligned}$$

and furthermore $\dot{V} = 0$ in $L^+(x_0)$.

Hence L is invariant, $\dot{V} = 0$ in L and every solution starting in D approaches L as $t \to \infty$.

Chetaev's instability theorem

Let $f(0) = 0$ and let $V : D \to R$ be a continuously differentiable function on a neighbourhood D of $x = 0$, such that $V(0) = 0$. Suppose that the set

$$U = \{x \in D : \|x\| < r, V(x) > 0\}$$

is nonempty for every $r > 0$. If $\dot{V} > 0$ in U, then $x = 0$ is unstable.

Example: Pendulum Revisited

The energy function

$$V(x) = (g/l)(1 - \cos x_1) + (1/2)x_2^2$$
$$V(x) \geq 0 \text{ with equality iff } x = (0,0)$$

Without friction:

$$\dot{V}(t) = 0 \quad \Rightarrow \quad x = (0,0) \text{ stable by Lyapunov}$$

Pendelum with friction

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -(g/l)\sin x_1 - (k/m)x_2$$

$\dot{V}(t) = -(k/m)x_2^2 \leq 0$

$D = \{x : V(x) \leq (2 - \epsilon)g/l\}$ for some small $\epsilon > 0$

$M = \{(0,0)\}$

$(x_1, x_2) \to M$ by LaSalle's theorem

Region of Attraction by LaSalle

If

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -4(x_1 + x_2) - \sin(2\pi(x_1 + x_2))$$

then by LaSalle's theorem, the region of attraction contains the set

$$\Omega = \{(x_1, x_2) : (x_1 + x_2)^2 \leq 1, \ (x_1 + x_2)^2 + x_1^2 \leq 10\}$$

Region of Attraction by Computer

For the system

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

We compute the following estimates by computer iteration.

Stability Analysis by Linearization

Suppose $f(0) = 0$. The equilibrium $x = 0$ of

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

is asymptotically stable if all eigenvalues of $\partial f / \partial x(0)$ have negative real part.

It is unstable if at least one eigenvalue has positive real part

Proof. Apply Lyapunov's theorem and Chetaev's theorem using quadratic Lyapunov functions from the linearization.

Center Manifold Theory

Assume

$$\begin{aligned}\dot{x}_1 &= A^0 x_1 + f^0(x_1, x_2) \\ \dot{x}_2 &= A^- x_2 + f^-(x_1, x_2)\end{aligned}$$

A^-: asymptotically stable

A^0: eigenvalues on imaginary axis

f^0 and f^- second order and higher terms.

Center Manifold Theorem

Assume $x = 0$ is an equilibrium point. For every $k \geq 2$ there exists a $\delta_k > 0$ and C^k mapping ϕ such that $\phi(0) = 0$ and $\phi'(0) = 0$ and the surface

$$x_2 = \phi(x_1) \quad \|x_1\| \leq \delta_k$$

is invariant under the dynamics above.

Proof Outline

For any continuously differentiable function ϕ_k, globally bounded together with its first partial derivative and with $\phi_k(0) = 0$, $\phi'(0) = 0$, let ϕ_{k+1} be defined by the equations

$$\dot{x}_1 = A^0 x_1 + f^0(x_1, \phi_k(x_1))$$
$$\dot{x}_2 = A^- x_2 + f^-(x_1, \phi_k(x_1))$$

$$\phi_{k+1}(x_1) = x_2$$

Under suitable assumptions, it can be verified that this defines ϕ_{k+1} uniquely, satisfying the assumptions for ϕ_k. Furthermore, the sequence $\{\phi_i\}$ is contractive in the norm $\sup_{x_1} \phi_i(x_1)$ and the limit ϕ satifies the conditions for a center manifold.

Usage

1) Determine $z_2 = \phi(z_1)$, at least approximately

2) The local stability for the entire system can be proved to be the same as for the dynamics restricted to a center manifold:

$$\dot{z}_1 = A^0 z_1 + f^0(z_1, \phi(z_1))$$

Example

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -x_2 + a x_1^2 + b x_1 x_2$$

Here $A^0 = 0$ and $A^- = -1$. $x_2 = \phi(x_1)$ gives

$$-\phi + a x_1^2 + b x_1 \phi - \phi' \phi = 0$$

hence

$$\phi(x_1) = a x_1^2 + O(|x_1|^3)$$

Substituting into the dynamics we get

$$\dot{x}_1 = a x_1^2 + O(|x_1|^3)$$

so $x = (0, 0)$ is unstable for $a \neq 0$.

Nonuniqueness

The center manifold need not be unique

Example

$$\dot{x}_1 = -x_1^3$$
$$\dot{x}_2 = -x_2$$

$x_2 = \phi(x_1)$ gives

$$\phi' x_1^3 = x_2 = \phi(x_1)$$

which has the solutions $\phi = Ce^{-1/x_2}$.

Nonautonomous systems

Let

$$\dot{x} = f(t, x), \tag{2}$$

where f is piecewise in $t \in [0, \infty)$ and locally Lipschitz in $x \in R^n$.

Definition of Uniform Stability

The equilibrium point $x = 0$ is

uniformly stable if $\forall \epsilon > 0 : \exists \delta > 0$ such that

$$\|x(t_0)\| < \delta \Rightarrow \|x(t)\| < \epsilon \quad \forall t \geq t_0 \geq 0$$

uniformly asymptotically stable if it is uniformly stable and $\exists \delta > 0$ such that

$$\|x(t_0)\| < \delta \Rightarrow \lim_{t \to \infty} x(t) = 0 \text{ uniformly in } t_0$$

exponentially stable if $\exists c, k, \gamma$ such that for $t \geq t_0 \geq 0$, $\|x(t_0)\| \leq c$ one has

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

Nonautonomous Stability Theorem

Let $f(\cdot, 0) = 0$, let $\alpha_1, \alpha_2, \alpha_3$ be strictly increasing functions on $[0, \infty)$ with $\alpha_1(0) = \alpha_2(0) = \alpha_3(0) = 0$ and let $V : R^2 \to R$ be a continuously differentiable function, such that

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

for $t \geq 0$, $\|x\| \leq r$. Then $x = 0$ is uniformly asymptotically stable.

If the conditions hold with

$$\alpha_i(r) = k_i r^c, \quad k_i > 0, c > 0, i = 1, 2, 3$$

then $x = 0$ is exponentially stable.

Proof

The second part is proved as follows. The first part is analogous, but less concrete.

$$\begin{aligned} \dot{V} &= \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t, x) \\ &\leq -k_3\|x\|^c \\ &\leq -\frac{k_3}{k_2} V \end{aligned}$$

$$\begin{aligned} V(t, x) &\leq V(t_0, x_0)e^{-(k_3/k_2)(t-t_0)} \\ &\leq k_2\|x_0\|^c e^{-(k_3/k_2)(t-t_0)} \end{aligned}$$

$$\begin{aligned} \|x(t)\| &\leq \left(\frac{V}{k_1}\right)^{1/c} \\ &\leq \left(\frac{k_2}{k_1}\right)^{1/c} \|x_0\|e^{-(k_3/k_2)(t-t_0)/c} \end{aligned}$$

Linear Time-varying Systems

A linear time-varying system is uniformly asymptotically stable if and only if it is globally exponentially stable.

Proof

Uniform asymptotic stability means that for any $x \in R^n$ there exists a function $\rho_x : R \to R$ such that

$$
\begin{aligned}
\rho_x(h) \searrow 0, \quad h \to \infty \\
\|\Phi(t,t_0)x\| \leq \rho_x(t-t_0), \quad \forall t \geq t_0
\end{aligned}
$$

Let $\rho(h) = \sum_{k=1}^{n} \rho_{e_k}(h)$ and choose $T > 0$ with $\rho(T) < 1/e$ for $h > T$. For $t_0 < t$, let N be the integer part of $(t - t_0)/T$. Then

$$
\|\Phi(t,t_0)\| = \max_{\|x\|=1} \|\Phi(t,t_0)x\| \leq \rho(t-t_0)
$$

$$
\begin{aligned}
&\|\Phi(t,t_0)\| \\
&= \|\Phi(t,t_0+NT)\cdots\Phi(t_0+2T,t_0+T)\Phi(t_0+T,t_0)\| \\
&\leq \rho(0)(1/e)^{N-1} \\
&\leq \rho(0)e^{-(t-t_0)/T}
\end{aligned}
$$

Input-Output Stability

Consider the dynamical system

$$
\begin{aligned}
\dot{x}(t) &= f(t,x,u) \\
y(t) &= h(t,x,u)
\end{aligned}
$$

where f is continuously differentiable and h is continuous. Let $x = 0$ be a globally exponentially stable equilibrium of

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

If the Jacobians $\frac{\partial f}{\partial x}(\cdot,\cdot,0)$ and $\frac{\partial f}{\partial u}$ are globally bounded and $\|h(t,x,u)\| \leq k_1\|x\| + k_2\|u\| + k_3$ for some $k_1, k_2, k_3 > 0$, then for $\|x(0)\| \leq \eta$, there exist $\gamma > 0, \beta \geq 0$ such that

$$
\sup_{t\geq 0} \|y(t)\| \leq \gamma \sup_{t \geq 0} \|u(t)\| + \beta.
$$

If $\eta = k_3 = 0$, then $\beta = 0$.

Proof outline

View $f(t,x,u)$ as perturbation of $f(t,x,0)$.

Global exponential stability plus bounded Jacobian $\frac{\partial f}{\partial x}(\cdot,\cdot,0)$ gives uniformly bounded Lyapunov function for $\dot{x} = f(t,x,0)$.

The global bound on $\frac{\partial f}{\partial u}$ gives a bound on the perturbation caused by the input.

Bounds on Lyapunov function and the bound on the perturbation together give the desired inequality.

Next week

Absolute Stability

Circle Criterion

Popov Criterion

Kalman - Yakubovich - Popov Lemma

Robustness Analysis

Nonlinear Control Theory 94

Lecture 11

- Absolute Stability
- Kalman - Yakubovich - Popov Lemma
- Circle Criterion
- Popov Criterion
- Simultaneous Lyapunov functions

pp. 237 - 268 + extra material on the K-Y-P Lemma

Global Sector Condition

Let $\psi(t,y) \in R$ be piecewise continuous in $t \in [0,\infty)$ and locally Lipschitz in $y \in R$.

Assume that ψ satisfies the *global sector condition*

$$\alpha \le \psi(t,y)/y \le \beta, \quad \forall t \ge 0, y \ne 0 \tag{1}$$

Absolute Stability

The system

$$\left\{ \begin{array}{rcl} \dot{x} & = & Ax + Bu, \quad t \ge 0 \\ y & = & Cx \\ u & = & -\psi(t,y) \end{array} \right. \tag{2}$$

with sector condition (1) is called *absolutely stable* if the origin is globally uniformly asymptotically stable for any nonlinearity ψ satisfying (1).

The Circle Criterion

The system (2) with sector condition (1) is absolutely stable if the origin is asymptotically stable for $\psi(t,y) = \alpha y$ and the Nyquist plot

$$C(j\omega I - A)^{-1}B + D, \quad \omega \in R$$

does not intersect the closed disc with diameter $[-1/\alpha, -1/\beta]$.

Special Case: Positivity

Let $M(j\omega) = C(j\omega I - A)^{-1}B + D$, where A is Hurwitz. The system

$$\begin{cases} \dot{x} &= Ax + Bu, \quad t \geq 0 \\ y &= Cx + Du \\ u &= -\psi(t,y) \end{cases}$$

with sector condition

$$\psi(t,y)/y \geq 0 \quad \forall t \geq 0, y \neq 0$$

is absolutely stable if

$$M(j\omega) + M(j\omega)^* > 0, \quad \forall \omega \in [0,\infty]$$

Proof

Set

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

where P is an $n \times n$ positive definite matrix. Then

$$\begin{aligned}\dot{V} &= 2x^T P\dot{x} \\ &= 2x^T P [A \quad B] \begin{bmatrix} x \\ -\psi \end{bmatrix} \\ &\leq 2x^T P [A \quad B] \begin{bmatrix} x \\ -\psi \end{bmatrix} + 2\psi y \\ &= 2 [x^T \quad -\psi] \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} \begin{bmatrix} x \\ -\psi \end{bmatrix}\end{aligned}$$

By the Kalman-Yakubovich-Popov Lemma, the inequality $M(j\omega) + M(j\omega)^* > 0$ guarantees that P can be chosen to make the upper bound for $\dot{V}$ strictly negative for all $(x,\psi) \neq (0,0)$.

Stability by Lyapunov's theorem.

The Kalman-Yakubovich-Popov Lemma

- Exists in numerous versions
- Idea: **Frequency dependence** is replaced by **matrix parameter** or vice versa

The K-Y-P Lemma, version I

Let $M(j\omega) = C(j\omega I - A)^{-1}B + D$, where A is Hurwitz. Then the following statements are equivalent.

(i) $M(j\omega) + M(j\omega)^* < 0$ for all $\omega \in [0,\infty]$

(ii) $\exists P = P^T > 0$ such that

$$\begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} A & B \\ C & D \end{bmatrix}^T \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} < 0$$

Compare Khalil (5.10-12):

$-M$ is strictly positive real if and only if $\exists P, W, L, \epsilon$:

$$\begin{bmatrix} PA + A^T P & PB + C^T \\ B^T P + C & D + D^T \end{bmatrix} = - \begin{bmatrix} \epsilon P + L^T L & L^T W \\ W^T L & W^T W \end{bmatrix}$$

The K-Y-P Lemma, version II

For

$$\begin{bmatrix} \Phi(s) \\ \widetilde{\Phi}(s) \end{bmatrix} = \begin{bmatrix} C \\ \widetilde{C} \end{bmatrix} (s\widetilde{A} - A)^{-1}(B - s\widetilde{B}) + \begin{bmatrix} D \\ \widetilde{D} \end{bmatrix},$$

with $s\widetilde{A} - A$ nonsingular for some $s \in C$, the following two statements are equivalent.

(i) $\Phi(j\omega)^* \widetilde{\Phi}(j\omega) + \widetilde{\Phi}(j\omega)^* \Phi(j\omega) \leq 0$ for all $\omega \in R$ with $\det(j\omega\widetilde{A} - A) \neq 0$.

(ii) There exists a nonzero pair $(p, P) \in R \times R^{n \times n}$ such that $p \geq 0$, $P = P^*$ and

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{bmatrix} + \begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq 0$$

The corresponding equivalence for strict inequalities holds with $p = 1$.

Some Notation Helps

Introduce

$$M = \begin{bmatrix} A & B \end{bmatrix}, \qquad \widetilde{M} = [I \quad 0],$$

$$N = \begin{bmatrix} C & D \end{bmatrix}, \qquad \widetilde{N} = [0 \quad I].$$

Then

$$y = [C(j\omega I - A)^{-1}B + D]u$$

if and only if

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} N \\ \widetilde{N} \end{bmatrix} w$$

for some $w \in C^{n+m}$ satisfying $Mw = j\omega \widetilde{M} w$.

Lemma 1

Given $y, z \in C^n$, there exists an $\omega \in [0, \infty]$ such that $y = j\omega z$, if and only if $yz^* + zy^* = 0$.

Proof Necessity is obvious. For sufficiency, assume that $yz^* + zy^* = 0$. Then

$$|v^*(y+z)|^2 - |v^*(y-z)|^2 = 2v^*(yz^* + zy^*)v = 0.$$

Hence $y = \lambda z$ for some $\lambda \in C \cup \{\infty\}$. The equality $yz^* + zy^* = 0$ gives that λ is purely imaginary.

Proof of the K-Y-P Lemma

(i) and (ii) can be connected by the following sequence of equivalent statements.

(a) $w^*(\widetilde{N}^*N + N^*\widetilde{N})w < 0$ for $w \neq 0$ satisfying $Mw = j\omega\widetilde{M}w$ with $\omega \in R$.

(b) $\Theta \cap \mathcal{P} = \emptyset$, where

$$\Theta = \left\{ \left(w^*(\widetilde{N}^*N + N^*\widetilde{N})w,\ \widetilde{M}ww^*M^* + Mww^*\widetilde{M}^* \right) : w^*w = 1 \right\}$$

$$\mathcal{P} = \{(r, 0) : r > 0\}$$

(c) $(\mathrm{conv}\ \Theta) \cap \mathcal{P} = \emptyset$.

(d) There exists a hyperplane in $R \times R^{n\times n}$ separating Θ from $\mathcal{P}$, i.e. $\exists P$ such that $\forall w \neq 0$

$$0 > w^* \left(\widetilde{N}^*N + N^*\widetilde{N} + M^*P\widetilde{M} + \widetilde{M}^*PM \right) w$$

Time-invariant Nonlinearity

Let $\psi(y) \in R$ be locally Lipschitz in $y \in R$.

Assume that ψ satisfies the *global sector condition*

$$\alpha \leq \psi(y)/y \leq \beta, \quad \forall t \geq 0, y \neq 0$$

Popov criterion

Let $M(j\omega) = C(j\omega I - A)^{-1}B$, where A is Hurwitz. The system

$$\begin{cases} \dot{x} &=& Ax + Bu, \quad t \geq 0 \\ y &=& Cx \\ u &=& -\psi(y) \end{cases}$$

with sector condition $0 \leq \psi(t, y)/y \leq k$, is absolutely stable if $\exists \eta \in R$ such that

$$1/k + \mathrm{Re}\,[(1 + j\omega\eta)M(j\omega)] > 0, \quad \forall \omega \in [0, \infty]$$

Popov proof I

Set

$$V(x) = x^TPx + 2\eta k \int_0^{Cx} \psi(\sigma)d\sigma$$

where P is an $n \times n$ positive definite matrix. Then

$$\begin{aligned} \dot{V} &= 2(x^TP + \eta k\psi C)\dot{x} \\ &= 2(x^TP + \eta k\psi C)\begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} x \\ -\psi \end{bmatrix} \\ &\leq 2(x^TP + \eta k\psi C)\begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} x \\ -\psi \end{bmatrix} - 2\psi(\psi - ky) \\ &= 2\begin{bmatrix} x^T & -\psi \end{bmatrix}\begin{bmatrix} PA & PB \\ -kC - \eta kCA & -1 - \eta kCB \end{bmatrix}\begin{bmatrix} x \\ -\psi \end{bmatrix} \end{aligned}$$

By the K-Y-P Lemma there is a P that makes the upper bound for $\dot{V}$ strictly negative for all $(x, \psi) \neq (0, 0)$.

Popov proof II

For $\eta \geq 0$, $V > 0$ is obvious for $x \neq 0$.

Stability for linear ψ gives $V \to 0$ and $\dot{V} < 0$, so V must be positive also for $\eta < 0$.

Stability for nonlinear ψ from Lyapunov's theorem.

Example

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -x_1 - x_2 - \psi(x_1), \quad 0 \leq \psi(x_1) \leq k
\end{aligned}
$$

Simultaneous Lyapunov functions

Let $\psi(t, x) \in R^n$ be piecewise continuous in $t \in [0, \infty)$ and locally Lipschitz in $x \in R^n$.

Given matrices $A_1, \ldots, A_N \in R^{n \times n}$, suppose that

$$\psi(t, x) \in \operatorname{conv}\{A_1 x, \ldots, A_N x\}, \quad \forall t, x$$

If there exists a $P = P^T > 0$ such that

$$PA_k + A_k^T P < 0 \quad k = 1, \ldots, N$$

then the origin is globally uniformly asymptotically stable for the system

$$\dot{x} = \psi(t, x)$$

Multi-loop Circle Criterion

Let $M(j\omega) = C(j\omega I - A)^{-1}B + D$, where A is Hurwitz. The system

$$
\left\{
\begin{array}{rcl}
\dot{x} & = & Ax + Bu, \quad t \geq 0 \\
y & = & Cx + Du \\
u_i & = & -\psi_i(t, y_i), \quad i = 1, \ldots, m
\end{array}
\right.
$$

with sector condition

$$\psi_i(t, y)/y \geq 0 \quad \forall t \geq 0, y \neq 0, i = 1, \ldots, m$$

is absolutely stable if $\exists X = \operatorname{diag}\{x_1, \ldots, x_m\}$ such that

$$XM(j\omega) + M(j\omega)^* X > 0, \quad \forall \omega \in [0, \infty]$$

Next week

Dissipativity

Multipliers

Slowly varying systems

Interconnected systems

Extra: Lemma 2

If R and S are matrices of the same size and satisfying $RR^* = SS^*$, then there exists a unitary matrix U such that $R = SU$.

Proof Introduce the polar decompositions

$$R = H_R U_R$$
$$S = H_S U_S$$

where H_R and H_S are hermitean and positive semidefinite, while U_R and U_S are unitary. Then

$$H_R = (RR^*)^{1/2} = (SS^*)^{1/2} = H_S$$

so the unitary matrix $U = U_S^* U_R$ satisfies $R = SU$.

Extra: Lemma 3

If

$$0 = \widetilde{M} W M^* + M W \widetilde{M}^* \qquad (3)$$

for some $W = W^* \geq 0$, then W has the form $W = \sum_{k=1}^{n+m} w_k {w_k}^*$, where

$$0 = \widetilde{M} w_k {w_k}^* M^* + M w_k {w_k}^* \widetilde{M}^*$$

for $k = 1, \ldots, n+m$.

Extra: Proof of Lemma 3

The equality $0 = \widetilde{M} W M^* + M W \widetilde{M}^*$ gives

$$(\widetilde{M} + M) W (\widetilde{M} + M)^* = (\widetilde{M} - M) W (\widetilde{M} - M)^*,$$

so by Lemma 2, there is a unitary matrix U such that $(M + \widetilde{M}) W^{1/2} = (M - \widetilde{M}) W^{1/2} U$.

Diagonalize

$$U = \sum_{k=1}^{n+m} e^{j\theta_k} u_k {u_k}^*$$

where $u_1, \ldots, u_{n+m} \in C^{n+m}$ and $\sum_k u_k u_k^* = I$.

Then $w_k = W^{1/2} u_k$, $k = 1, \ldots, m+n$ fulfil the conditions.

Nonlinear Control Theory 94

Lecture 12

- Dissipativity
- Integral Quadratic Constraints
- Comments on interconnections and slowly varying systems

pp. 268 - 286 + extra material on dissipativity and integral quadratic constraints

Dissipativity

Consider a nonlinear system

$$\begin{cases} \dot{x}(t) &= f(t, x(t), u(t)), \quad t \geq 0 \\ y(t) &= h(t, x(t), u(t)) \end{cases}$$

and a locally integrable function

$$w(t) = w(t, u(t), y(t)).$$

The system is said to be *dissipative* with respect to the *supply rate* w if there exists a *storage function* $S(t, x)$ such that for all t_0, t_1 and inputs u on $[t_0, t_1]$

$$S(t_0, x(t_0)) + \int_{t_0}^{t_1} w(t)dt \geq S(t_1, x(t_1))$$

Example: Memoryless Nonlinearity

The memoryless nonlinearity $u = -\psi(t, y)$ with sector condition

$$\alpha \leq \psi(t, y)/y \leq \beta, \quad \forall t \geq 0, y \neq 0$$

is dissipative with respect to the supply rate

$$w(t) = -[u(t) + \alpha y(t)][u(t) + \beta y(t)]$$

with storage function

$$S(t, x) \equiv 0$$

Linear System Dissipativity

The linear system (minimal realization)

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t), \quad t \geq 0 \\ y(t) &= Cx(t) + Du(t) \end{cases}$$

is dissipative with respect to the supply rate

$$w(t) = u(t)^T y(t)$$

if and only if $G(s) = C(sI - A)^{-1}B + D$ is positive real, i.e.

$$G(s) + G(s)^* \geq 0 \text{ for } \operatorname{Re} s \geq 0$$

A storage function is given by

$$S(x) = \frac{1}{2} x^T P x$$

for any $P = P^T > 0$ satisfying

$$\begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} + \begin{bmatrix} A & B \\ -C & -D \end{bmatrix}^T \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \leq 0$$

Proof

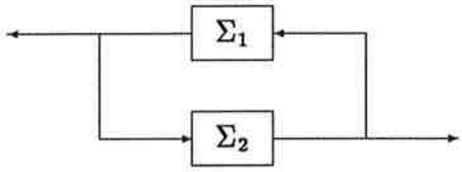
$$
\begin{aligned}
\dot{S} - u^T y &= x^T P \dot{x} - u^T y \\
&= x^T P [A \quad B] \begin{bmatrix} x \\ u \end{bmatrix} - u^T y \\
&= [x^T \quad u^T] \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ -C & -D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
&\le 0
\end{aligned}
$$

by K-Y-P Lemma. Hence

$$S(t_1) - S(t_0) - \int_{t_0}^{t_1} w(t)dt = \int_{t_0}^{t_1} (\dot{S} - u^T y)dt \le 0$$

which is the desired dissipation inequality.

Feedback Dissipativity I



For two dissipative systems Σ_1, Σ_2 with corresponding inputs u_1, u_2, outputs y_1, y_2, supply rates w_1, w_2 and storage functions S_1, S_2, suppose that the feedback interconnection

$$u_2 = y_1 \quad u_1 = y_2$$

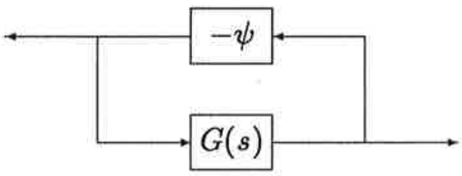
is "well posed" and that

$$0 = w_1(u, y) + w_2(y, u) \qquad \text{(neutral interconnection)}$$

Then the feedback system is dissipative with respect to the supply rate $w(y_1, y_2) \equiv 0$ and the storage function

$$S_1(x_1) + S_2(x_2)$$

Linear system with memoryless feedback



A "well posed" interconnection of the memoryless nonlinearity $u = -\psi(t, y)$ satisfying

$$\psi(t, y)y \ge 0 \quad \forall t \ge 0$$

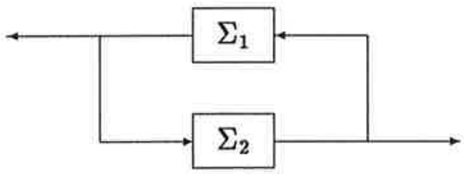
with the positive real system

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), & t \ge 0 \\ y(t) = Cx(t) + Du(t) \end{cases}$$

gives a closed loop system that is dissipative with respect to zero supply rate, with storage function

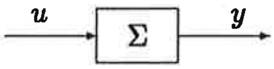
$$S(x) = \frac{1}{2} x^T P x$$

Applicability?



How find supply rate that make Σ_1 and Σ_2 simultaneously dissipative?

Integral Quadratic Constraint

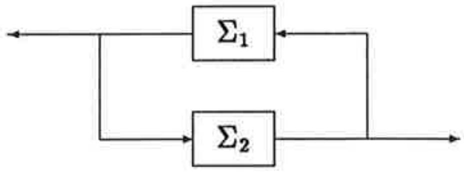


The nonlinear bounded operator $\Sigma : L_2^m \rightarrow L_2^m$ is said to *satisfy the IQC defined by* Π, if

$$\int_{-\infty}^{\infty} \begin{bmatrix} \widehat{u}(j\omega) \\ \widehat{y}(j\omega) \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} \widehat{u}(j\omega) \\ \widehat{y}(j\omega) \end{bmatrix} d\omega \geq 0$$

for $u, y \in L_2^m$ with $y = \Sigma(u)$.

Feedback Dissipativity II



Suppose $\Pi_{11}(j\omega) \geq 0$ and $\Pi_{22}(j\omega) \leq 0$.

If Σ_1 satisfies the IQC defined by

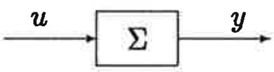
$$\begin{bmatrix} \Pi_{11}(j\omega) & \Pi_{12}(j\omega) \\ \Pi_{21}(j\omega) & \Pi_{22}(j\omega) \end{bmatrix}$$

and Σ_2 satisfies the IQC defined by

$$-\begin{bmatrix} \Pi_{22}(j\omega) & \Pi_{21}(j\omega) \\ \Pi_{12}(j\omega) & \Pi_{11}(j\omega) \end{bmatrix}$$

then the feedback system, if "well posed", is dissipative with supply rate zero.

Contractiveness

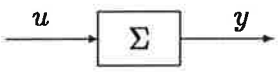


A contractive operator satisfies the IQC defined by

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

I denotes the unit matrix of appropriate dimension.

Passivity



A passive operator satisfies the IQC defined by

$$\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

Linear Time-invariant Dynamics

$$\xrightarrow{u} \boxed{\Sigma(j\omega)} \xrightarrow{y}$$

A linear time-invariant contractive operator satisfies any IQC defined by

$$\begin{bmatrix} x(j\omega)I & 0 \\ 0 & -x(j\omega)I \end{bmatrix}$$

where $x(j\omega) \geq 0$ is a bounded measurable function.

Linear Time-invariant Scalar Dynamics

$$\xrightarrow{u} \boxed{\delta(j\omega)I} \xrightarrow{y}$$

An operator, defined by multiplication in frequency domain with a scalar contractive transfer function, satisfies any IQC defined by

$$\begin{bmatrix} X(j\omega) & 0 \\ 0 & -X(j\omega) \end{bmatrix}$$

where $X(j\omega) = X(j\omega)^* \geq 0$ is a bounded measurable matrix function.

Constant Real Scalar

$$\xrightarrow{u} \boxed{\delta I} \xrightarrow{y}$$

Multiplication with a real constant satisfies IQC's defined by matrix functions of the form

$$\begin{bmatrix} 0 & Y(j\omega) \\ Y(j\omega)^* & 0 \end{bmatrix}$$

where $Y(j\omega) = -Y(j\omega)^*$ is bounded and measurable.

(Upper bound for structured singular values)

Time-varying Real Scalar

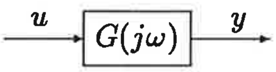
$$\xrightarrow{u} \boxed{\delta(t)I} \xrightarrow{y}$$

Multiplication in the time-domain with a scalar function $\delta \in L_\infty$ with $\|\delta\|_\infty \leq 1$ satisfies IQC's defined by

$$\begin{bmatrix} X & Y \\ Y^T & -X \end{bmatrix}$$

where $X = X^T \geq 0$ and $Y = -Y^T$ are real matrices [Feron, 1994!].

IQC for Transfer Function

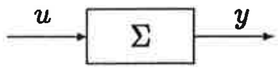


The transfer matrix $G(s)$ satisfies the IQC defined by Π if and only if it is stable and

$$\begin{bmatrix} I \\ G(j\omega) \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} I \\ G(j\omega) \end{bmatrix} \geq 0$$

for $\omega \in [0, \infty]$.

Dissipativity from IQC's



Suppose

$$\begin{aligned} 0 &\geq \Pi_{22}(j\omega), \quad \forall \omega \in R \\ \Pi(j\omega) &= \Phi(j\omega) + \Phi(j\omega)^* \\ \Phi(j\omega) &= C_\Phi(sI - A_\Phi)^{-1}B_\Phi + D_\Phi \\ &= \int_0^\infty e^{-j\omega t}\phi(t)dt \end{aligned}$$

Let $x(0) = 0$ and $x_\Phi(0)$ and

$$\left\{ \begin{array}{rcl} \dot{x} &=& f(t, x, u) \\ \dot{x}_\Phi &=& A_\Phi x_\Phi + B_\Phi [\, u^T \quad y^T \,]^T \\ y &=& h(t, x, u) \\ y_\Phi &=& C_\Phi x_\Phi + D_\Phi [\, u^T \quad y^T \,]^T \end{array} \right. \tag{1}$$

If the map from u to y satisfies the IQC defined by Π, then (1) with supply rate

$$w = [\, u^T \quad y^T \,]\, y_\Phi$$

is dissipative with storage function

$$S(t, x(t), x_\Phi(t)) = \inf_u \int_0^t w(\tau)d\tau$$

Proof

Given u and $y = \Sigma(u)$, define

$$(u_t(\tau), y_t(\tau)) = \begin{cases} (u(\tau), y(\tau)), & \tau \leq t \\ (0, 0), & \tau > t \end{cases}$$

Then

$$\int_0^t w(\tau)d\tau = \int_0^t \begin{bmatrix} u \\ y \end{bmatrix}^T \left(\phi * \begin{bmatrix} u \\ y \end{bmatrix}\right) d\tau$$

$$= \int_0^\infty \begin{bmatrix} u_t \\ \Sigma(u_t) \end{bmatrix}^T \left(\phi * \begin{bmatrix} u_t \\ \Sigma(u_t) \end{bmatrix}\right) d\tau$$
$$- \int_0^\infty \begin{bmatrix} 0 \\ \Sigma(u_t) - y_t \end{bmatrix}^T \left(\phi * \begin{bmatrix} 0 \\ \Sigma(u_t) - y_t \end{bmatrix}\right) d\tau$$

$$= \frac{1}{2}\int_{-\infty}^\infty \begin{bmatrix} \widehat{u}_t \\ \widehat{\Sigma(u_t)} \end{bmatrix}^* \Pi \begin{bmatrix} \widehat{u}_t \\ \widehat{\Sigma(u_t)} \end{bmatrix} d\omega$$
$$- \frac{1}{2}\int_{-\infty}^\infty \begin{bmatrix} 0 \\ \widehat{\Sigma(u_t)} - \widehat{y}_t \end{bmatrix}^* \Pi \begin{bmatrix} 0 \\ \widehat{\Sigma(u_t)} - \widehat{y}_t \end{bmatrix} d\omega$$

$$\geq \frac{1}{2}\int_{-\infty}^\infty \begin{bmatrix} \widehat{u}_t \\ \widehat{\Sigma(u_t)} \end{bmatrix}^* \Pi \begin{bmatrix} \widehat{u}_t \\ \widehat{\Sigma(u_t)} \end{bmatrix} d\omega \geq 0$$

"*" denotes convolution

Proof of Feedback Dissipativity II

$\Sigma_1 : u \to y$ is dissipative with supply rate

$$\begin{aligned} w &= [\, u^T \quad y^T \,]\, y_\Phi \\ &= [\, u^T \quad y^T \,] \left(\phi * \begin{bmatrix} u \\ y \end{bmatrix}\right) \end{aligned}$$

$\Sigma_2 : y \to u$ is dissipative with supply rate

$$w = -[\, u^T \quad y^T \,] \left(\phi * \begin{bmatrix} u \\ y \end{bmatrix}\right)$$

Feedback Dissipativity II therefore follows from Feedback Dissipativity I.

Example: Friction model

Interconnected systems

$$\begin{cases} \dot{x}_1 &= f_1(t, x_1) + g_1(t, x) \\ & \vdots \\ \dot{x}_m &= f_m(t, x_m) + g_m(t, x) \end{cases}$$

Suppose $V_1(t, x_1), \ldots, V_m(t, x_m) > 0$ with

$$\begin{aligned} \frac{\partial V_i}{\partial t} + \frac{\partial V_i}{\partial x_i} f_i(t, x_i) &\leq -\alpha_i \phi_i^2(x_i) \\ \left\| \frac{\partial V_i}{\partial x_i} \right\| &\leq \beta_i \phi_i(x_i) \\ \|g_i(t, x)\| &\leq \sum_{i=1}^{m} \gamma_{ij} \phi_j(x_j) \end{aligned}$$

Then, with $D = \operatorname{diag}\{d_1, \ldots, d_m\}$ and properly defined S

$$\begin{aligned} \sum_i d_i \dot{V}(t, x) &\leq \sum_i d_i \left[-\alpha_i \phi_i^2(x_i) + \sum_j \beta_i \gamma_{ij} \phi_i(x_i) \phi_j(x_j) \right] \\ &= -\phi^T (DS + S^T D) \phi / 2 \end{aligned}$$

Stability if $DS + S^T D > 0$ for some D!

Slowly varying systems

Suppose

$$\dot{x} = f(x, u(t)) \tag{2}$$

where $\|\dot{u}\|$ is small. Equilibrium

$$0 = f(h(u), u)$$

Change variables $z = x - h(u)$. Analyse

$$\dot{z} = f(z + h(u), u)$$

for "frozen" u. Bounds on Lyapunov function plus bounds on $\partial h / \partial u$ gives results on (2).

Next week

Oscillations

Describing functions

Stability analysis

Application: Delta-Sigma Modulators

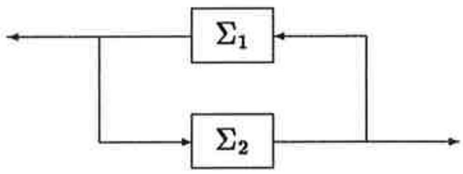
Nonlinear Control Theory 94

Lecture 13

- Oscillations
- Poincare-Bendixon Theorem
- Describing functions

Khalil pp. 339 - 407.

Last week



For two dissipative systems Σ_1, Σ_2 with corresponding inputs u_1, u_2, outputs y_1, y_2, supply rates w_1, w_2 and storage functions S_1, S_2, suppose that the feedback interconnection

$$u_2 = y_1 \quad u_1 = y_2$$

is "well posed" and that

$$0 = w_1(u, y) + w_2(y, u) \qquad \text{(neutral interconnection)}$$

Then the feedback system is dissipative with respect to the supply rate $w(y_1, y_2) \equiv 0$ and the storage function

$$S_1(x_1) + S_2(x_2)$$

Yakubovich's Oscillation Condition

Periodic Orbits

- Second Order Systems
- Describing functions
- Averaging

Poincare-Bendixon Theorem

Suppose that $x(t)$ satisfies

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

and stays for $t \geq 0$ in a compact set $K \subset \Omega$, that contains no equilibrium. Then, either x is a perodic orbit, or it spirals towards a perodic orbit as $t \to \infty$.

Proof

Limit point y by Bolzano-Weierstrass.

Periodic orbit through y:

x spirals towards the periodic orbit:

Example

Consider

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

and

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

Poincare-Bendixon can be applied to the region

$$\{x : 1 - \epsilon < V(x) < 1 + \epsilon\}$$

for any $\epsilon \in (0, 1)$.

Problem Statement

Detect periodic solutions of

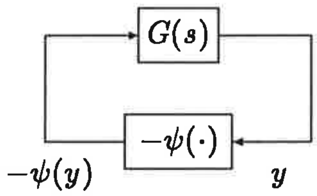
$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx \\ u &= -\psi(y) \end{cases}$$

Assumptions

$$\psi(-y) = -\psi(y) \qquad \forall y$$

$$\alpha(y_2 - y_1) \leq \psi(y_2) - \psi(y_1) \leq \beta(y_2 - y_1) \qquad \forall y_1 < y_2$$

Describing function



$$y = a\sin(\omega t) + y_h$$

$$\psi(a\sin(\omega t)) = \sum_{k\,\text{odd}} b_k e^{j\omega k t}$$

Describing function

$$\Psi(a) = \frac{b_1}{a/2j} = \frac{2\omega}{\pi a}\int_0^{\pi/\omega} \psi(a\sin\omega t)\sin\omega t dt$$

Describing function idea

A solution (ω, a) of

$$G(j\omega) = -\frac{1}{\Psi(a)}$$

indicates a periodic orbit with $a\sin(\omega t)$ as the output of the linear part.

Notation

Loss-pass measure

$$\rho(\omega) = \inf_{k>1, k\,\text{odd}} \left|\frac{\alpha+\beta}{2} + \frac{1}{G(jk\omega)}\right|$$

Error circle frequencies

$$\Omega = \{\omega |\ \rho(\omega) > (\beta-\alpha)/2\}$$

Error circle radius

$$\sigma(\omega) = \frac{\left(\frac{\beta-\alpha}{2}\right)^2}{\rho(\omega) - \frac{\beta-\alpha}{2}}, \quad \omega \in \Omega$$

High frequency interval

$$\tilde{\Omega} = \left\{\omega : \left|\frac{\alpha+\beta}{2} + \frac{1}{G(jk\omega)}\right| > \frac{\beta-\alpha}{2},\ k = 1,3,5,\ldots\right\}$$

Complete Intersection

A *region of complete intersection* Γ is a connected component of the set

$$\left\{(\omega, a) : \left|\Psi(a) + \frac{1}{G(j\omega)}\right| < \sigma(\omega)\right\}$$

that contains a unique pair (ω_s, a_s) satisfying

$$\Psi(a_s) + \frac{1}{G(j\omega_s)} = 0$$

$$\left.\frac{d}{da}\Psi(a)\right|_{a=a_s} \neq 0$$

$$\left.\frac{d}{d\omega}\operatorname{Im} G(j\omega)\right|_{\omega=\omega_s} \neq 0$$

Theorem on Describing Functions

(a) There exist no half-wave symmetric periodic solutions with frequency $\omega \in \tilde{\Omega}$

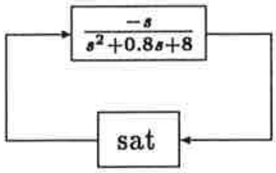
(b) There exist no half-wave symmetric periodic solutions with frequency $\omega \in \Omega'$ unless for some a

$$\left|\Psi(a) + \frac{1}{G(j\omega)}\right| \leq \sigma(\omega)$$

(c) There exists at least one half-wave symmetric periodic solution for each region of complete intersection

[see Mees, Dynamics of Feedback Systems, 1981, pp. 126-160]

Example



Unique intersection at $(\omega_s, a_s) = (2\sqrt{2}, 1.455)$ with region

$$\Gamma = [1.37, 1.54] \times [2.81, 2.85]$$

and regularity conditions satisfied.

Hence periodic orbit exists.

Equations for a Periodic Orbit

For $y(t) = \sum_{k\,\mathrm{odd}} a_k e^{j\omega k t}$, define

$$||y||^2 = 2 \sum_{k\,\mathrm{odd}} |a_k|^2$$

$$y_1(t) = Py(t) = a_{-1}e^{-j\omega kt} + a_1 e^{j\omega kt}$$

$$y_h(t) = (I - P)y(t) = \sum_{|k|>1, k\,\mathrm{odd}} a_k e^{j\omega t}$$

$$gy(t) = \sum_{k\,\mathrm{odd}} G(j\omega k) a_k e^{j\omega kt}$$

A periodic orbit solves $0 = y + g\psi(y)$, or componentwise

$$\begin{cases} 0 = y_1 + Pg\psi(y_1 + y_h) \\ 0 = y_h + (I - P)g\psi(y_1 + y_h) \end{cases}$$

Proof of (b)

Set $-\alpha = \beta = 1$. Given y_1, the operator

$$Ty_h = (P - I)g\psi(y_1 + y_h)$$

satisfies the Lipschitz condition

$$\begin{aligned} ||Ty^{(2)} - Ty^{(1)}|| &\leq \max_{k>1} |G(j\omega k)|\, ||y^{(2)} - y^{(1)}|| \\ &= \frac{1}{\rho(\omega)} ||y^{(2)} - y^{(1)}|| \end{aligned}$$

A periodic orbit with $y_1(t) = a \sin \omega t$ satisfies $Ty_h = y_h$ and $T(-y_1) = 0$, so

$$||y_h|| = ||Ty_h - T(-y_1)|| \leq \frac{1}{\rho(\omega)}(||y_h|| + a)$$

$$||y_h|| \leq \frac{a}{\rho(\omega) - 1} = a\sigma(\omega)$$

Hence a necessary condition is that

$$\begin{aligned} |a + G(j\omega)\Psi(a)a| &= ||y_1 + Pg\psi(y_1)|| \\ &= ||Pg[\psi(y_1) - \psi(y_1 + y_h)]|| \\ &\leq |G(j\omega)|\, ||y_h|| \leq |G(j\omega)|\sigma(\omega)a \end{aligned}$$

$$\left|\Psi(a) + \frac{1}{G(j\omega)}\right| \leq \sigma(\omega)$$

A Homotopy Approach to (c)

Define

$$\phi_\mu(y_1, y_h) = \begin{bmatrix} y_1 + (1-\mu)Pg\psi(y_1) + \mu Pg\psi(y_1 + y_h) \\ y_h + \mu(I-P)g\psi(y_1 + y_h) \end{bmatrix}$$

A periodic orbit solves

$$\phi_\mu(y_1, y_h) = 0$$

for $\mu = 1$. Harmonic balance

$$0 = \Psi(a) + \frac{1}{G(j\omega)}$$

means solvability for $\mu = 0$.

Degree Theory Review

Suppose D is a bounded open subspace of a normed space E. Let ϕ be compact perturbation of the identity, with $p \notin \phi(\partial D)$. Define the *degree of ϕ relative to D* by

$$d(\phi, D, p) = \sum_{x_i \in \phi^{-1}(p)} \text{sign det}\left[\frac{\partial \phi}{\partial x}(x_i)\right]$$

Then the degree has the following two basic properties.

- If $d(\phi, D, p) \neq 0$, then $\phi(x) = p$ has at least one solution in D.
- Suppose $\phi_\mu = I + h_\mu$, where h_μ is compact for $\mu \in [0,1]$ and depends continuously on μ. If $p \notin \phi_\mu(\partial D)$ for all $\mu \in [0,1]$, then $d(\phi_\mu, D, p)$ is independent of μ on $[0,1]$.

Proof of (c)

Without resctriction, set $-\alpha = \beta = 1$. For any complete intersection Γ, define

$$D = \{(a \sin \omega t, y_h) : (\omega, a) \in \Gamma, \; \|y_h\| \leq \sigma(\omega)a\}$$

Then $d(\phi_0, D, 0) \neq 0$. It remains to prove that $\phi_\mu(y_1, y_h) \neq 0$ for $(y_1, y_h) \in \partial D$. On ∂D, we get from the definition of D that

$$\|y_1 + Pg\psi(y_1)\| \leq G(j\omega)\sigma(\omega)a$$
$$\|y_h\| \leq \sigma(\omega)a$$

where at least one inequality holds as equality.

Proof of (c) Continued

For $(y_1, y_h) \in \partial D$ we get

$$\begin{aligned}
&\|y_1 + (1-\mu)Pg\psi(y_1) + \mu Pg\psi(y_1 + y_h)\| \\
&\quad \geq \|y_1 + Pg\psi(y_1)\| - \mu\|Pg\psi(y_1 + y_h) - Pg\psi(y_1)\| \\
&\quad \geq \|y_1 + Pg\psi(y_1)\| - \mu G(j\omega)\|y_h\| \\
&\quad \geq \|y_1 + Pg\psi(y_1)\| - \mu G(j\omega)\sigma(\omega)a
\end{aligned}$$

Using that $\rho(\omega)(1 + \sigma(\omega)) = \sigma(\omega)$ we also get

$$\begin{aligned}
&\|y_h + \mu(I-P)g\psi(y_1 + y_h)\| \\
&\quad \geq \|y_h\| - \mu\|(I-P)g\psi(y_1)\| - \mu\|(I-P)g\psi(y_h)\| \\
&\quad \geq \|y_h\| - \mu\rho(\omega)a - \mu\rho(\omega)\sigma(\omega)a \\
&\quad \geq \|y_h\| - \mu\sigma(\omega)a
\end{aligned}$$

At least one of the right hand sides must be strictly positive for $\mu < 1$, so $\phi_\mu(y_1, y_h) \neq 0$ on ∂D and the proof is complete.

Delta-Sigma Modulators

- Fast digital circuits rather than precise analog circuits
- Feedback shapes noise spectrum
- Feedback may cause instability
- Describing functions help design

Stability Analysis

Indication of stable oscillations

Indication of unstable oscillations

Warning Example

System in polar coordinates:

$$\dot{r} = \frac{(1-r^2)^3}{r}$$
$$\dot{\theta} = 1 + (1-r^2)^2$$

has solution

$$r(t) = \left[1 - \frac{1-r_0^2}{\sqrt{1+4t(1-r_0^2)^2}}\right]^{1/2}$$
$$\theta(t) = \theta_0 + t + \frac{1}{4}\ln[1+4t(1-r_0^2)^2]$$

Should the orbit $x_1 = \cos t$, $x_2 = \sin t$ be considered stable?

Next Week

Averaging

Singular Perturbations

Nonlinear Control Theory 94

Lecture 14

- Periodic Perturbations
- Averaging
- Singular Perturbations

Khalil pp. 408 - 476.

Last week

(a) There exist no half-wave symmetric periodic solutions with frequency $\omega \in \tilde{\Omega}$

(b) There exist no half-wave symmetric periodic solutions with frequency $\omega \in \Omega'$ unless for some a

$$\left| \Psi(a) + \frac{1}{G(j\omega)} \right| \leq \sigma(\omega)$$

(c) There exists at least one half-wave symmetric periodic solution for each region of complete intersection

Today: Two Time-scales

Averaging

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

The state x moves slowly compared to f.

Singular perturbations

$$\begin{aligned} \dot{x} &= f(t, x, z, \epsilon) \\ \epsilon \dot{z} &= g(t, x, z, \epsilon) \end{aligned}$$

The state x moves slowly compared to z.

Example: Vibrating Pendulum I

Newtons law in tangential direction

$$\begin{aligned} &m(l\ddot{\theta} - a\omega^2 \sin \omega t \sin \theta) \\ &\quad = -mg \sin \theta - k(l\dot{\theta} + a\omega \cos \omega t \sin \theta) \end{aligned}$$

With $\epsilon = a/l, \tau = \omega t, \alpha = \omega_0 l/\omega a,\ \beta = k/m\omega_0$

$$\begin{aligned} x_1 &= \theta \\ x_2 &= \epsilon^{-1}(d\theta/d\tau) + \cos \tau \sin \theta \\ f_1(\tau, x) &= x_2 - \cos \tau \sin x_1 \\ f_2(\tau, x) &= -\alpha\beta x_2 - \alpha^2 \sin x_1 \\ &\qquad + x_2 \cos \tau \cos x_1 - \cos^2 \tau \sin x_1 \cos x_1 \end{aligned}$$

the state equation is given by

$$\frac{dx}{d\tau} = \epsilon f(\tau, x)$$

Averaging Assumptions

Consider the system

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

where f and its derivatives up to second order are continuous and bounded.

Let x_{av} be defined by the equations

$$\dot{x}_{\text{av}}(t, \epsilon) = \epsilon f_{\text{av}}(x_{\text{av}}(t, \epsilon)), \quad x_{\text{av}}(0, \epsilon) = x_0$$
$$f_{\text{av}}(x) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(\tau, x, 0) d\tau$$

Example: Vibrating Pendulum II

The averaged system

$$\begin{aligned}\dot{x} &= \epsilon f_{\text{av}}(x) \\ &= \epsilon \begin{bmatrix} x_2 \\ -\alpha\beta x_2 - \alpha^2 \sin x_1 - \frac{1}{4}\sin 2x_1 \end{bmatrix}\end{aligned}$$

has

$$\frac{\partial f_{\text{av}}}{\partial x}(\pi, 0) = \begin{bmatrix} 0 & 1 \\ \alpha^2 - 0.5 & -\alpha\beta \end{bmatrix}$$

which is Hurwitz for $0 < \alpha < 1/\sqrt{2}$, $\beta > 0$.

Can this be used for rigorous conclusions?

Periodic Averaging Theorem

Let f be periodic in t with period T.

Let $x = 0$ be an exponentially stable equilibrium of $\dot{x}_{\text{av}} = \epsilon f(x_{\text{av}}, 0)$.

If $|x_0|$ is sufficiently small, then

$$x(t, \epsilon) = x_{\text{av}}(t, \epsilon) + O(\epsilon) \text{ for all } t \in [0, \infty]$$

Furthermore, for sufficiently small $\epsilon > 0$, the equation $\dot{x} = \epsilon f(t, x, \epsilon)$ has a unique exponentially stable periodic solution of period T in an $O(\epsilon)$ neighbourhoodd of $x = 0$.

General Averaging Theorem

Under certain conditions on the convergence of

$$f_{\text{av}}(x, \epsilon) = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(\tau, x, \epsilon) d\tau$$

there exists a $C > 0$ such that for sufficiently small $\epsilon > 0$

$$|x(t, \epsilon) - x_{\text{av}}(t, \epsilon)| < C\epsilon$$

for all $t \in [0, 1/\epsilon]$.

Example: Vibrating Pendulum III

The Jacobian of the averaged system is Hurwitz for $0 < \alpha < 1/\sqrt{2}$, $\beta > 0$.

For a/l sufficiently small and

$$\omega > \sqrt{2}\omega_0 l/a$$

the unstable pendulum equilibrium $(\theta, \dot{\theta}) = (\pi, 0)$ is therefore stabilized by the vibrations.

Periodic Perturbation Theorem

Consider

$$\dot{x} = f(x) + \epsilon g(t, x, \epsilon)$$

where f, g, $\partial f/\partial x$ and $\partial g/\partial x$ are continuous and bounded.

Let g be periodic in t with period T.

Let $x = 0$ be an exponentially stable equilibrium point for $\epsilon = 0$.

Then, for sufficiently small $\epsilon > 0$, there is a unique periodic solution

$$\bar{x}(t, \epsilon) = O(\epsilon)$$

which is exponentially stable.

Proof ideas of Periodic Perturbation Theorem

Let $\phi(t, x_0, \epsilon)$ be the solution of

$$\dot{x} = f(x) + \epsilon g(t, x, \epsilon), \quad x(0) = x_0$$

- Exponential stability of $x = 0$ for $\epsilon = 0$, plus bounds on the magnitude of g, shows existence of a bounded solution $\bar{x}$ for small $\epsilon > 0$.
- The implicit function theorem shows solvability of

  $$x = \phi(T, 0, x, \epsilon)$$

  for small ϵ. This gives periodicity of $\bar{x}$.
- Put $z = x - \bar{x}$. Exponential stability of $x = 0$ for $\epsilon = 0$ gives exponential stability of $z = 0$ for small $\epsilon > 0$.

Proof idea of Averaging Theorem

For small $\epsilon > 0$ define u and y by

$$u(t, x) = \int_0^t [f(\tau, x, 0) - f_{\text{av}}(x)] d\tau$$
$$x = y + \epsilon u(t, y)$$

Then

$$\left[I + \epsilon \frac{\partial u}{\partial y}\right] \dot{y} = \epsilon f(t, y + \epsilon u, \epsilon) - \epsilon \frac{\partial u}{\partial t}(t, y)$$
$$= \epsilon f_{\text{av}}(y) + \epsilon^2 p(t, y, \epsilon)$$

With $s = \epsilon t$,

$$\frac{dy}{ds} = f_{\text{av}}(y) + \epsilon q\left(\frac{s}{\epsilon}, y, \epsilon\right)$$

which has a unique and exponentially stable periodic solution for small ϵ. This gives the desired result.

Application: Second Order Oscillators

For the second order system

$$\ddot{y} + \omega^2 y = \epsilon g(y, \dot{y}) \tag{1}$$

introduce

$$\begin{aligned} y &= r \sin\phi \\ \dot{y}/\omega &= r\cos\phi \\ f(\phi, r, \epsilon) &= \frac{g(r\sin\phi, \omega r\cos\phi)\cos\phi}{\omega^2 - (\epsilon/r) g(r\sin\phi, \omega r\cos\phi)\sin\phi} \\ f_{\mathrm{av}}(r) &= \frac{1}{2\pi}\int_0^{2\pi} f(\phi, r, 0) d\phi \\ &= \frac{1}{2\pi\omega^2}\int_0^{2\pi} g(r\sin\phi, \omega r\cos\phi)\cos\phi d\phi \end{aligned}$$

Then (1) is equivalent to

$$\frac{dr}{d\phi} = \epsilon f(\phi, r, \epsilon)$$

and the periodic averaging theorem may be applied.

Illustration: Van der Pol Oscillator I

Example: Van der Pol Oscillator I

The vacuum tube cicuit equation

$$\ddot{y} + y = \epsilon\dot{y}(1 - y^2)$$

gives

$$\begin{aligned} f_{\mathrm{av}}(r) &= \frac{1}{2\pi}\int_0^{2\pi} r\cos\phi(1 - r^2\sin^2\phi)\cos\phi d\phi \\ &= \frac{1}{2}r - \frac{1}{8}r^3 \end{aligned}$$

The averaged system

$$\frac{dr}{d\phi} = \epsilon\left(\frac{1}{2}r - \frac{1}{8}r^3\right)$$

has equilibria $r = 0$, $r = 2$ with

$$\left.\frac{df_{\mathrm{av}}}{dr}\right|_{r=2} = -1$$

so small ϵ give a stable limit cycle, which is close to circular with radius $r = 2$.

Singular Perturbations

Consider equations of the form

$$\begin{aligned} \dot{x} &= f(t, x, z, \epsilon), \quad x(0) = x_0 \\ \epsilon\dot{z} &= g(t, x, z, \epsilon) \quad z(0) = z_0 \end{aligned}$$

For small $\epsilon > 0$, the first equation discribes the *slow dynamics*, while the second equation defines the *fast dynamics*.

The main idea will be to approximate x with the solution of the *reduced problem*

$$\dot{\bar{x}} = f(t, x, h(t, \bar{x}), 0) \quad \bar{x}(0) = x_0$$

where $h(t, \bar{x})$ is defined by the equation

$$0 = g(t, x, h(t, x), 0)$$

Example: DC Motor I

$$J\frac{d\omega}{dt} = ki$$
$$L\frac{di}{dt} = -k\omega - Ri + u$$

With $x = \omega$, $z = i$ and $\epsilon = Lk^2/JR^2$ we get

$$\dot{x} = z$$
$$\epsilon\dot{z} = -x - z + u$$

Linear Singular Perturbation Theorem

Let the matrix A_{22} have nonzero eigenvalues $\gamma_1, \dots, \gamma_m$ and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A_0 = A_{11} - A_{12}A_{22}^{-1}A_{21}$.

Then, $\forall\delta > 0\ \exists\epsilon_0 > 0$ such that the eigenvalues $\alpha_1, \dots, \alpha_{n+m}$ of the matrix

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21}/\epsilon & A_{22}/\epsilon \end{bmatrix}$$

satisfy the bounds

$$|\lambda_i - \alpha_i| < \delta, \quad i = 1, \dots, n$$
$$|\gamma_{i-n} - \epsilon\alpha_i| < \delta, \quad i = n+1, \dots, n+m$$

for $0 < \epsilon < \epsilon_0$.

Proof

A_{22} is invertible, so it follows from the implicit function theorem that for sufficiently small ϵ the Riccati equation

$$\epsilon A_{11}P_\epsilon + A_{12} - \epsilon P_\epsilon A_{21}P_\epsilon - P_\epsilon A_{22} = 0$$

has a uniqe solution $P_\epsilon = A_{12}A_{22}^{-1} + O(\epsilon)$.

The desired result now follows from the similarity transformation

$$\begin{bmatrix} I & -\epsilon P_\epsilon \\ 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21}/\epsilon & A_{22}/\epsilon \end{bmatrix}\begin{bmatrix} I & \epsilon P_\epsilon \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} I & -\epsilon P_\epsilon \\ 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} + \epsilon A_{11}P_\epsilon \\ A_{21}/\epsilon & A_{22}/\epsilon + A_{21}P_\epsilon \end{bmatrix}$$
$$= \begin{bmatrix} A_0 + O(\epsilon) & 0 \\ * & A_{22}/\epsilon + O(1) \end{bmatrix}$$

Example: DC Motor II

In the example

$$\dot{x} = z$$
$$\epsilon\dot{z} = -x - z + u$$

we have

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$$

$$A_{11} - A_{12}A_{22}^{-1}A_{21} = -1$$

so stability of the DC motor model for small

$$\epsilon = \frac{Lk^2}{JR^2}$$

is verified.

The Boundary-Layer System

For fixed (t, x) the *boundary layer system*

$$\frac{d\hat{y}}{d\tau} = g(t, x, \hat{y} + h(t, x), 0), \quad \hat{y}(0) = z_0 - h(0, x_0)$$

describes the fast dynamics, disregarding variations in the slow variables t, x.

Tikhonov's Theorem

Consider a singular perturbation problem with $f, g, h, \partial g/\partial x \in C^1$. Assume that the reduced problem has a unique bounded solution $\bar{x}$ on $[0, T]$ and that the equilibrium $\hat{y} = 0$ of the boundary layer problem is exponentially stable uniformly in (t, x). Then

$$x(t, \epsilon) = \bar{x}(t) + O(\epsilon)$$
$$z(t, \epsilon) = h(t, \bar{x}(t)) + \hat{y}(t/\epsilon) + O(\epsilon)$$

uniformly for $t \in [0, T]$.

Example: High Gain Feedback

Closed loop system

$$\dot{x}_p = Ax_p + Bu_p$$
$$\frac{1}{k_1}\dot{u}_p = \psi(u - u_p - k_2 C x_p)$$

Reduced model

$$\dot{x}_p = (A - Bk_2C)x_p + Bu$$

Proof ideas of Tikhonov's Theorem

Replace f and g with F and G that are identical for $|x| < r$, but nicer for large x.

For small ϵ, $G(t, x, y, \epsilon)$ is close to $G(t, x, y, 0)$.

y-bound for $G(\cdot, \cdot, \cdot, 0)$-equation
$\Rightarrow$ y-bound for G-equation
$\Rightarrow$ x, y-bound for F, G-equations

For small $\epsilon > 0$, the x, y-solutions of the F, G-equations will satisfy $|x| < r$. Hence, they also solve the f, g-equations

The Slow Manifold

For small $\epsilon > 0$, the system

$$\dot{x} = f(x, z)$$
$$\epsilon\dot{z} = g(x, z)$$

has the invariant manifold

$$z = H(x, \epsilon)$$

It can often be computed approximately by Taylor expansion

$$H(x, \epsilon) = H_0(x) + \epsilon H_1(x) + \epsilon^2 H_2(x) + \cdots$$

where H_0 satisfies

$$0 = g(x, H_0)$$

The Fast Manifold

All solutions of

$$\dot{x} = f(x, z)$$
$$\epsilon\dot{z} = g(x, z)$$

approach the slow maifold along a fast manifold approximately satisfying

$$x = \text{constant}$$

Example: Van der Pol Oscillator III

Consider

$$\frac{d^2v}{ds^2} - \mu(1 - v^2)\frac{dv}{ds} + v = 0$$

With

$$x = -\frac{1}{\mu}\frac{dv}{ds} + v - \frac{1}{3}v^3$$
$$z = v$$
$$t = s/\mu$$
$$\epsilon = 1/\mu^2$$

we have the system

$$\dot{x} = z$$
$$\epsilon\dot{z} = -x + z - \frac{1}{3}z^3$$

with slow manifold

$$x = z - \frac{1}{3}z^3$$

Illustration: Van der Pol III

Course Review of Second Part

- Lyapunov functions
- Absolute stability
- Dissipativity
- Describing functions
- Averaging

Research needs

- Effective computations
- Model hierarchies
- Performance measures
- Controller suboptimization
- Adaptation

A Note on the Kalman-Yacubovich-Popov Lemma

Anders Rantzer

1 Statement of the Main result

The purpose of this note is to present a new elementary proof for the so called Kalman-Yakubovich-Popov lemma [And67]. The basic idea has been suggested in connection to the theory for structured singular values [PD93]. Here we give the details and refinements needed to cover all main versions of the classical result.

Theorem 1 *For*

$$\begin{bmatrix} \Phi(s) \\ \widetilde{\Phi}(s) \end{bmatrix} = \begin{bmatrix} C \\ \widetilde{C} \end{bmatrix} (s\widetilde{A} - A)^{-1}(B - s\widetilde{B}) + \begin{bmatrix} D \\ \widetilde{D} \end{bmatrix},$$

with $A, \widetilde{A} \in \mathbf{R}^{n\times n}$, $B, \widetilde{B} \in \mathbf{R}^{n\times m}$, $C, \widetilde{C} \in \mathbf{R}^{l\times n}$, $D, \widetilde{D} \in \mathbf{R}^{l\times m}$ and $s\widetilde{A} - A$ nonsingular for some $s \in \mathbf{C}$, the following two statements are equivalent.

(*i*) $\Phi(j\omega)^*\widetilde{\Phi}(j\omega) + \widetilde{\Phi}(j\omega)^*\Phi(j\omega) \leq 0$ *for all* $\omega \in \mathbf{R}$ *with* $\det(j\omega\widetilde{A} - A) \neq 0$.

(*ii*) *There exists a nonzero pair $(p, P) \in \mathbf{R}\times\mathbf{R}^{n\times n}$ such that $p \geq 0$, $P = P^*$ and*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{bmatrix} + \begin{bmatrix} \widetilde{A} & \widetilde{B} \\ \widetilde{C} & \widetilde{D} \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \leq 0$$

The corresponding equivalence for strict inequalities holds with $p = 1$.

Remark There is another version of the theorem, where the condition in (*i*) is instead that $\Phi(s)^*\widetilde{\Phi}(s) + \widetilde{\Phi}(s)^*\Phi(s) \leq 0$ for all $s \in \mathbf{C}$ with $\operatorname{Re} s \geq 0$ and $\det(s\widetilde{A} - A) \neq 0$. The corresponding change in (*ii*) is the additional condition $P \geq 0$. This version can be proved analogously.

It is instructive to state also the following equivalent theorem, which applies to discrete time systems.

Theorem 2 *With notation from Theorem 1, the following two statements are equivalent.*

(*i*) $\Phi(e^{j\omega})^*\Phi(e^{j\omega}) + \tilde{\Phi}(e^{j\omega})^*\tilde{\Phi}(e^{j\omega}) \leq 0$ *for* $\omega \in \mathbf{R}$ *with* $\det(e^{j\omega}\tilde{A} - A) \neq 0$.

(*ii*) *There exists a nonzero pair* $(p, P) \in \mathbf{R}\times\mathbf{R}^{n\times n}$ *such that* $p \geq 0$, $P = P^*$ *and*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix}^* \begin{bmatrix} P & 0 \\ 0 & pI \end{bmatrix} \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} \leq 0$$

The corresponding equivalence for strict inequalities holds with $p = 1$.

This theorem follows directly from Theorem 1, by substituting $e^{j\omega}$ with $(j\omega + 1)/(j\omega - 1)$ and noting that

$$\begin{aligned} &\Phi\left((j\omega+1)/(j\omega-1)\right) \\ &= C\left((j\omega+1)\tilde{A} - (j\omega-1)A\right)^{-1}\left((j\omega-1)B - (j\omega+1)\tilde{B}\right) + D \\ &= C\left(j\omega(A-\tilde{A}) - (A+\tilde{A})\right)^{-1}\left((B+\tilde{B}) - j\omega(B-\tilde{B})\right) + D. \end{aligned}$$

Alternatively, it can be proved analogously to Theorem 1.

2 The New Proof

First recall the following standard result, which we prove here for completeness.

Lemma 3 *Let* R *and* S *be matrices of the same size. Then* $RR^* = SS^*$, *if and only if there exists a unitary matrix* U *such that* $R = SU$.

Similarly, $RS^* + SR^* = 0$ *if and only if there exists a unitary matrix* U *such that* $R(I + U) = S(I - U)$.

In particular, given nonzero $r, s \in \mathbf{C}^n$, *there is an* $\omega \in \mathbf{R}$ *such that* $r = j\omega s$, *if and only if* $rs^* + sr^* = 0$.

Proof. Each of the first two statements follows from the other by replacement of R and S with $R + S$ and $S - R$. The third statement follows from the second since $r(1 + U) = s(1 - U) \neq 0$ and $(1 - U)(1 + U)^{-1}$ is purely imaginary when $U^*U = 1$.

It remains to prove the first statement. Let the size of R and S be $k \times l$. Consider first square matrices, i.e. the case $k = l$. Introduce the polar decompositions

$$\begin{aligned} R &= H_R U_R \\ S &= H_S U_S \end{aligned}$$

where H_R and H_S are hermitean and positive semidefinite, while U_R and U_S are unitary. Then

$$H_R = (RR^*)^{1/2} = (SS^*)^{1/2} = H_S$$

so the unitary matrix $U = U_S^* U_R$ satisfies $R = SU$.

The case $k < l$ follows immediately by extending R and S with zero rows to square matrices.

If $k > l$, then let R_1 be a submatrix of R with the same rank, but a minimal number of rows. Let S_1 be defined by the corresponding rows in S_1. Then $R_1 R_1^* = S_1 S_1^*$ and we have proved the existence of a unitary matrix U such that $R_1 = S_1 U$. In fact, since all rows of R and S are linear combinations of the rows in R_1 and S_1, the desired equality $R = SU$ is proved. □

The next lemma is central for the proof.

Lemma 4 *Let M and $\widetilde{M}$ be matrices of the same size. If*

$$0 = \widetilde{M} W M^* + M W \widetilde{M}^*$$

for some $W = W^* \geq 0$, then W has the form $W = \sum_{k=1}^{n+m} w_k {w_k}^*$, where

$$0 = \widetilde{M} w_k {w_k}^* M^* + M w_k {w_k}^* \widetilde{M}^*, \quad w_k \in \mathbf{C}^{n+m} \tag{1}$$

for $k = 1, \ldots, n+m$.

Proof. By Lemma 3, there is a unitary matrix U such that

$$M W^{1/2}(I + U) = \widetilde{M} W^{1/2}(I - U)$$

Being unitary, $U = \sum_{k=1}^{n+m} e^{j\theta_k} u_k {u_k}^*$ where $u_k \in \mathbf{C}^{n+m}$ and $\sum_k u_k u_k^* = I$. With $w_k = W^{1/2} u_k$, we get $\sum_{k=1}^{n+m} w_k {w_k}^* = W$ and

$$\begin{aligned} M w_k (1 + e^{j\theta_k}) &= M W^{1/2}(I + U) u_k \\ &= \widetilde{M} W^{1/2}(I - U) u_k \\ &= \widetilde{M} w_k (1 - e^{j\theta_k}) \end{aligned}$$

for $k = 1, \ldots, n+m$. Another application of Lemma 3 gives (1) and the proof is complete. □

Proof of Theorem 1. The two statements will be connected by a sequence of equivalent reformulations. Introduce

$$M = \begin{bmatrix} A & B \end{bmatrix}, \qquad \widetilde{M} = [\tilde{A} \quad \tilde{B}],$$

$$N = \begin{bmatrix} C & D \end{bmatrix}, \qquad \tilde{N} = [\tilde{C} \quad \tilde{D}].$$

For any $u \in \mathbf{C}^m$, $\omega \in \mathbf{R}$ with $\det(j\omega\tilde{A} - A) \neq 0$, we have by definition

$$\begin{bmatrix} \Phi(j\omega) \\ \tilde{\Phi}(j\omega) \end{bmatrix} u = \begin{bmatrix} N \\ \tilde{N} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

where $x = (j\omega\tilde{A} - A)^{-1}(B - j\omega\tilde{B})u$, or equivalently

$$M \begin{bmatrix} x \\ u \end{bmatrix} = j\omega\widetilde{M} \begin{bmatrix} x \\ u \end{bmatrix}.$$

Hence (i) can be rewritten as

(a) $w^*(\tilde{N}^*N + N^*\tilde{N})w \leq 0$ for all w satisfying $Mw = j\omega\widetilde{M}w$ with $\omega \in \mathbf{R}$ and $\det(j\omega\tilde{A} - A) \neq 0$.

At this point, the condition $\det(j\omega\tilde{A} - A) \neq 0$ does not make any difference. Application of Lemma 3 gives the equivalent statement

(b) The sets

$$\Theta = \left\{ \left(w^*(\tilde{N}^*N + N^*\tilde{N})w,\ \widetilde{M}ww^*M^* + Mww^*\widetilde{M}^* \right) : w \in \mathbf{C}^{n+m} \right\}$$
$$\mathcal{P} = \{(r, 0) : r > 0\}$$

are disjoint.

At this point, we take the main step of the proof, replacing Θ by its convex hull. Indeed, any element of conv Θ has the form

$$(\mathrm{trace}\,[W(\tilde{N}^*N + N^*\tilde{N})], \widetilde{M}WM^* + MW\widetilde{M}^*) \quad W \geq 0$$

If it also belongs to $\mathcal{P}$, then by Lemma 4 it can be written as

$$\sum_k (w_k{}^*(\tilde{N}^*N + N^*\tilde{N})w_k, \widetilde{M}w_kw_k{}^*M^* + Mw_kw_k{}^*\widetilde{M}^*),$$

for some $w_1, \ldots, w_{n+m} \in \mathbf{C}^{n+m}$. At least one term of this sum is in $\mathcal{P} \cap \Theta$. This gives the following reformulation of **(b)** .

(c) The convex hull of Θ does not intersect $\mathcal{P}$.

The non-intersection is equivalent to existence of a hyperplane, separating Θ from $\mathcal{P}$. Equivalently, there is a nonzero functional on $\mathbf{R}\times\mathbf{R}^{n\times n}$, that is non-positive on Θ and non-negative on $\mathcal{P}$. Let this linear functional be defined by the nonzero pair $(p, P) \in \mathbf{R} \times \mathbf{R}^{n\times n}$ and the standard scalar product. Then the non-negativity on $\mathcal{P}$ means that $p \geq 0$ and the non-positivity of Θ that

$$\begin{aligned} 0 &\geq pw^*(\widetilde{N}^*N + N^*\widetilde{N})w + \operatorname{trace}[P(\widetilde{M}ww^*M^* + Mww^*\widetilde{M}^*)] \\ &= w^*\left(p\widetilde{N}^*N + pN^*\widetilde{N} + M^*P\widetilde{M} + \widetilde{M}^*PM\right)w. \end{aligned}$$

This demonstrates the equivalence between **(c)** and *(ii)* .

The argument for strict inequalities is analogous. Furthermore, it is clear that $p \neq 0$ is necessary for strict inequality in *(ii)* . Without restriction, we can therefore let $p = 1$. □

References

[And67] B.D.O. Anderson. A system theory criterion for positive real matrices. *SIAM Journal of Control*, 5:171–182, 1967.

[PD93] A. Packard and J.C. Doyle. The complex structured singular value. *Automatica*, pages 71–109, 1993.

Integral Quadratic Constraints as a Unifying Concept

A. Rantzer

Dept. of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 LUND
SWEDEN

A. Megretski

Dept. of Electrical Engineering
Iowa State University
Ames, IA 50011
USA

33rd CDC, draft

Keywords: Robustness, Stability, Passivity

Abstract

It is demonstrated how a number of widely used tools for stability analysis can be conveniently unified and generalized using integral quadratic constraints (IQC's). A purely IQC based stability theorem is presented, which covers classical small gain conditions with anti-causal multipliers, but gains flexibility by avoiding extended spaces and truncated signals.

1 Introduction

In this paper, the term IQC (Integral Quadratic Constraint) is used for an inequality describing a relation between variables in a dynamical system. The system variables appear in a weighted quadratic frequency integral. Implicitly, IQC's have been present in the control literature ever since the 60s, when passivity and circle criteria were introduced in the stability analysis of nonlinear feedback systems [Zames, 1966, Willems, 1971, Desoer and Vidyasagar, 1975]. For example, positivity of a map from u to y, can be expressed by the IQC $\int_{-\infty}^{\infty} y(j\omega)^* u(j\omega) d\omega \geq 0$.

To enhance flexibility of the approach, so-called *multipliers* were introduced to exploit the fact that several different IQC's may be applicable to the same operator. Most of this classical theory was devoted to scalar feedback systems. This lead to conveniently visualizable stability criteria based on the Nyquist diagram, which was particularly important in times when computers were less accessible.

In the 70-s, IQC's were used (and named so) by Yakubovich to treat the stability problem for systems with complex nonlinearities, including amplitude and frequency modulation systems [Yakubovich, 1967, Yakubovich, 1971, Yakubovich, 1973]. Some new IQC's, unrelated to the passivity or small gain arguments, were introduced, and the so-called S-procedure was applied to the case of multiple IQC's.

An important step towards modern robust control was the introduction of analysis methods which essentially rely on the use of computers. One example is the theory for quadratic stabilization [Leitmann, 1979, Gutman, 1979, Corless and Leitmann, 1981], another is the multiloop generalization of the circle criterion based on D-scaling, [Safonov and Athans, 1981, Doyle, 1982]. At the same time, the H^∞ control methods were introduced for synthesis of robust controllers, see for example [Zames, 1981, Tannenbaum, 1982]. Again the results can be viewed in terms of integral quadratic constraints. Both the search for a Lyapunov function and the search for D-scales can be interpreted as optimization of parameters in an IQC, and optimal design with respect to an IQC leads to H^∞ optimization.

During the last decade, a variety of methods were developed within the area of robust control. As was pointed out in [Megretski, 1993] and further emphasized below, many of them can be reformulated to fall within the framework of IQC's. Two recent developments are particularly noticeable. Firstly, the development of interior point methods for solving systems of Linear Matrix Inequalities (LMI:s) strongly improves the possibilities to analyze complex systems using combinations of different IQC's. Secondly, a series of necessity results for IQC-related stability conditions [Khammash and Pearson, 1991, Shamma, 1992, Megretski and Treil, 1993, Tikku and Poolla, 1993] further emphasize the role of IQC's and their analogs as being the most elementary units of uncertainty description.

2 IQC's for Simple Operators

Let us first define the term IQC more exactly. Suppose the function $\omega \mapsto \Pi(j\omega) = \Pi(j\omega)^* = \overline{\Pi(-j\omega)} \in \mathbf{C}^{2m\times 2m}$ is bounded and measurable. A given bounded causal map $\Delta : \mathbf{L}_2^m \to \mathbf{L}_2^m$ is said to *satisfy the IQC defined by* Π, if

$$\int_{-\infty}^{\infty} \begin{bmatrix} \widehat{u}(j\omega) \\ \widehat{v}(j\omega) \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} \widehat{u}(j\omega) \\ \widehat{v}(j\omega) \end{bmatrix} d\omega \geq 0$$

for any $\widehat{u}, \widehat{v}$ being the Fourier transforms of $u, v \in \mathbf{L}_2$ with $v = \Delta(u)$.

The purpose of this section is to review a number of IQC's, that are satisfied by elements appearing in models of uncertain systems.

Linear Time-invariant Dynamics

Let Δ be any linear time-invariant operator on $\mathbf{L}_2^m$ with gain (H_∞ norm) less than one. Then Δ satisfies any IQC defined by a matrix function of the form

$$\begin{bmatrix} x(j\omega)I & 0 \\ 0 & -x(j\omega)I \end{bmatrix}$$

where $x(j\omega) \geq 0$ is a bounded measurable function and I denotes the unit matrix of appropriate dimension.

Linear Time-invariant Scalar Dynamics

Suppose Δ is defined by multiplication in frequency domain with a scalar transfer function having H_∞-norm less than one. Then, Δ satisfies IQC's defined by matrix functions of the form

$$\begin{bmatrix} X(j\omega) & 0 \\ 0 & -X(j\omega) \end{bmatrix}$$

where $X(j\omega) = X(j\omega)^* \geq 0$ is a bounded measurable matrix function.

Constant Real Scalar

If Δ is defined by multiplication with a real number, then it satisfies IQC's defined by matrix functions of the form

$$\begin{bmatrix} 0 & Y(j\omega) \\ Y(j\omega)^* & 0 \end{bmatrix}$$

where $Y(j\omega) = -Y(j\omega)^*$ is bounded and measurable.

These first three types of IQC's are the basis for standard upper bounds for structured singular values [Fan *et al.*, 1991].

Time-varying Real Scalar

Let Δ be defined by multiplication in the time-domain with a scalar function $\delta \in \mathbf{L}_\infty$ with $\|\delta\|_\infty \leq 1$. Then Δ satisfies IQC's defined by a matrix of the form

$$\begin{bmatrix} X & Y \\ Y^T & -X \end{bmatrix}$$

where $X = X^T \geq 0$ and $Y = -Y^T$ are real matrices [Feron, 1994].

Periodic Real Scalar

Let Δ be defined by multiplication in the time-domain with a periodic scalar function $\delta \in \mathbf{L}_\infty$ with $\|\delta\|_\infty \leq 1$ and period T. Then Δ satisfies IQC's defined by matrix functions of the form

$$\begin{bmatrix} X(j\omega) & Y(j\omega) \\ Y(j\omega)^* & -X(j\omega) \end{bmatrix}$$

where X and Y are bounded, measurable matrix functions satisfying

$$\begin{aligned} X(j\omega) &= X(j(\omega + 2\pi/T)) = X(j\omega)^* \geq 0 \\ Y(j\omega) &= Y(j(\omega + 2\pi/T)) = -Y(j\omega)^*. \end{aligned}$$

This set of IQC's can be used to prove the result by Willems on stability of systems with uncertain periodic gains [Willems, 1971].

Memoryless Nonlinearity

Suppose Δ operates on scalar signals according to the nonlinear map $v(t) = \delta(t, u(t))$, where δ is a time-varying function on $\mathbf{R}$ satisfying

$$\alpha x^2 \leq x\delta(t,x) \leq x^2/\beta$$

for some constants $\alpha \leq 0 \leq \beta$. Then Δ satisfies the IQC defined by the 2×2 matrix

$$\begin{bmatrix} -2\alpha & \alpha\beta + 1 \\ \alpha\beta + 1 & -2\beta \end{bmatrix}.$$

The Sign Function

If $v = \Delta(u)$ is defined by the sign function, $v_i(t) = 1$ if $u_i(t) \geq 0$, else $v_i(t) = -1$, then Δ satisfies the IQC defined by

$$\begin{bmatrix} 0 & I + H(j\omega) \\ I + H(j\omega)^* & 0 \end{bmatrix},$$

where H is an arbitrary transfer function with $\|H\|_{L_1} \leq 1$. (A proof will be included in the final version of the paper.)

Although the collection of IQC's presented in this section is far from being complete, it supports the idea that many important properties can be characterized by IQC's.

3 IQC's for Signals

Performance of a linear control system is often measured in terms of disturbance attenuation. An important issue is then the definition of the set of expected disturbances. Here again, integral quadratic constraints can be used as a flexible tool.

A signal $f \in \mathbf{L}_2$ is said to satisfy the IQC defined by Π if

$$\int_{-\infty}^{\infty} \hat{f}(j\omega)^* \Pi(j\omega) \hat{f}(j\omega) d\omega \geq 0.$$

Next follows some examples of signal properties that can be described by IQC's.

Bounds on Auto Correlation

A bound on the auto correlation of f can be expressed as

$$\int_{-\infty}^{\infty} f(t)^* f(t-T)dt \leq \alpha \int_{-\infty}^{\infty} f(t)^* f(t)dt,$$

where $\alpha \leq 1$. This IQC is defined by $\Pi(j\omega) = 2\alpha - e^{j\omega T} - e^{-j\omega T}$.

Dominant Harmonics

The integral

$$\int_a^b |\widehat{f}(j\omega)|^2 d\omega$$

measures intuitively how much energy of f that is concentrated on the spectral interval $[a, b]$. Therefore, the assumption that f is dominated to a certain degree by harmonics in the interval $[a, b]$ may be expressed by an IQC defined by $\Pi(j\omega) = \rho(j\omega)I$, being small positive for $\pm\omega \in [a, b]$ and large negative on the rest of spectrum.

Finite Sets of Signals

In some applications, f is considered as a reference signal from a given finite set $\mathcal{F} = \{f_s\}_{s=1}^N$ of functions. Then a set of IQC's describing $\mathcal{F}$ can be found by solving the system of linear inequalities

$$\int_{-\infty}^{\infty} \widehat{f}_k(j\omega)^* \Pi(j\omega) \widehat{f}_k(j\omega) d\omega \geq 0 \quad k = 1, \ldots, N.$$

4 A General Stability Condition

It will now be demonstrated how IQC's can be applied in stability theory of feedback systems. For most dynamical systems considered in the theory of robust control, any reasonable type of stability is equivalent to invertibility of an operator. The invertibility is often implied by some quadratic inequalities relating its input and output. For example, the operator $I + D$ on $\mathbf{L}_2$, where I denotes identity, is invertible if D is contractive, i.e. $||Df||^2 \leq (1-\epsilon)||f||^2$ for $f \in \mathbf{L}_2$ and some $\epsilon > 0$. The main idea of this section is to combine many IQC's obtained from various sources to get a quadratic inequality which implies invertibility and therefore stability.

Consider the feedback system

$$\begin{cases} v &= G(s)u + e \\ u &= \Delta(v) + f, \end{cases} \tag{1}$$

where G is a stable strictly proper linear transfer function and Δ is a bounded nonlinear operator on $\mathbf{L}_2^m$. The feedback system is said to be $\mathbf{L}_2$-stable if there is a constant $C > 0$ such that for every $e, f \in \mathbf{L}_2^m$, the equations have unique solutions $u, v \in \mathbf{L}_2^m$ and

$$\int_{-\infty}^{\infty} |u|^2 + |v|^2 dt \leq C \int_{-\infty}^{\infty} |e|^2 + |f|^2 dt.$$

The following theorem covers a large number of small gain and passivity type theorems, even with so called anti-causal multipliers. However, unlike most formulations of such results, it is purely based on IQC's and no signal truncations are involved. This allows for more flexibility than the traditional results.

Theorem 1 *Let $\Pi_1, \ldots, \Pi_m$ be bounded measurable matrix functions, that are Hermitean and indefinite on the imaginary axis and satisfy $[I \quad 0]\Pi_i(j\omega)[I \quad 0]^T \geq 0$ and $[0 \quad I]\Pi_i(j\omega)[0 \quad I]^T \leq 0$. Then the feedback system (1) is L_2-stable for all bounded continuous causal operators Δ satisfying the IQC's defined by $\Pi_1, \ldots, \Pi_m$, if and only if $\exists x_1, \ldots, x_m \geq 0, \epsilon > 0$ such that $\forall \omega \in [0, \infty]$*

$$\begin{bmatrix} G(j\omega) \\ I \end{bmatrix}^* \left[\sum_{i=1}^{m} x_i \Pi_i(j\omega)\right] \begin{bmatrix} G(j\omega) \\ I \end{bmatrix} < -\epsilon G(j\omega)^* G(j\omega) \tag{2}$$

Proof outline. Let us first consider the sufficiency. Suppose Δ is an operator satisfying the IQC's defined by $\Pi_1, \ldots, \Pi_m$. The conditions on Π_i are sufficient to make sure that for any $\tau \in [0,1]$, also the operator $\tau\Delta$ satisfies the IQC's. In particular, the IQC defined by $\Pi = \sum_i x_i \Pi_i$ is satisfied.

We will prove that the operator $I - \tau\Delta G$ has a bounded inverse for $\tau \in [0,1]$. The inequality (2) and the boundedness of Π gives straightforwardly that there exists an $\epsilon > 0$ such that

$$||u - \tau\Delta(Gu)|| \geq \epsilon ||Gu|| \quad \text{for any } u \in \mathbf{L}_2, \tau \in [0,1]. \tag{3}$$

This proves injectivity of the operator. The idea is now to prove that the set of τ-values with surjectivity, is both closed and open in $[0, 1]$. This completes the sufficiency proof, because the invertibility, is trivial for $\tau = 0$. Closedness follows from the compactness of the operator $P_T \Delta G$, where P_T means truncation at time T. Openness is shown by application of the Schauder-Tichonov fixed point theorem. The causality of G and Δ is crucial, both for closedness and openness.

The necessity part is essentially analogous to the necessity proof of the multi-loop circle criterion in [Megretski and Treil, 1991, Shamma, 1992]. □

5 Computations Based on LMI:s

In practical applications of Theorem 1, it is useful to introduce state space representations of $\Pi_1, \ldots, \Pi_m$ and G. The search for $x_1, \ldots, x_m$ can then be converted using the Kalman-Yakubovich-Popov Lemma into a convex optimization problem defined by linear matrix inequalities (LMI:s). For such problems there has recently been a strong development of numerical algorithms based on interior point methods [Nesterov and Nemirovski, 1993, Boyd *et al.*, 1993].

To concretize this idea, let G have the state space representation

$$G(j\omega) = C(j\omega I - A)^{-1}B + D$$

and let $\Pi_i = \widetilde{\Phi}_i^* \Phi_i + \Phi_i^* \widetilde{\Phi}_i$, where Φ_i and $\widetilde{\Phi}_i$ have the state space representations

$$\begin{aligned} \Phi_i(j\omega) &= G_i(j\omega I - E)^{-1}F + H_i \\ \widetilde{\Phi}_i(j\omega) &= \widetilde{G}_i(j\omega I - E)^{-1}F + \widetilde{H}_i \end{aligned}$$

In fact, any rational matrix that is hermitean and bounded on the imaginary axis can be written this way with $\widetilde{\Phi}_i = I$. However, in parametrizing a class of Π_i:s, the flexibility in $\widetilde{\Phi}_i$ is sometimes useful.

Finally, introduce the notation

$$K = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & I & 0 \\ \hline C & 0 & D \\ 0 & 0 & I \end{array}\right] \qquad \widetilde{K} = \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline C & 0 & D \\ 0 & 0 & I \end{array}\right]$$

$$L_i = \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & E & F \\ \hline 0 & G_i & H_i \end{array}\right] \qquad \widetilde{L}_i = \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & \widetilde{G}_i & \widetilde{H}_i \end{array}\right]$$

for $i = 1, \ldots, m$. Then the following proposition is useful for the application of Theorem 1.

Proposition 2 *Given $x_1, \ldots, x_m \geq 0$, the inequality (2) is equivalent to existence of a symmetric matrix P such that*

$$\sum_{i=1}^{m} \left\{ \widetilde{K}^T \widetilde{L}_i^T \left[\begin{array}{c|c} P & 0 \\ \hline 0 & x_i I \end{array}\right] L_i K + K^T L_i^T \left[\begin{array}{c|c} P & 0 \\ \hline 0 & x_i I \end{array}\right] \widetilde{L}_i \widetilde{K} \right\} < 0.$$

Note that the inequality is linear in $(P, x_1, \ldots, x_m)$.

5.1 Example: LTI Uncertainty with Nonlinear Gain

to be written...

6 Acknowledgement

The authors are greatful to a large number of people, for comments and suggestions about this work. In particular, thanks to John Doyle for stimulating discussions and continuous encouragement.

Boyd, S., L. E. Ghaoui, E. Feron, and V. Balakrishnan. (1993):. "Linear matrix inequalities in systems and control theory.". June 1993 draft.

Corless, M. and G. Leitmann (1981): "Continuous state - feedback guaranteeing uniform ultimate boundedness for uncertain dynamic systems." *IEEE Transactions on Automatic Control*, **26:5**.

Desoer, C. and M. Vidyasagar (1975): *Feedback Systems: Input-Output Properties*. Academic Press, New York.

Doyle, J. C. (1982): "Analysis of feedback systems with structured uncertainties." In *IEE Proceedings*, volume D-129, pp. 242–251.

Fan, M., A. Tits, and J. Doyle (1991): "Robustness in presence of mixed parametric uncertainty and unmodeled dynamics." *IEEE Transactions on Automatic Control*, **36:1**, pp. 25–38.

Feron, E. (1994):. "Improving on Scalings in Quadratic Stability of Systems.". Submitted as a short paper for the 1994 ACC.

Gutman, S. (1979): "Uncertain dynamical systems - a Lyapunov min-max approach." *IEEE Transactions on Automatic Control*, **24:3**.

Khammash, M. and J. Pearson (1991): "Performance robustness of discrete-time systems with structured uncertainty." *IEEE Transactions on Automatic Control*, **36:4**, pp. 398–412.

Leitmann, G. (1979): "Guaranteed asymptotic stability for some linear systems with bounded uncertainties." *Journal of Dynamic Systems, measurement, and Control*, **101:3**.

Megretski, A. (1993): "Power distribution approach in robust control." In *Proceedings of IFAC Congress*.

Megretski, A. and S. Treil (1991): "S-procedure and power distribution inequalities: A new method in optimization and robustness of uncertain systems." Technical Report 1, Mittag-Leffler Institute. Submitted to Journal of Mathematical Systems, Estimation and Control.

Megretski, A. and S. Treil (1993): "Power distribution inequalities in optimization and robustness of uncertain systems." *Journal of Mathematical Systems, Estimation and Control*, **3:3**, pp. 301–319.

Nesterov, Y. and A. Nemirovski (1993): *Interior point polynomial methods in convex programming*. SIAM books.

Safonov, M. (1980): *Stability and Robustness of Multivariable Feedback Systems*. MIT Press.

Safonov, M. and M. Athans (1981): "A multi-loop generalization of the circle criterion for stability margin analysis." *IEEE Transactions on Automatic Control*, **26**, pp. 415–422.

Shamma, J. (1992): "Robustness analysis for time-varying systems." In *Proc. of 31st IEEE Conference on Decision and Control*.

Tannenbaum, A. (1982): "Modified Nevanlinna-Pick interpolation of linear plants with uncertainty in the gain factor." *Int. Journal of Control*, **36**, pp. 331–336.

Tikku, A. and K. Poolla (1993): "Robust performance against slowly time-varying structured perturbations." In *Proceedings of 32nd IEEE Conference on Decision and Control*, pp. 990–995.

Willems, J. (1971): *The Analysis of Feedback Systems.* MIT Press.

Willems, J. (1972): "Dissipative dynamical systems, part i: General theory; part ii: Linear systems with quadratic supply rates." *Arch. Rational Mechanics and Analysis*, **45:5**, pp. 321–393.

Yakubovich, V. (1967): "Frequency conditions for the absolute stability of control systems with several nonlinear or linear nonstationary units." *Autom. Telemech.*, pp. 5–30.

Yakubovich, V. (1971): "S-procedure in nonlinear control theory." *Vestnic Leningrad University, Matematika.*

Yakubovich, V. (1973): "Minimization of quadratic functionals under the quadratic constraints and the necessity of a frequency condition in the quadratic criterium for absolute stability of nonlinear control systems." *Soviet Mathematical Doklady*, **14**, pp. 593–597.

Zames, G. (1966): "On the input-output stability of nonlinear time-varying feedback systems—Part I: Conditions derived using concepts of loop gain." *IEEE Transactions on Automatic Control*, **11**, April, pp. 228–238.

Zames, G. (1981): "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses." *IEEE Transactions on Automatic Control*, **26**, pp. 301–320.

... to be continued...