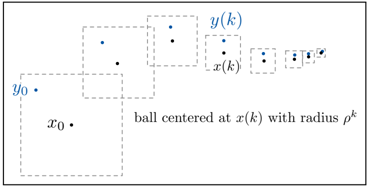
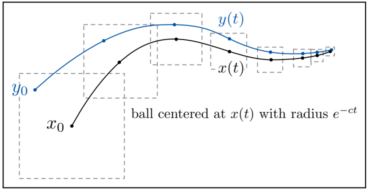
Contraction Theory in Systems and Control

Francesco Bullo

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
http://motion.me.ucsb.edu

2022 IEEE Conference in Decision and Control, Cancun, México



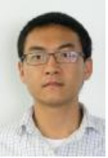


Acknowledgments


Saber Jafarpour
GeorgiaTech


Alexander Davydov
UC Santa Barbara


Kevin D. Smith
UC Santa Barbara


Xiaoming Duan
Shanghai Jiao Tong


Pedro Cisneros-Velarde
University of Illinois


Veronica Centorrino
Scuola Sup Meridionale


Robin Delabays
HES-SO Sion


Anton Proskurnikov
Politecnico Torino


Giovanni Russo
Univ Salerno


John W. Simpson-Porco
University of Toronto

AFOSR ARO DTRA/ERDC

Frederick Leve @AFOSR FA9550-22-1-0059
Edward Palazzolo @ARO W911NF-22-1-0233
Paul Tandy @DTRA W912HZ-22-2-0010
Donald Wagner @AFOSR FA9550-21-1-0203

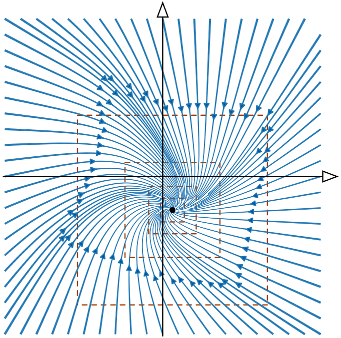
contractivity = robust computationally-friendly stability

fixed point theory + Lyapunov stability theory + geometry of metric spaces

contractivity = robust computationally-friendly stability
fixed point theory + Lyapunov stability theory + geometry of metric spaces

highly-ordered transient and asymptotic behavior:

1. unique globally exponential stable equilibrium
   & two natural Lyapunov functions
2. robustness properties
   - bounded input, bounded output (iss)
   - finite input-state gain
   - robustness margin wrt unmodeled dynamics
   - robustness margin wrt delayed dynamics
3. periodic input, periodic output
4. modularity and interconnection properties
5. accurate numerical integration and equilibrium point computation

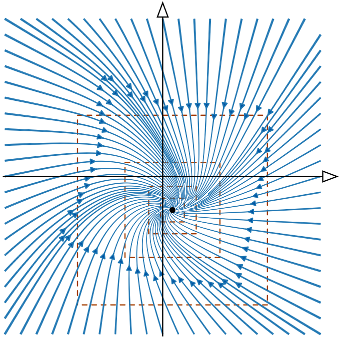


contractivity = robust computationally-friendly stability

fixed point theory + Lyapunov stability theory + geometry of metric spaces

highly-ordered transient and asymptotic behavior:

1. unique globally exponential stable equilibrium
   & two natural Lyapunov functions
2. robustness properties
   - bounded input, bounded output (iss)
   - finite input-state gain
   - robustness margin wrt unmodeled dynamics
   - robustness margin wrt delayed dynamics
3. periodic input, periodic output
4. modularity and interconnection properties
5. accurate numerical integration and equilibrium point computation



search for contraction properties

design engineering systems to be contracting

Contraction theory: historical notes

- **Origins**

  S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3(1):133–181, 1922.

Contraction theory: historical notes

- **Origins**

  S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3(1):133–181, 1922.

- **Dynamics**:

  G. Dahlquist. *Stability and error bounds in the numerical integration of ordinary differential equations*. PhD thesis, (Reprinted in Trans. Royal Inst. of Technology, No. 130, Stockholm, Sweden, 1959), 1958

  S. M. Lozinskii. Error estimate for numerical integration of ordinary differential equations. I. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, 5:52–90, 1958. URL **http://mi.mathnet.ru/eng/ivm2980**. (in Russian)

- **Computation**:

  C. A. Desoer and H. Haneda. The measure of a matrix as a tool to analyze computer algorithms for circuit analysis. *IEEE Transactions on Circuit Theory*, 19(5):480–486, 1972.

Contraction theory: historical notes

- **Origins**

  S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3(1):133–181, 1922.

- **Dynamics**:

  G. Dahlquist. *Stability and error bounds in the numerical integration of ordinary differential equations*. PhD thesis, (Reprinted in Trans. Royal Inst. of Technology, No. 130, Stockholm, Sweden, 1959), 1958

  S. M. Lozinskii. Error estimate for numerical integration of ordinary differential equations. I. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, 5:52–90, 1958. URL **http://mi.mathnet.ru/eng/ivm2980**. (in Russian)

- **Computation**:

  C. A. Desoer and H. Haneda. The measure of a matrix as a tool to analyze computer algorithms for circuit analysis. *IEEE Transactions on Circuit Theory*, 19(5):480–486, 1972.

- **Systems and control**:

  W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6): 683–696, 1998.

- **Incomplete list of contributors who influenced me**
  Aminzare, Arcak, Chung, Coogan, Di Bernardo, Manchester, Margaliot, Pavlov, Pham, Proskurnikov, Russo, Sepulchre, Slotine, Sontag, ...

- **Incomplete list of contributors who influenced me**
  Aminzare, Arcak, Chung, Coogan, Di Bernardo, Manchester, Margaliot, Pavlov, Pham, Proskurnikov, Russo, Sepulchre, Slotine, Sontag, ...

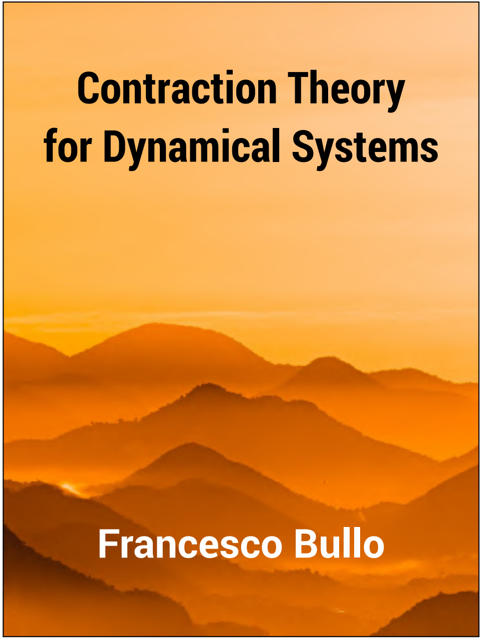
- **Surveys**:

Z. Aminzare and E. D. Sontag. Contraction methods for nonlinear systems: A brief introduction and some open problems. In *IEEE Conf. on Decision and Control*, pages 3835–3847, Dec. 2014.

M. Di Bernardo, D. Fiore, G. Russo, and F. Scafuti. Convergence, consensus and synchronization of complex networks via contraction theory. In *Complex Systems and Networks*. Springer, 2016.

H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine. Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview. *Annual Reviews in Control*, 52:135–169, 2021.

P. Giesl, S. Hafstein, and C. Kawan. Review on contraction analysis and computation of contraction metrics. *Journal of Computational Dynamics*, 10(1):1–47, 2023.



Contraction Theory for Dynamical Systems, Francesco Bullo, KDP, 1.0 edition, 2022, ISBN 979-8836646806

1. Textbook with exercises and answers. Format: textbook, slides, and paperbook
2. Content:
   Fixed point theory
   Theory of contracting dynamics on vector spaces
   Applications to nonlinear and interconnected systems
3. Self-Published and Print-on-Demand at:
   https://www.amazon.com/dp/B0B4K1BTF4
4. PDF Freely available at
   http://motion.me.ucsb.edu/book-ctds
5. 10h minicourse on youtube:
   https://youtu.be/RvR47ZbqJjc
6. Future version to include: systems on Riemannian manifolds, homogeneous spaces, and solid cones

   "Continuous improvement is better than delayed perfection"
   Mark Twain

Outline

Linear algebra: induced norms

Vector norm	Induced matrix norm	Induced matrix log norm
$\|x\|_1 = \sum_{i=1}^n \lvert x_i\rvert$	$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^n \lvert a_{ij}\rvert$	$\mu_1(A) = \max_{j\in\{1,\dots,n\}} \left(a_{jj} + \sum_{i=1,i\neq j}^n \lvert a_{ij}\rvert\right)$ $=$ max column "absolute sum" of A
$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\|A\|_2 = \sqrt{\lambda_{\text{max}}(A^\top A)}$	$\mu_2(A) = \lambda_{\text{max}}\left(\frac{A+A^\top}{2}\right)$
$\|x\|_\infty = \max_{i\in\{1,\dots,n\}} \lvert x_i\rvert$	$\|A\|_\infty = \max_{i\in\{1,\dots,n\}} \sum_{j=1}^n \lvert a_{ij}\rvert$	$\mu_\infty(A) = \max_{i\in\{1,\dots,n\}} \left(a_{ii} + \sum_{j=1,j\neq i}^n \lvert a_{ij}\rvert\right)$ $=$ max row "absolute sum" of A

Discrete-time dynamics and Lipschitz constants

$$x_{k+1} = \mathsf{F}(x_k) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced norm } \|\cdot\|$$

Discrete-time dynamics and Lipschitz constants

$$x_{k+1} = \mathsf{F}(x_k) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced norm } \|\cdot\|$$

Lipschitz constant

$$\begin{aligned}\mathsf{Lip}(\mathsf{F}) &= \inf\{\ell > 0 \text{ such that } \|\mathsf{F}(x) - \mathsf{F}(y)\| \leq \ell\|x - y\| \quad \text{for all } x, y\} \\ &= \sup_x \|\mathsf{J}_{\mathsf{F}}(x)\|\end{aligned}$$

For **scalar map** f, $\mathsf{Lip}(f) = \sup_x |f'(x)|$

Discrete-time dynamics and Lipschitz constants

$$x_{k+1} = \mathsf{F}(x_k) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced norm } \|\cdot\|$$

Lipschitz constant

$$\begin{aligned}\mathsf{Lip}(\mathsf{F}) &= \inf\{\ell > 0 \text{ such that } \|\mathsf{F}(x) - \mathsf{F}(y)\| \leq \ell \|x - y\| \quad \text{for all } x, y\} \\ &= \sup_x \|\mathsf{J}_{\mathsf{F}}(x)\|\end{aligned}$$

For **scalar map** f, $\quad \mathsf{Lip}(f) = \sup_x |f'(x)|$

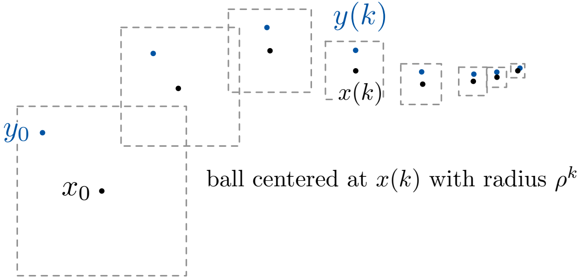
For **affine map** $\mathsf{F}_A(x) = Ax + a$

$$\begin{aligned}\|x\|_{2,P} &= (x^\top P x)^{1/2} & \mathsf{Lip}_{2,P}(\mathsf{F}_A) &= \|A\|_{2,P} \leq \ell & \iff \quad & A^\top P A \preceq \ell^2 P \\ \|x\|_{\infty,\eta} &= \max_i |x_i|/\eta_i & \mathsf{Lip}_{\infty,\eta}(\mathsf{F}_A) &= \|A\|_{\infty,\eta} \leq \ell & \iff \quad & \eta^\top |A| \leq \ell \eta^\top\end{aligned}$$

Banach contraction theorem for discrete-time dynamics:

If $\rho := \text{Lip}(F) < 1$, then

1. F is **contracting** = distance between trajectories decreases exp fast (ρ^k)
2. F has a unique, glob exp stable equilibrium x^*



From discrete to continuous time

The **induced log norm** of $A \in \mathbb{R}^{n \times n}$ wrt to $\|\cdot\|$:

$$\mu(A) := \lim_{h \to 0^+} \frac{\|I_n + hA\| - 1}{h}$$

subadditivity: $\mu(A + B) \leq \mu(A) + \mu(B)$

scaling: $\mu(bA) = b\mu(A), \quad \forall b \geq 0$

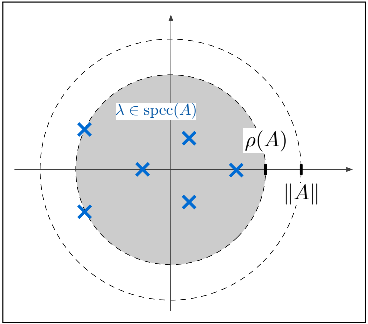
From discrete to continuous time

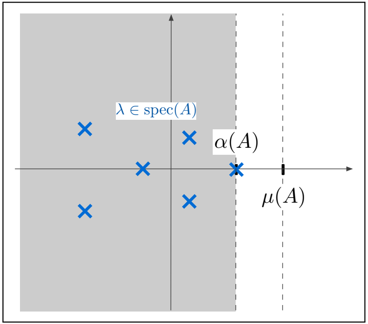
The **induced log norm** of $A \in \mathbb{R}^{n\times n}$ wrt to $\|\cdot\|$:

$$\mu(A) := \lim_{h\to 0^+} \frac{\|I_n + hA\| - 1}{h}$$

subadditivity: $\mu(A+B) \le \mu(A) + \mu(B)$

scaling: $\mu(bA) = b\mu(A), \quad \forall b \ge 0$





Example induced log norms

Vector norm	Induced matrix norm	Induced matrix log norm
$\|x\|_1 = \sum_{i=1}^n \|x_i\|$	$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^n \|a_{ij}\|$	$\mu_1(A) = \max_{j\in\{1,\dots,n\}} \Big(a_{jj} + \sum_{i=1, i\neq j}^n \|a_{ij}\|\Big)$ $=$ max column "absolute sum" of A
$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$	$\|A\|_2 = \sqrt{\lambda_{\mathsf{max}}(A^\top A)}$	$\mu_2(A) = \lambda_{\mathsf{max}}\Big(\frac{A+A^\top}{2}\Big)$
$\|x\|_\infty = \max_{i\in\{1,\dots,n\}} \|x_i\|$	$\|A\|_\infty = \max_{i\in\{1,\dots,n\}} \sum_{j=1}^n \|a_{ij}\|$	$\mu_\infty(A) = \max_{i\in\{1,\dots,n\}} \Big(a_{ii} + \sum_{j=1, j\neq i}^n \|a_{ij}\|\Big)$ $=$ max row "absolute sum" of A

Continuous-time dynamics and one-sided Lipschitz constants

$\dot{x} = \mathsf{F}(x)$ on $\mathbb{R}^n$ with norm $\|\cdot\|$ and induced log norm $\mu(\cdot)$

Continuous-time dynamics and one-sided Lipschitz constants

$$\dot{x} = \mathsf{F}(x) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced log norm } \mu(\cdot)$$

One-sided Lipschitz constant

$$\begin{aligned}\mathrm{osLip}(\mathsf{F}) &= \inf\{b \in \mathbb{R} \text{ such that } \langle\!\langle \mathsf{F}(x) - \mathsf{F}(y), x - y \rangle\!\rangle \leq b\|x - y\|^2 \quad \text{for all } x, y\} \\ &= \sup_x \mu(\mathsf{J}_{\mathsf{F}}(x))\end{aligned}$$

For **scalar map** f, $\quad \mathrm{osLip}(f) = \sup_x f'(x)$

Continuous-time dynamics and one-sided Lipschitz constants

$$\dot{x} = \mathsf{F}(x) \qquad \text{on } \mathbb{R}^n \text{ with norm } \|\cdot\| \text{ and induced log norm } \mu(\cdot)$$

One-sided Lipschitz constant

$$\begin{aligned} \mathrm{osLip}(\mathsf{F}) &= \inf\{b \in \mathbb{R} \text{ such that } \langle\langle \mathsf{F}(x) - \mathsf{F}(y), x - y \rangle\rangle \leq b\|x-y\|^2 \quad \text{for all } x, y\} \\ &= \sup_x \mu(\mathsf{J}_{\mathsf{F}}(x)) \end{aligned}$$

For **scalar map** f, $\quad \mathrm{osLip}(f) = \sup_x f'(x)$

For **affine map** $\mathsf{F}_A(x) = Ax + a$

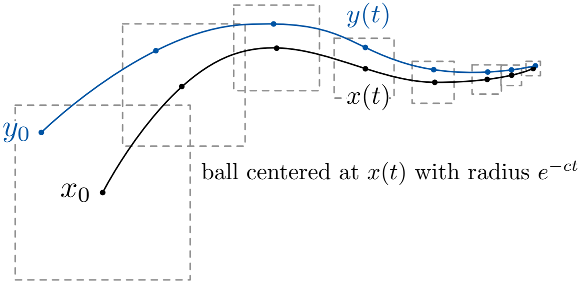
$$\mathrm{osLip}_{2,P}(\mathsf{F}_A) = \mu_{2,P}(A) \leq \ell \qquad \iff \qquad A^\top P + AP \preceq 2\ell P$$

$$\mathrm{osLip}_{\infty,\eta}(\mathsf{F}_A) = \mu_{\infty,\eta}(A) \leq \ell \qquad \iff \qquad a_{ii} + \sum_{j \neq i} |a_{ij}| \eta_i / \eta_j \leq \ell$$

Banach contraction theorem for continuous-time dynamics:

If $-c := \mathrm{osLip}(\mathsf{F}) < 0$, then

1. F is **infinitesimally contracting** = distance between trajectories decreases exp fast (e^{-ct})
2. F has a unique, glob exp stable equilibrium x^*



From inner products to weak pairings

$$\frac{1}{2}\frac{d}{dt}\|x(t)\|_2^2 \;=\; \dot{x}^\top x \;=\; \langle\!\langle \dot{x}, x \rangle\!\rangle$$

$$\implies \qquad \tfrac{1}{2}D^+\|x(t)\|^2 \;=:\; [\![\dot{x}, x]\!]$$

- D^+ is upper-right Dini derivative

From inner products to weak pairings

$$\frac{1}{2}\frac{d}{dt}\|x(t)\|_2^2 \;=\; \dot{x}^\top x \;=\; \langle\langle \dot{x}, x\rangle\rangle$$

$$\implies \qquad \tfrac{1}{2}D^+\|x(t)\|^2 \;=:\; [\![\dot{x}, x]\!]$$

- D^+ is upper-right Dini derivative
- **weak pairing** $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ exists for each norm, i.e.,

$$[\![y, x]\!]_1 \;:=\; \|x\|_1 \operatorname{sign}(x)^\top y \qquad \text{(sign pairing)}$$

$$[\![y, x]\!]_\infty \;:=\; \max_{i \in \mathcal{A}_\infty(x)} x_i y_i \qquad \text{for } \mathcal{A}_\infty(x) = \{i \mid |x_i| = \|x\|_\infty\} \qquad \text{(max pairing)}$$

From inner products to weak pairings

$$\tfrac{1}{2}\frac{d}{dt}\|x(t)\|_2^2 \;=\; \dot{x}^\top x \;=\; \langle\!\langle \dot{x}, x \rangle\!\rangle$$

$$\implies \qquad \tfrac{1}{2}D^+\|x(t)\|^2 \;=:\; [\![\dot{x}, x]\!]$$

- D^+ is upper-right Dini derivative
- **weak pairing** $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ exists for each norm, i.e.,

$$[\![y, x]\!]_1 \;:=\; \|x\|_1 \operatorname{sign}(x)^\top y \qquad \textbf{(sign pairing)}$$

$$[\![y, x]\!]_\infty \;:=\; \max_{i \in \mathcal{A}_\infty(x)} x_i y_i \qquad \text{for } \mathcal{A}_\infty(x) = \{i \mid |x_i| = \|x\|_\infty\} \qquad \textbf{(max pairing)}$$

theory of weak pairings: computational properties and applications to monotone operators

Log norm bounds	Demidovich conditions	One-sided Lipschitz conditions
$\mu_{2,P}(\mathsf{J_F}(x)) \leq -c$	$P\mathsf{J_F}(x) + \mathsf{J_F}(x)^\top P \preceq -2cP$	$(x-y)^\top P\big(\mathsf{F}(x) - \mathsf{F}(y)\big) \leq -c\\|x-y\\|_{P^{1/2}}^2$
$\mu_1(\mathsf{J_F}(x)) \leq -c$	$\operatorname{sign}(v)^\top \mathsf{J_F}(x)v \leq -c\\|v\\|_1$	$\operatorname{sign}(x-y)^\top(\mathsf{F}(x) - \mathsf{F}(y)) \leq -c\\|x-y\\|_1$
$\mu_\infty(\mathsf{J_F}(x)) \leq -c$	$\max_{i \in \mathcal{A}_\infty(v)} v_i\,(\mathsf{J_F}(x)v)_i \leq -c\\|v\\|_\infty^2$	$\max_{i \in \mathcal{A}_\infty(x-y)} (x_i-y_i)(\mathsf{F}_i(x)-\mathsf{F}_i(y)) \leq -c\\|x-y\\|_\infty^2$

Each row = three equivalent statements. To be understood for all $x, y \in \mathbb{R}^n$ and all $v \in \mathbb{R}^n$.

One sided Lipschitz conditions

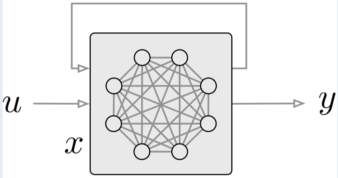
1. **simple sufficient condition** for uniqueness of continuous ODEs in: A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Kluwer, 1988. ISBN 902772699X (Chapter 1, page 5, citing Krasnosel'skii and Krein 1955)
2. **one-sided Lipschitz maps** in: E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer, 1993. (Section I.10)

3. **uniformly decreasing maps** in: L. Chua and D. Green. A qualitative analysis of the behavior of dynamic nonlinear networks: Stability of autonomous networks. *IEEE Transactions on Circuits and Systems*, 23(6):355–379, 1976.
4. **maps with negative nonlinear measure** in: H. Qiao, J. Peng, and Z.-B. Xu. Nonlinear measures: A new approach to exponential stability analysis for Hopfield-type neural networks. *IEEE Transactions on Neural Networks*, 12(2): 360–370, 2001.
5. **dissipative Lipschitz maps** in: T. Caraballo and P. E. Kloeden. The persistence of synchronization under environmental noise. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461 (2059):2257–2267, 2005.
6. **maps with negative lub log Lipschitz constant** in: G. Söderlind. The logarithmic norm. History and modern theory. *BIT Numerical Mathematics*, 46(3):631–652, 2006.
7. **QUAD maps** in: W. Lu and T. Chen. New approach to synchronization analysis of linearly coupled ordinary differential systems. *Physica D: Nonlinear Phenomena*, 213(2):214–230, 2006.
8. **incremental quadratically stable maps** in: L. D'Alto and M. Corless. Incremental quadratic stability. *Numerical Algebra, Control and Optimization*, 3:175–201, 2013.

Advantages of non-Euclidean approaches

1. *well suited for certain class of systems*
   ℓ_1 for monotone flow systems
2. *computational advantages*
   ℓ_1/ℓ_∞ constraints lead to LPs, whereas ℓ_2 constraints leads to LMIs
3. *robustness to structural perturbations*
   ℓ_1/ℓ_∞ contractions are connectively robust (i.e., edge removal)
4. *adversarial input-output analysis*
   ℓ_∞ better suited for the analysis of adversarial examples than ℓ_2
5. *asynchronous distributed computation*
   ℓ_∞ contractions converge under fully asynchronous distributed execution

Application: ℓ_∞-contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b) \qquad (\textit{recurrent NN})$$

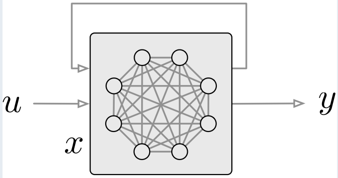
$$x = \Phi(Ax + Bu + b) \qquad (\textit{implicit NN})$$

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) \qquad (\textit{forward Euler})$$

If

$$\mu_\infty(A) < 1 \qquad \left(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\right)$$

Application: ℓ_∞-contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b) \qquad \text{(recurrent NN)}$$

$$x = \Phi(Ax + Bu + b) \qquad \text{(implicit NN)}$$

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) \qquad \text{(forward Euler)}$$

If

$$\mu_\infty(A) < 1 \qquad \left(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\right)$$

- recurrent NN is contracting with rate $1 - \mu_\infty(A)_+$
- implicit NN is well posed
- forward Euler is contracting with factor $1 - \dfrac{1-\mu_\infty(A)_+}{1-\min_i(a_{ii})_-}$ at $\alpha = \dfrac{1}{1-\min_i(a_{ii})_-}$

Detour: convexity and fixed point theory

For differentiable $V : \mathbb{R}^n \to \mathbb{R}$, equivalent statements:

1. V is **strongly convex** with parameter m
2. $-\,\mathrm{grad} V$ is m-**strongly infinitesimally contracting** with respect to $\| \cdot \|_2$

Forward Euler theorem for contracting dynamics

Given arbitrary norm $\|\cdot\|$, equivalent statements

1. $\dot{x} = \mathsf{F}(x)$ is infinitesimally contracting
2. there exists $\alpha > 0$ such that $x_{k+1} = x_k + \alpha \mathsf{F}(x_k)$ is contracting

Forward Euler theorem for contracting dynamics

Given arbitrary norm $\|\cdot\|$, equivalent statements

1. $\dot{x} = \mathsf{F}(x)$ is infinitesimally contracting
2. there exists $\alpha > 0$ such that $x_{k+1} = x_k + \alpha \mathsf{F}(x_k)$ is contracting

Given *contraction rate* c and *Lipschitz constant* ℓ, define *condition number* $\kappa = \dfrac{\ell}{c} \geq 1$

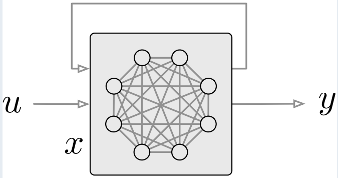
1. $\mathsf{Id} + \alpha \mathsf{F}$ is contracting for

$$0 < \alpha < \frac{1}{c\kappa(1+\kappa)}$$

2. the optimal step size minimizing and minimum contraction factor:

$$\alpha^* = \frac{1}{c}\Big(\frac{1}{2\kappa^2} - \frac{3}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)\Big)$$
$$\ell^* = 1 - \frac{1}{4\kappa^2} + \frac{1}{8\kappa^3} + \mathcal{O}\Big(\frac{1}{\kappa^4}\Big)$$

Application: ℓ_∞-contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b) \qquad \text{(*recurrent NN*)}$$

$$x = \Phi(Ax + Bu + b) \qquad \text{(*implicit NN*)}$$

$$x_{k+1} = (1-\alpha)x_k + \alpha\Phi(Ax_k + Bu + b) \qquad \text{(*forward Euler*)}$$

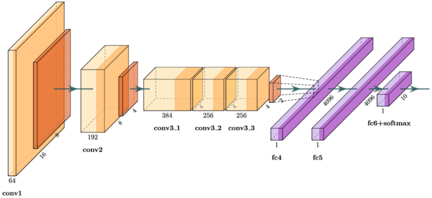
If

$$\mu_\infty(A) < 1 \qquad \Big(\text{i.e., } a_{ii} + \sum_j |a_{ij}| < 1 \text{ for all } i\Big)$$

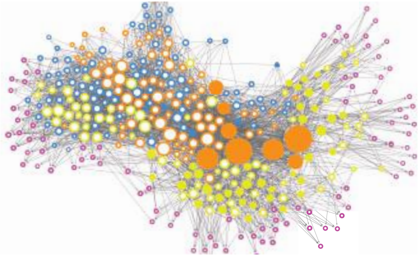
- recurrent NN is contracting with rate $1 - \mu_\infty(A)_+$
- implicit NN is well posed
- forward Euler is contracting with factor $1 - \dfrac{1-\mu_\infty(A)_+}{1-\min_i(a_{ii})_-}$ at $\alpha^* = \dfrac{1}{1-\min_i(a_{ii})_-}$

Motivation: ℓ_∞-contracting neural networks

While most ML architectures are feedforward,
biological neural networks are recurrent and recent interest for implicit ML architectures



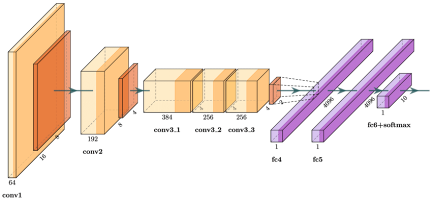
artificial neural network AlexNet ’12



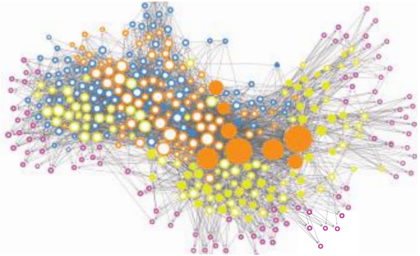
C. elegans connectome ’17

Motivation: ℓ_∞-contracting neural networks

While most ML architectures are feedforward,
biological neural networks are recurrent and recent interest for implicit ML architectures



artificial neural network AlexNet '12



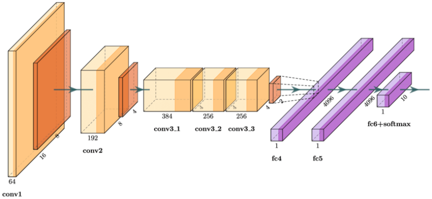
C. elegans connectome '17

For recurrent NN, ℓ_∞-contractivity characterizes the synaptic weights to ensure:

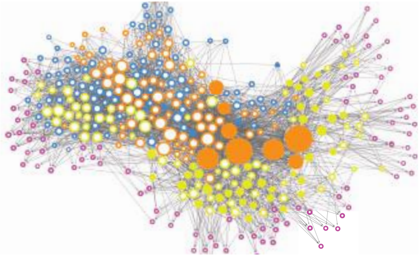
- reproducible & robust behavior
- highly-ordered transient+asymptotic dynamic behavior
- efficient computational methods

Motivation: ℓ_∞-contracting neural networks

While most ML architectures are feedforward,
biological neural networks are recurrent and recent interest for implicit ML architectures



artificial neural network AlexNet ’12



C. elegans connectome ’17

For recurrent NN, ℓ_∞-contractivity characterizes the synaptic weights to ensure:

- reproducible & robust behavior
- highly-ordered transient+asymptotic dynamic behavior
- efficient computational methods

A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 2012

G. Yan, P. E. Vértes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási. Network control principles predict neuron function in the Caenorhabditis elegans connectome. *Nature*, 550(7677):519–523, 2017.

Outline

$\#1$: From closed to open systems

Incremental ISS and input-state gain

Given normed spaces $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$, consider

$$\dot{x} = \mathsf{F}(x, u(t)), \qquad x_0 \in \mathcal{X}, \qquad u(t) \in \mathcal{U}$$

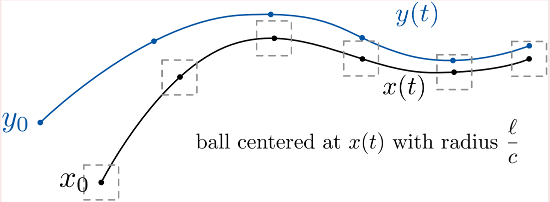
Assume:

- **contractivity wrt x**: $\quad \mathrm{osLip}_x(\mathsf{F}) \leq -c < 0,\quad$ uniformly in u
- **Lipschitz wrt u**: $\quad \mathrm{Lip}_u(\mathsf{F}) \leq \ell,\quad$ uniformly in x

Then

1. any soltns: $x(t)$ with input u_x and $y(t)$ with input u_y

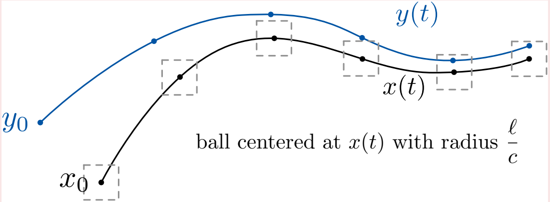
$$D^+\|x(t)-y(t)\|_{\mathcal{X}} \leq -c\|x(t)-y(t)\|_{\mathcal{X}} + \ell\|u_x(t)-u_y(t)\|_{\mathcal{U}}$$



Then

1. any soltns: $x(t)$ with input u_x and $y(t)$ with input u_y

$$D^+\|x(t)-y(t)\|_{\mathcal{X}} \;\le\; -c\|x(t)-y(t)\|_{\mathcal{X}} \;+\; \ell\|u_x(t)-u_y(t)\|_{\mathcal{U}}$$



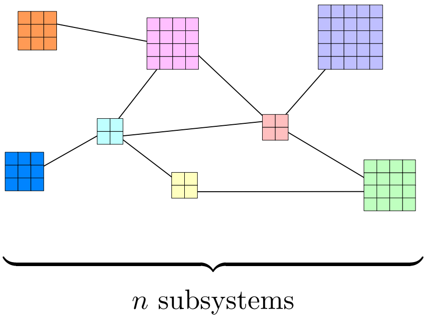
2. F is **incrementally ISS**, that is, for all x_0, y_0

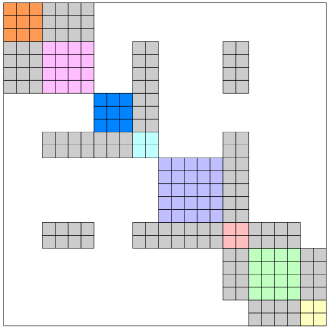
$$\|x(t)-y(t)\|_{\mathcal{X}} \;\le\; e^{-ct}\|x_0-y_0\|_{\mathcal{X}} \;+\; \frac{\ell(1-e^{-ct})}{c}\sup_{\tau\in[0,t]}\|u_x(\tau)-u_y(\tau)\|_{\mathcal{U}}$$

Consider n interconnected subsystems

$$\dot{x}_i = \mathsf{F}_i(x_i, x_{-i}), \qquad \text{for } i \in \{1, \dots, n\}$$

with state $x_i \in \mathbb{R}^{N_i}$
with states of connected subsystems $x_{-i} \in \mathbb{R}^{N-N_i}$, and
consider n *local norms* $\|\cdot\|_i$ on $\mathbb{R}^{N_i}$





Assume for each node i:

- **contractivity wrt x_i**: $\mathsf{osLip}_{x_i}(\mathsf{F}_i) \leq -c_i < 0$, uniformly in x_{-i}
- **Lipschitz wrt x_j**: $\mathsf{Lip}_{x_j}(\mathsf{F}_i) \leq \ell_{ij}$, uniformly in x_{-j}

Assume for each node i:

- **contractivity wrt x_i:** $\quad \mathsf{osLip}_{x_i}(\mathsf{F}_i) \leq -c_i < 0, \quad$ uniformly in x_{-i}
- **Lipschitz wrt x_j:** $\quad \mathsf{Lip}_{x_j}(\mathsf{F}_i) \leq \ell_{ij}, \quad$ uniformly in x_{-j}

Network contraction theorem

If the Lipschitz constants matrix $\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix}$ is **Hurwitz**

$\implies$ the **interconnected system** is infinitesimally contracting

History: interconnection of stable systems, method of vector Lyapunov functions, connective stability via M-matrix theory – Matrosov and Bellman 1962, Ström, Siljak, Russo/DiBernardo/Sontag, ...

$$\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix} \text{ is } \textbf{Metzler}$$

Hurwitzness depends upon both topology and edge weights

- Hurwitz iff there exists a positive ξ such that $M\xi < \mathbb{0}_n$ (power method)
- Hurwitz iff Lyapunov diagonally stable

$$\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix} \text{ is } \textbf{Metzler}$$

Hurwitzness depends upon both topology and edge weights

- Hurwitz iff there exists a positive ξ such that $M\xi < \mathbb{0}_n$ (power method)
- Hurwitz iff Lyapunov diagonally stable

- for $n = 2$, Hurwitz if and only if **small gain condition**

$$\text{cycle gain} := \frac{\ell_{12}}{c_1}\frac{\ell_{21}}{c_2} < 1$$

and, for $n \geq 3$, **network small-gain theorem for Metzler matrices**

X. Duan, S. Jafarpour, and F. Bullo. Graph-theoretic stability conditions for Metzler matrices and monotone systems. *SIAM Journal on Control and Optimization*, 59(5):3447–3471, 2021.

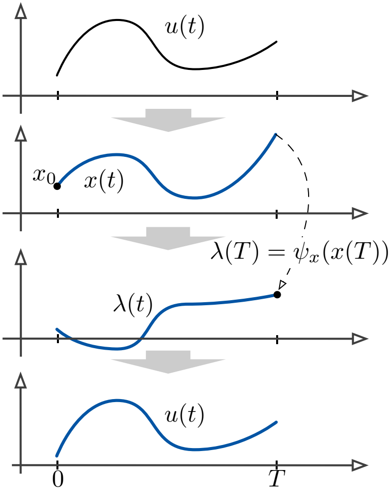
#3: From closed to systems with optimal controls

For $\dot{x} = \mathsf{F}(x, u)$, compute $u : [0, T] \to \mathbb{R}^k$ to minimize $\psi(x(T)) + \int_0^T \phi(x, u) dt$

#3: From closed to systems with optimal controls

For $\dot{x} = \mathsf{F}(x, u)$, compute $u : [0, T] \to \mathbb{R}^k$ to minimize $\quad \psi(x(T)) + \int_0^T \phi(x, u)dt$

Pontryagin Minimum Principle: $\quad u = \mathcal{FBS}[u]$



$\mathcal{F}: \quad \dot{x} = \mathsf{F}(x, u)$

$\mathcal{B}: \quad \dot{\lambda} = -\mathsf{J}_{\mathsf{F}}^{\top}(x, u)\lambda - \phi_x(x, u)$

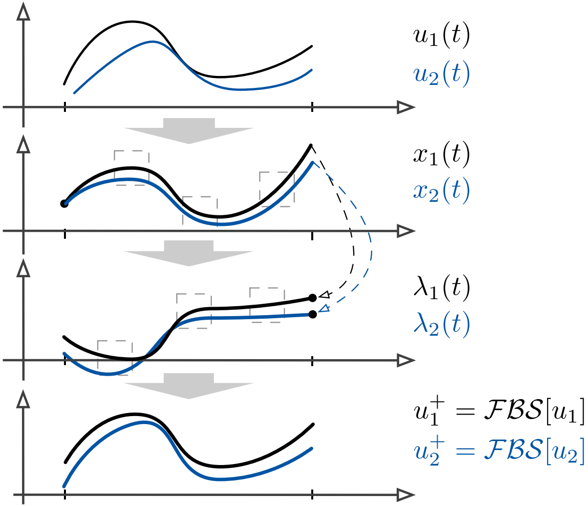
$\mathcal{S}: \quad u = \mathrm{argmin}_{\tilde{u}} \underbrace{\lambda^{\top}\mathsf{F}(x, \tilde{u}) + \phi(x, \tilde{u})}_{\mathcal{H}(x, \tilde{u}, \lambda)}$

To compute a solution to:

$$u = \mathcal{FBS}[u]$$

adopt

$$u^+ = \mathcal{FBS}[u]$$

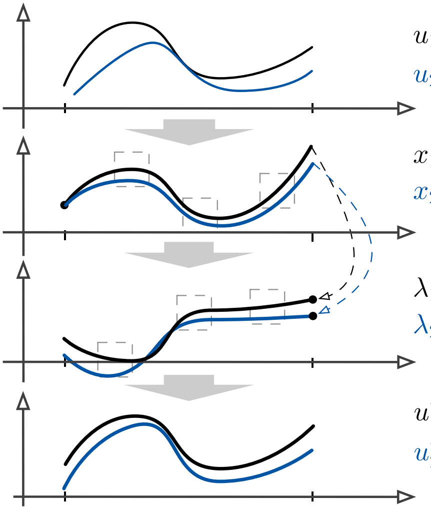


To compute a solution to:

$$u = \mathcal{FBS}[u]$$

adopt

$$u^+ = \mathcal{FBS}[u]$$



If $\text{osLip}_x(\text{F}) = -c$ and all other maps are Lipschitz,

1. $\text{osLip}_\lambda(\text{Adjoint}(\text{F})) \;=\; \text{osLip}_x(\text{F})$
2. $\text{Lip}(\mathcal{FBS}) = \text{constant} \times \dfrac{1 - \mathrm{e}^{-cT}}{c}$

$\mathcal{FBS}$ contracting for short T or large c

Summary

contractivity = robust computationally-friendly stability

fixed point theory + Lyapunov stability theory + geometry of metric spaces

From closed to open, interconnected and optimal systems:

1. iISS
2. network small gain theorems
3. numerical optimal control

Summary

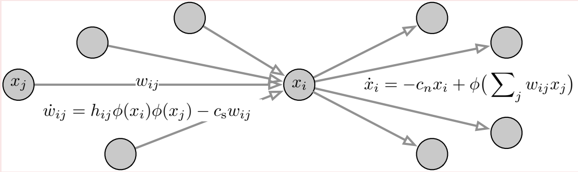
contractivity = robust computationally-friendly stability

fixed point theory + Lyapunov stability theory + geometry of metric spaces

From closed to open, interconnected and optimal systems:

1. iISS
2. network small gain theorems
3. numerical optimal control

Applications coupled neural-synaptic dynamics and ML via optimal control



Outline

From nominal to uncertain systems

Given a norm $\|\cdot\|$, consider

$$\dot{x} = \mathsf{F}(x) + \Delta(x)$$

Assume:

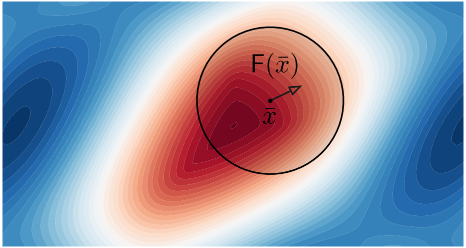
- **contractivity**: $\text{osLip}(\mathsf{F}) \leq -c < 0$
- **bounded disturbance**: $\text{osLip}(\Delta) \leq d < c$

Then

1. $\mathsf{F} + \Delta$ is strongly contracting with rate $c - d$
2. the unique equilibria x^*_{F} of F and $x^*_{\mathsf{F}+\Delta}$ of $\mathsf{F} + \Delta$ satisfy

$$\|x^*_{\mathsf{F}} - x^*_{\mathsf{F}+\Delta}\| \leq \frac{\|\Delta(x^*_{\mathsf{F}})\|}{c - d}$$

From global to local contractivity



Given a norm $\|\cdot\|$, consider

$$\dot{x} = \mathsf{F}(x)$$

Assume:

- **contractivity over closed set D**: $\text{osLip}(\mathsf{F}|_D) \leq -c < 0$
- **existence of almost equilibrium**: D contains the closed B at $\bar{x}$ of radius $r \geq \|\mathsf{F}(\bar{x})\|/c$

Then

1. B is forward invariant
2. $\mathsf{F}|_B$ is strongly infinitesimally contracting

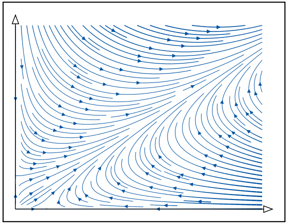
From strongly to weakly contracting systems

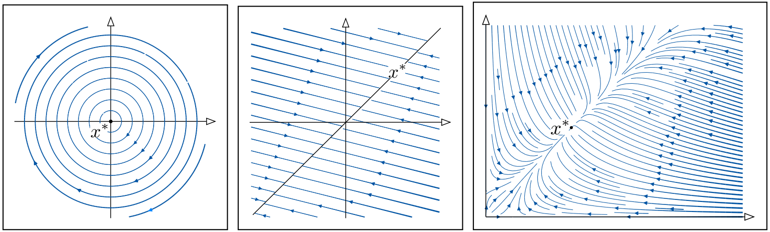
Given a norm $\|\cdot\|$, consider

$$\dot{x} = \mathsf{F}(x) \qquad \text{satisfying} \qquad \mathsf{osLip}(\mathsf{F}) = 0$$

Dichotomy for weakly-contracting systems

1. no equilibrium and every trajectory is unbounded, or
2. at least one equilibrium, every trajectory is bounded, and local asy stability $\implies$ global





Outline

1. Contractivity of dynamical systems
   - From discrete-time to continuous-time dynamics
   - Table of infinitesimal contractivity conditions
   - Application to recurrent neural networks
   - Connection with convex optimization
2. From closed to open, interconnected and optimal systems
   - Incremental input-to-state stability
   - Interconnected contracting systems
   - Contractivity in indirect optimal control
3. Additional robustness, computational and stability properties
4. Conclusions and Future Research

Robust and computationally-friendly stability theory

1. contractivity conditions on normed vector spaces
2. convexity and fixed point methods
3. disturbances, interconnections and optimal control

	Lyapunov Theory	**Contraction Theory for Dynamical Systems**
	F admits global Lyapunov function	F is strongly contracting
existence of equilibrium	assumed	implied + computational methods
Lyapunov function	arbitrary	distance to trajectory (+ norm of vector field)
inputs	ISS via $\mathcal{KL}$ and $\mathcal{L}$ functions	iISS via explicit formulas

search for contraction properties
design engineering systems to be contracting

References (1/2)

Contraction theory on normed spaces:

- A. Davydov, S. Jafarpour, and F. Bullo. Non-Euclidean contraction theory for robust nonlinear stability. *IEEE Transactions on Automatic Control*, 67(12):6667–6681, 2022b.
- S. Jafarpour, A. Davydov, and F. Bullo. Non-Euclidean contraction theory for monotone and positive systems. *IEEE Transactions on Automatic Control*, 2023.

Contractivity in optimal control:

- K. D. Smith and F. Bullo. Contractivity of the method of successive approximations for optimal control. *IEEE Control Systems Letters*, Nov. 2022.

Contracting neural networks and fixed point theory:

- S. Jafarpour, A. Davydov, A. V. Proskurnikov, and F. Bullo. Robust implicit networks via non-Euclidean contractions. In *Advances in Neural Information Processing Systems*, Dec. 2021.
- A. Davydov, A. V. Proskurnikov, and F. Bullo. Non-Euclidean contractivity of recurrent neural networks. In *American Control Conference*, pages 1527–1534, Atlanta, USA, May 2022d.
- F. Bullo, P. Cisneros-Velarde, A. Davydov, and S. Jafarpour. From contraction theory to fixed point algorithms on Riemannian and non-Euclidean spaces. In *IEEE Conf. on Decision and Control*, Dec. 2021.

References (2/2)

Here at CDC 2022

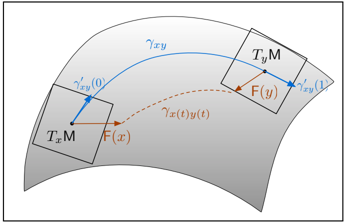
- A. Davydov, S. Jafarpour, A. V. Proskurnikov, and F. Bullo. Non-Euclidean monotone operator theory with applications to recurrent neural networks. In *IEEE Conf. on Decision and Control*, Dec. 2022c.
- A. Davydov, S. Jafarpour, M. Abate, F. Bullo, and S. Coogan. Comparative analysis of interval reachability for robust implicit and feedforward neural networks. In *IEEE Conf. on Decision and Control*, 2022a. URL **https://arxiv.org/abs/2204.00187**
- V. Centorrino, F. Bullo, and G. Russo. Contraction analysis of Hopfield neural networks with Hebbian learning. In *IEEE Conf. on Decision and Control*, Dec. 2022.
- R. Ofir, F. Bullo, and M. Margaliot. Minimum effort decentralized control design for contracting network systems. *IEEE Control Systems Letters*, 6:2731–2736, 2022.

Resources on contraction theory for dynamics, control and learning

1. tutorial session "Contraction Theory for Machine Learning" at the 2021 IEEE CDC conference:
   https://sites.google.com/view/contractiontheory
2. free online book and 10h minicourse
   http://motion.me.ucsb.edu/book-ctds
   https://youtu.be/RvR47ZbqJjc
3. upcoming Workshop on "Contraction Theory for Systems, Control, and Learning" at the 2023 American Control Conference in San Diego, California (under review):
   http://motion.me.ucsb.edu/contraction-workshop-2023

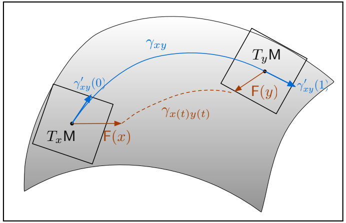
Theoretical frontiers

- higher order contraction
- relationship with monotone operator theory
- metric spaces: seminorms, Hilbert metrices ...



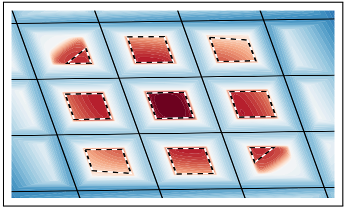
Theoretical frontiers

- higher order contraction
- relationship with monotone operator theory
- metric spaces: seminorms, Hilbert metrices ...



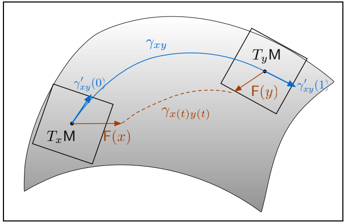
Limitations: not all stable systems are contractive:

- Lyapunov-diagonally-stable networks
- multistable systems
- biochemical networks



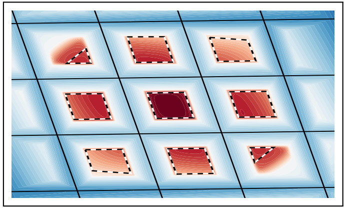
Theoretical frontiers

- higher order contraction
- relationship with monotone operator theory
- metric spaces: seminorms, Hilbert metrices ...



Limitations: not all stable systems are contractive:

- Lyapunov-diagonally-stable networks
- multistable systems
- biochemical networks



Application to control and learning

1. control: optimization-based control design
2. ML: implicit models and energy-based learning
3. neuroscience: robust dynamical modeling