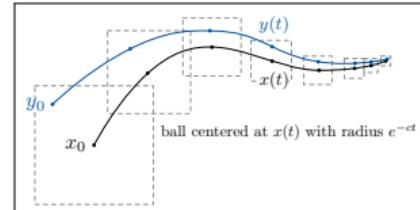
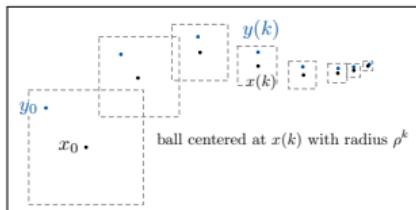


Contracting Dynamical Systems: A Tutorial on Theory and Applications

Francesco Bullo

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
<https://fbullo.github.io>

Tutorial (based on lectures in Napoli Nov '22 and San Diego Jun '23). This version: 2023/06/17

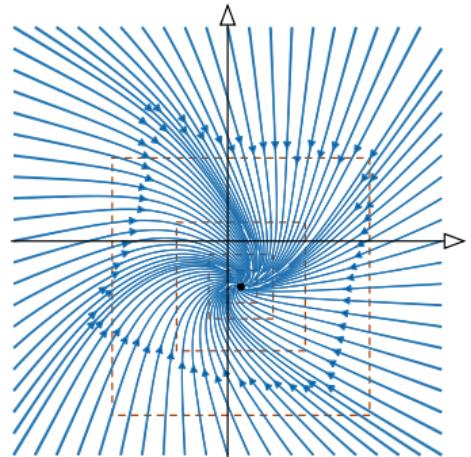


contractivity = robust computationally-friendly stability

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

Acknowledgments



Veronica Centorrino
Scuola Sup
Meridionale



Pedro Cisneros-Velarde
UIUC



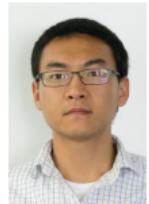
Alex Davydov
UC Santa Barbara



Giulia De Pasquale
ETH



Robin Delabays
HES-SO Sion



Xiaoming Duan
Shanghai Jiao Tong



Anand Gokhale
UC Santa Barbara



Saber Jafarpour
GeorgiaTech



Anton Proskurnikov
Politecnico Torino



Giovanni Russo
Univ Salerno



John W.
Simpson-Porco
University of Toronto



Kevin D. Smith
Utilidata



Elena Valcher
Universita di Padova

Outline

1 History and resources

2 Basic definitions

- Discrete- and continuous-time dynamics on vector spaces
- Dynamics on Riemannian manifolds

3 Examples

- Optimization-based dynamics
- Recurrent neural network dynamics

4 Properties of contracting dynamics

- iISS
- Periodic systems
- Composite norms and interconnected systems
- Contractivity of delay dynamics
- Forward Euler theorem

5 Generalizations

6 Conclusions and future research

7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality

- Systems with invariance/conservation properties
- Induced seminorms and duality

8 Advanced Topics: Time-varying convex optimization via contracting dynamics

- Tracking equilibrium trajectories

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. doi: 

- **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. doi: 

- **Systems and control:**

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

- **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. 2014b. doi: 

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. doi: 

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. doi: 

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. doi: 

The Banach Contraction Theorem is also referred to as the *Picard-Banach-Caccioppoli*, because of the earlier work by Picard (1890) on the “method of successive approximations” and the later independent work by Renato Caccioppoli (1930).



Figure: Renato Caccioppoli (Napoli, 20 gennaio 1904 – Napoli, 8 maggio 1959) was an Italian mathematician

1921-1932 student and researcher @ Napoli

1931-1934 professor @ Padova

1934-1959 professor @ Napoli

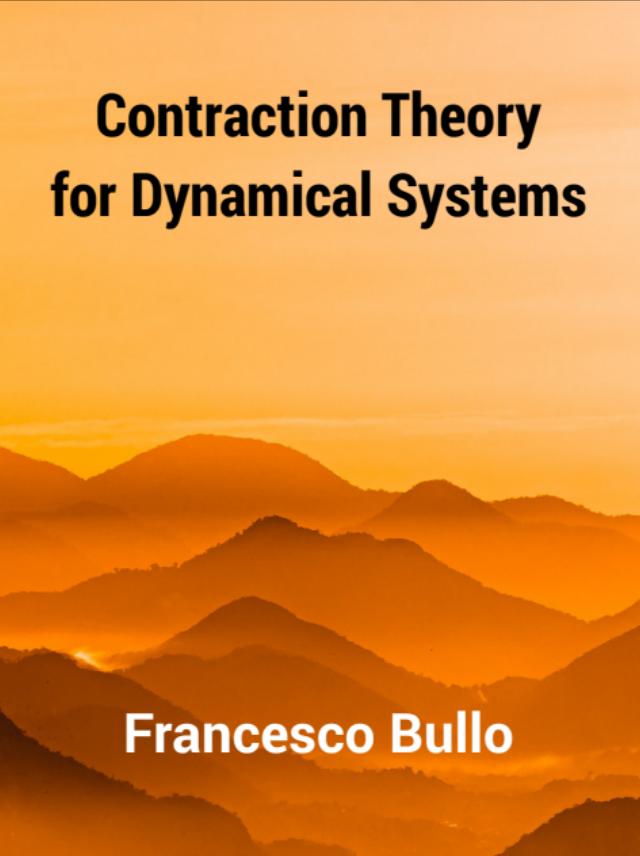
R. Caccioppoli. Un teorema generale sull'esistenza di elementi uniti in una trasformazione funzionale. *Rendiconti dell'Accademia Nazionale dei Lincei*, 11:794–799, 1930

Contraction conditions without Jacobians

- ① **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. 
- ② no-name in: A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Kluwer, 1988. ISBN 902772699X (Chapter 1, page 5)
- ③ **one-sided Lipschitz maps** in: E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer, 1993.  (Section 1.10, Exercise 6)
- 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. 

Links to recent related educational and research events

- 2023 ACC Workshop on "Contraction Theory for Systems, Control, and Learning"
<http://motion.me.ucsb.edu/contraction-workshop-2023>
- Tutorial session: <https://sites.google.com/view/contractiontheory> "Contraction Theory for Machine Learning" (PDFs and youtube videos) at the 2021 IEEE CDC conference, by Soon-Jo Chung, Jean-Jacques Slotine, and Hiroyasu Tsukamoto
- Tutorial paper at CDC2021 "Contraction-Based Methods for Stable Identification and Robust Machine Learning: a Tutorial" by Ian Manchester and coauthors: <https://arxiv.org/abs/2110.00207>,
<https://ieeexplore.ieee.org/abstract/document/9683128>
- Plenary presentation: (Slides
<https://fbullo.github.io/talks/2022-12-FBullo-ContractionSystemsControl-CDC.pdf>) "Contraction Theory in Systems and Control" by Francesco Bullo at the 2022 IEEE CDC
- Youtube lectures: "Lectures on Nonlinear Systems" by Jean-Jacques Slotine, Fall 2013:
<https://web.mit.edu/nsl/www/videos/lectures.html>, Lectures 14-20 (approximately 1h20min each)
- Youtube lectures: "Minicourse on Contraction Theory" by Francesco Bullo, Fall 2022. Youtube lectures
<https://youtu.be/RvR47ZbqJjc>: 10h in 4 lectures, with chapters
- Textbook: Contraction Theory for Dynamical Systems, Francesco Bullo, rev 1.1, Mar 2023. (Book and slides freely available) <https://fbullo.github.io/ctds>



Contraction Theory for Dynamical Systems

Francesco Bullo

Contraction Theory for Dynamical Systems, Francesco Bullo,
KDP, 1.1 edition, 2023, ISBN 979-8836646806

- ➊ Textbook with exercises and answers. Format: textbook, slides, and paperbook
- ➋ Content:
 - Fixed point theory
 - Theory of contracting dynamics on vector spaces
 - Applications to nonlinear and interconnected systems
- ➌ Self-Published and Print-on-Demand at:
<https://www.amazon.com/dp/B0B4K1BTF4>
- ➍ PDF Freely available at
<https://fbullo.github.io/ctds>
- ➎ 10h minicourse on youtube:
<https://youtu.be/RvR47ZbqJjc>
- ➏ Future version to include: systems on Riemannian manifolds, homogeneous spaces, and solid cones
 - "Continuous improvement is better than delayed perfection"
 - Mark Twain**

Selected references from my group

Contraction theory on normed spaces and Riemannian manifolds:

- A. Davydov, S. Jafarpour, and F. Bullo. Non-Euclidean contraction theory for robust nonlinear stability. *IEEE Transactions on Automatic Control*, 67(12):6667–6681, 2022a. doi: 
- S. Jafarpour, A. Davydov, and F. Bullo. Non-Euclidean contraction theory for monotone and positive systems. *IEEE Transactions on Automatic Control*, 2023. doi: To appear
- J. W. Simpson-Porco and F. Bullo. Contraction theory on Riemannian manifolds. *Systems & Control Letters*, 65:74–80, 2014. doi: 

Contracting neural networks:

- 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. doi: 
- 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 2022c. doi: 
- V. Centorrino, A. Gokhale, A. Davydov, G. Russo, and F. Bullo. Euclidean contractivity of neural networks with symmetric weights. *IEEE Control Systems Letters*, 7:1724–1729, 2023. doi: 

Weak and semicontraction theory:

- S. Jafarpour, P. Cisneros-Velarde, and F. Bullo. Weak and semi-contraction for network systems and diffusively-coupled oscillators. *IEEE Transactions on Automatic Control*, 67(3):1285–1300, 2022. doi: 
- G. De Pasquale, K. D. Smith, F. Bullo, and M. E. Valcher. Dual seminorms, ergodic coefficients, and semicontraction theory. *IEEE Transactions on Automatic Control*, 2022. doi: Submitted
- R. Delabays and F. Bullo. Semicontraction and synchronization of Kuramoto-Sakaguchi oscillator networks. *IEEE Control Systems Letters*, 7:1566–1571, 2023. doi: 

Optimization:

- 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. doi: 
- 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*, Cancún, México, Dec. 2022b. doi: 
- A. Davydov, V. Centorrino, A. Gokhale, G. Russo, and F. Bullo. Contracting dynamics for time-varying convex optimization. *IEEE Transactions on Automatic Control*, June 2023. doi: Submitted

Outline

1 History and resources

2 Basic definitions

- Discrete- and continuous-time dynamics on vector spaces
- Dynamics on Riemannian manifolds

3 Examples

- Optimization-based dynamics
- Recurrent neural network dynamics

4 Properties of contracting dynamics

- iISS
- Periodic systems
- Composite norms and interconnected systems
- Contractivity of delay dynamics
- Forward Euler theorem

5 Generalizations

6 Conclusions and future research

7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality

- Systems with invariance/conservation properties
- Induced seminorms and duality

8 Advanced Topics: Time-varying convex optimization via contracting dynamics

- Tracking equilibrium trajectories

Banach Contraction Theorem Let (\mathcal{X}, d) be a *complete metric space*

If $T : \mathcal{X} \rightarrow \mathcal{X}$ is Lipschitz with constant $\ell < 1$ (called the *contraction factor*), then

- ① T has a unique fixed point x^* in \mathcal{X}
- ② the sequence $\{x_k\}_{k \in \mathbb{N}}$ generated by the *Picard iteration* $x_{k+1} = T(x_k)$ converges to x^* for all initial conditions $x_0 \in \mathcal{X}$
- ③ the following error estimates hold for all $k \in \mathbb{N}$:

(geometric convergence):

$$d(x_k, x^*) \leq \ell^k d(x_0, x^*)$$

(a-priori upper bound):

$$d(x_k, x^*) \leq \frac{\ell^k}{1 - \ell} d(x_0, x_1)$$

(a-posteriori upper bound):

$$d(x_k, x^*) \leq \frac{\ell}{1 - \ell} d(x_{k-1}, x_k)$$

Proof

For $x_{k+1} = T(x_k)$

- sequence $\{x_k\}_{k \in \mathbb{N}}$ is Cauchy

$$\begin{aligned} d(x_{k+h}, x_k) &\leq d(x_{k+h}, x_{k+h-1}) + \cdots + d(x_{k+1}, x_k) \\ &\leq (\ell^{h-1} + \cdots + 1)d(x_{k+1}, x_k) \\ &\leq \frac{1}{1-\ell}d(x_{k+1}, x_k) \\ &\leq \frac{\ell^k}{1-\ell}d(x_1, x_0) \end{aligned}$$

- since \mathcal{X} is complete, sequence converges to a point x^*
- uniqueness from $\ell < 1$
- geometric convergence

$$d(x_k, x^*) = d(T(x_{k-1}), x^*) \leq \ell d(x_{k-1}, x^*) \leq \ell^k d(x_0, x^*)$$

Linear algebra: induced norms

Vector norm

$$\|x\|_1 = \sum_{i=1}^n |x_i|$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$$

$$\|x\|_\infty = \max_{i \in \{1, \dots, n\}} |x_i|$$

Induced matrix norm

$$\|A\|_1 = \max_{j \in \{1, \dots, n\}} \sum_{i=1}^n |a_{ij}|$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$$

$$\|A\|_\infty = \max_{i \in \{1, \dots, n\}} \sum_{j=1}^n |a_{ij}|$$

Induced matrix log norm

$$\begin{aligned}\mu_1(A) &= \max_{j \in \{1, \dots, n\}} \left(a_{jj} + \sum_{i=1, i \neq j}^n |a_{ij}| \right) \\ &= \text{max column "absolute sum" of } A\end{aligned}$$

$$\mu_2(A) = \lambda_{\max}\left(\frac{A + A^\top}{2}\right)$$

$$\begin{aligned}\mu_\infty(A) &= \max_{i \in \{1, \dots, n\}} \left(a_{ii} + \sum_{j=1, j \neq i}^n |a_{ij}| \right) \\ &= \text{max row "absolute sum" of } A\end{aligned}$$

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

Lipschitz constant

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

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

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

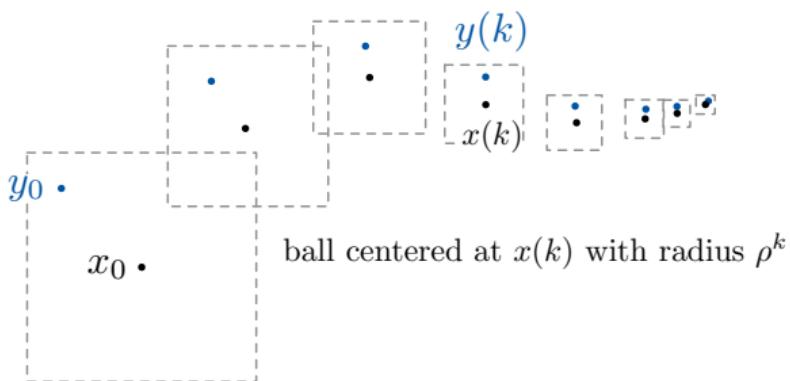
$$\|x\|_{2,P} = (x^\top Px)^{1/2} \quad \text{Lip}_{2,P}(\mathsf{F}_A) = \|A\|_{2,P} \leq \ell \iff A^\top PA \preceq \ell^2 P$$

$$\|x\|_{\infty,\eta} = \max_i |x_i|/\eta_i \quad \text{Lip}_{\infty,\eta}(\mathsf{F}_A) = \|A\|_{\infty,\eta} \leq \ell \iff \eta^\top |A| \leq \ell \eta^\top$$

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 induced norms to induced log norms

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

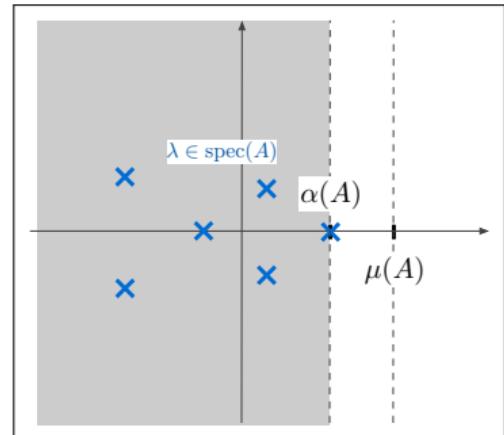
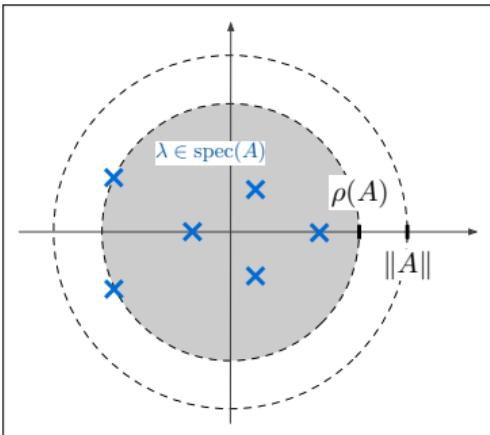
$$\mu(A) := \lim_{h \rightarrow 0^+} \frac{\|I_n + hA\| - 1}{h}$$

subadditivity:

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

scaling:

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



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

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

One-sided Lipschitz constant

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

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

For **affine map** $F_A(x) = Ax + a$

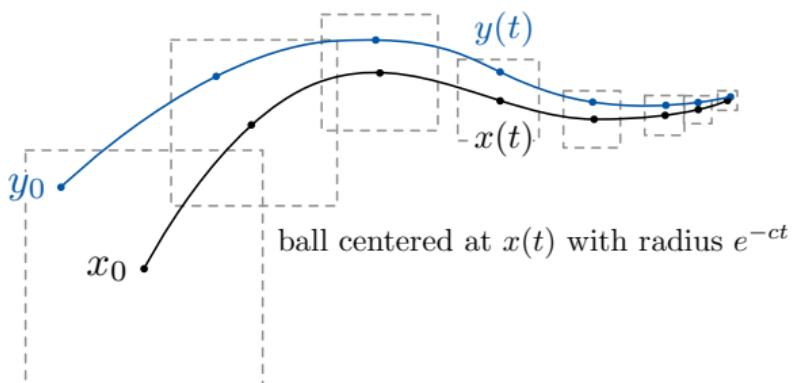
$$\text{osLip}_{2,P}(F_A) = \mu_{2,P}(A) \leq \ell \iff A^\top P + AP \preceq 2\ell P$$

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

Banach contraction theorem for continuous-time dynamics:

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

A **weak pairing** is $\llbracket \cdot, \cdot \rrbracket : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ satisfying

- ① $\llbracket x_1 + x_2, y \rrbracket \leq \llbracket x_1, y \rrbracket + \llbracket x_2, y \rrbracket$ and $x \mapsto \llbracket x, y \rrbracket$ is continuous,
- ② $\llbracket bx, y \rrbracket = \llbracket x, by \rrbracket = b \llbracket x, y \rrbracket$ for $b \geq 0$ and $\llbracket -x, -y \rrbracket = \llbracket x, y \rrbracket$,
- ③ $\llbracket x, x \rrbracket > 0$, for all $x \neq 0_n$,
- ④ $|\llbracket x, y \rrbracket| \leq \llbracket x, x \rrbracket^{1/2} \llbracket y, y \rrbracket^{1/2}$,

Given norm $\|\cdot\|$, compatibility: $\llbracket x, x \rrbracket = \|x\|^2$ for all x

Key properties

Curve norm derivative formula:

$$\frac{1}{2} D^+ \|x(t)\|^2 = \llbracket \dot{x}(t), x(t) \rrbracket$$

Sup of non-Euclidean numerical range:

$$\mu(A) = \sup_{\|x\|=1} \llbracket Ax, x \rrbracket$$

Example weak pairings

Norms

From inner products to sign and max pairings

From LMIs to log norms

$$\|x\|_{2,P^{1/2}}^2 = x^\top Px$$

$$[\![x, y]\!]_{2,P^{1/2}} = x^\top Py$$

$$\mu_{2,P^{1/2}}(A) = \min\{b \mid A^\top P + PA \preceq 2bP\}$$

$$\|x\|_1 = \sum_i |x_i|$$

$$[\![x, y]\!]_1 = \|y\|_1 \text{sign}(y)^\top x$$

$$\mu_1(A) = \max_j \left(a_{jj} + \sum_{i \neq j} |a_{ij}| \right)$$

$$\|x\|_\infty = \max_i |x_i|$$

$$[\![x, y]\!]_\infty = \max_{i \in I_\infty(y)} y_i x_i$$

$$\mu_\infty(A) = \max_i \left(a_{ii} + \sum_{j \neq i} |a_{ij}| \right)$$

where $I_\infty(x) = \{i \in \{1, \dots, n\} \text{ such that } |x_i| = \|x\|_\infty\}$

Table of continuous-time contractivity conditions

Log Norm bound	Demidovich condition	One-sided Lipschitz condition
$\mu_{2,P}(D\mathbf{F}(x)) \leq b$	$P D\mathbf{F}(x) + D\mathbf{F}(x)^\top P \preceq 2bP$	$(x - y)^\top P(\mathbf{F}(x) - \mathbf{F}(y)) \leq b\ x - y\ _{P^{1/2}}^2$
$\mu_1(D\mathbf{F}(x)) \leq b$	$\text{sign}(v)^\top D\mathbf{F}(x)v \leq b\ v\ _1$	$\text{sign}(x - y)^\top (\mathbf{F}(x) - \mathbf{F}(y)) \leq b\ x - y\ _1$
$\mu_\infty(D\mathbf{F}(x)) \leq b$	$\max_{i \in I_\infty(v)} v_i (D\mathbf{F}(x)v)_i \leq b\ v\ _\infty^2$	$\max_{i \in I_\infty(x-y)} (x_i - y_i)(\mathbf{F}_i(x) - \mathbf{F}_i(y)) \leq b\ x - y\ _\infty^2$

Equivalent contractivity conditions

J. A. Jacquez and C. P. Simon. Qualitative theory of compartmental systems. *SIAM Review*, 35(1):43–79, 1993. [doi](#)

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. [doi](#)

G. Como, E. Lovisari, and K. Savla. Throughput optimality and overload behavior of dynamical flow networks under monotone distributed routing. *IEEE Transactions on Control of Network Systems*, 2(1):57–67, 2015. [doi](#)

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 *reachability analysis via mixed-monotone embeddings*

ℓ_∞ suited for mixed-monotone embeddings

- ⑥ *asynchronous distributed computation*

ℓ_∞ contractions converge under fully asynchronous distributed execution

NonEuclidean contractions: biological transcriptional systems (Russo et al., 2010), Hopfield neural networks (Fang and Kincaid, 1996; Qiao et al., 2001), chemical reaction networks (Al-Radhawi and Angeli, 2016), traffic networks (Coogan and Arcak, 2015; Como et al., 2015; Coogan, 2019), multi-vehicle systems (Monteil et al., 2019), and coupled oscillators (Russo et al., 2013; Aminzare and Sontag, 2014a)

Contraction dynamics on Riemannian manifolds

Contraction theory on Riemannian manifolds originates in

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

A formal coordinate-free analysis (with connection to monotone operators) is given in

J. W. Simpson-Porco and F. Bullo. Contraction theory on Riemannian manifolds. *Systems & Control Letters*, 65:74–80, 2014. doi:

In the differential geometry literature, geodesically monotonic vector fields are studied by

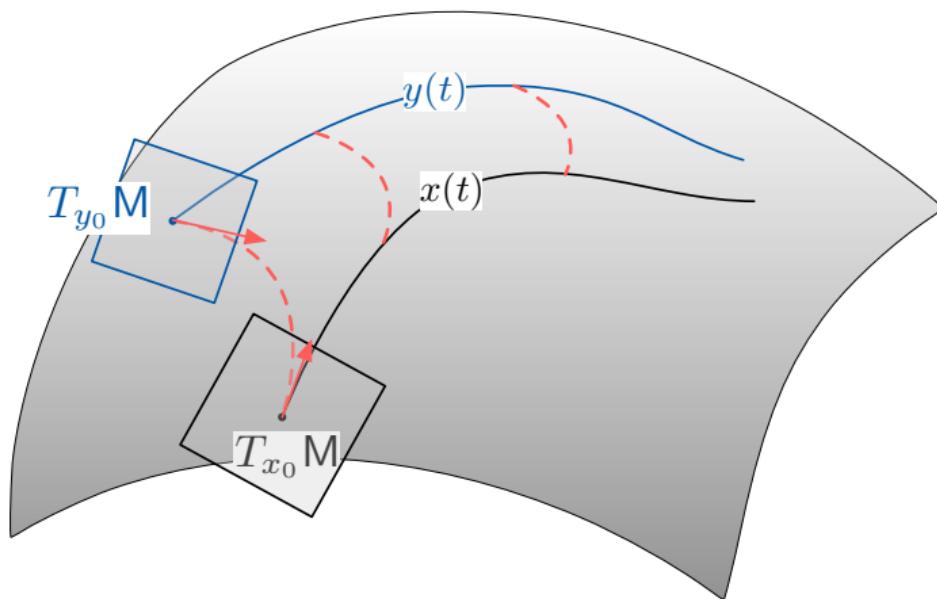
S. Z. Németh. Geodesic monotone vector fields. *Lobachevskii Journal of Mathematics*, 5:13–28, 1999. URL
<http://mi.mathnet.ru/eng/ljm145>

J. X. Da Cruz Neto, O. P. Ferreira, and L. R. Lucambio Pérez. Contributions to the study of monotone vector fields. *Acta Mathematica Hungarica*, 94(4):307–320, 2002. doi:

J. H. Wang, G. López, V. Martín-Márquez, and C. Li. Monotone and accretive vector fields on Riemannian manifolds. *Journal of Optimization Theory and Applications*, 146(3):691–708, 2010. doi:

Assume existence and uniqueness of geodesic curve between each (x, y)

F **contracting** if geodesic distances from x to y diminishes along the flow of F



integral test: the inner product between F and the geodesic velocity vector γ' at x and y

differential test: condition on covariant differential of F

Given vector field F on a Riemannian manifold (M, \mathbb{G}) and $c > 0$, equivalent statements:

- ① **integral condition:** for each $x, y \in M$ and geodesic $\gamma : [0, 1] \rightarrow M$ with $\gamma(0) = x, \gamma(1) = y$,

$$\langle\langle F(y), \gamma'(1) \rangle\rangle_{\mathbb{G}} - \langle\langle F(x), \gamma'(0) \rangle\rangle_{\mathbb{G}} \leq -c d_{\mathbb{G}}(x, y)^2$$

or, equivalently, using the parallel transport map $P_{y \rightarrow x} : T_y M \rightarrow T_x M$,

$$\langle\langle P_{y \rightarrow x} F(y) - F(x), \gamma'(0) \rangle\rangle_{\mathbb{G}} \leq -c d_{\mathbb{G}}(x, y)^2$$

- ② **differential condition:** for all $v_x \in T_x M$

$$\langle\langle \nabla_{v_x} F(x), v_x \rangle\rangle_{\mathbb{G}} \leq -c \|v_x\|_{\mathbb{G}}^2,$$

where ∇ is the Levi-Civita connection. In components:

$$\mathbb{G}(x) D F(x) + D F(x)^\top \mathbb{G}(x) + \mathcal{L}_F \mathbb{G}(x) \preceq -2c \mathbb{G}(x)$$

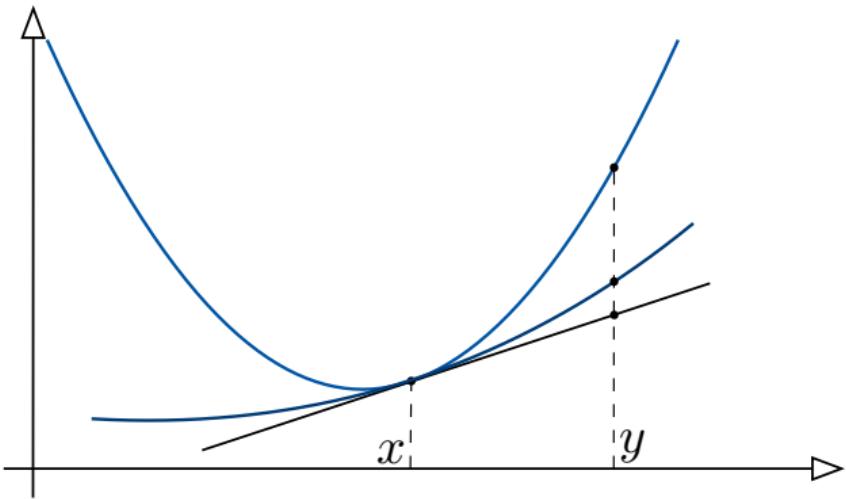
- ③ **trajectory condition:** for all solutions $x(\cdot), y(\cdot)$

$$D^+ d_{\mathbb{G}}(x(t), y(t)) \leq -c d_{\mathbb{G}}(x(t), y(t))$$

Outline

- 1 History and resources
- 2 Basic definitions
 - Discrete- and continuous-time dynamics on vector spaces
 - Dynamics on Riemannian manifolds
- 3 Examples
 - Optimization-based dynamics
 - Recurrent neural network dynamics
- 4 Properties of contracting dynamics
 - iISS
 - Periodic systems
 - Composite norms and interconnected systems
 - Contractivity of delay dynamics
 - Forward Euler theorem
- 5 Generalizations
- 6 Conclusions and future research
- 7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality
 - Systems with invariance/conservation properties
 - Induced seminorms and duality
- 8 Advanced Topics: Time-varying convex optimization via contracting dynamics
 - Tracking equilibrium trajectories

Optimization-based dynamics



$$V(y) \geq V(x) + \text{grad } V(x)^\top (y - x) + \frac{m}{2} \|x - y\|_2^2$$

Example #1: Gradient flow for strongly convex function

Given strongly convex $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with parameter μ , **gradient dynamics**

$$\dot{x} = F_G(x) := -\nabla f(x)$$

F_G is infinitesimally contracting wrt $\|\cdot\|_2$ with rate μ

If f is twice-differentiable, then $\text{Hess } f(x) \succeq \mu I_n$ for all x

$$\begin{aligned} D(-\nabla f)(x) &= -\text{Hess } f(x) \preceq -\mu I_n \\ \iff I_n D(-\nabla f)(x) + D(-\nabla f)(x)^\top I_n &\preceq -2\mu I_n \end{aligned}$$

Convexity and contractivity

Kachurovskii's Theorem: For differentiable $f : \mathbb{R}^n \rightarrow \mathbb{R}$, equivalent statements:

- ① f is **strongly convex** with parameter m
- ② $-\text{grad}f$ is **m -strongly infinitesimally contracting**, that is

$$(-\text{grad}f(x) + \text{grad}f(y))^\top (x - y) \leq -m\|x - y\|_2^2$$

Also: global minimum of f = globally-exponentially stable equilibrium of $-\nabla f$

For map $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, equivalent statements:

- ① F is a **monotone operator^a** (or a **coercive operator**) with parameter m ,
- ② $-F$ is **m -strongly contracting**

^a $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a **m -strongly monotone operator** if $\langle F(x) - F(y), x - y \rangle \geq m\|x - y\|_2^2$

Example #2: Saddle dynamics

Assume $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$

- $x \mapsto f(x, y)$ is μ_x -strongly convex, uniformly in y
- $y \mapsto f(x, y)$ is μ_y -strongly concave, uniformly in x

saddle dynamics (primal-descent / dual-ascent):

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \mathsf{F}_S(x, y) := \begin{bmatrix} -\nabla_x f(x, y) \\ \nabla_y f(x, y) \end{bmatrix}$$

F_S is infinitesimally contracting wrt $\|\cdot\|_2$ with rate $\min\{\mu_x, \mu_y\}$

If f is twice-differentiable, then

$$\begin{aligned} \mu_2(D\mathsf{F}_S(x, y)) &= \mu_2 \left(\begin{bmatrix} -\text{Hess}_x f(x, y) & -D_y \nabla_x f(x, y) \\ D_x \nabla_y f(x, y) & \text{Hess}_y f(x, y) \end{bmatrix} \right) \\ &\stackrel{\mu_2(A)=\mu_2(\frac{A+A^\top}{2})}{=} \mu_2 \left(\begin{bmatrix} -\text{Hess}_x f(x, y) & 0 \\ 0 & \text{Hess}_y f(x, y) \end{bmatrix} \right) = -\min\{\mu_x, \mu_y\}. \end{aligned}$$

Example #3: Primal-dual gradient dynamics

strongly convex function f

$$\text{s.t. } 0 \prec \mu_{\min} I_n \preceq \text{Hess } f \preceq \mu_{\max} I_n$$

constraint matrix A

$$\text{s.t. } 0 \prec a_{\min} I_m \preceq AA^\top \preceq a_{\max} I_m$$

linearly constrained optimization:

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{s.t. } Ax = b \end{aligned}$$

primal-dual gradient dynamics:

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \mathsf{F}_{\text{PDG}}(x, \lambda) := \begin{bmatrix} -\nabla f(x) - A^\top \lambda \\ Ax - b \end{bmatrix}$$

F_{PDG} is infinitesimally contracting wrt weighted $\|\cdot\|_{2,P^{1/2}}$ with rate c

$$P = \begin{bmatrix} I_n & \alpha A^\top \\ \alpha A & I_m \end{bmatrix}, \quad \alpha = \frac{1}{3} \min \left\{ \frac{1}{\mu_{\max}}, \frac{\mu_{\min}}{a_{\max}} \right\}, \quad \text{and} \quad c = \frac{5}{18} \min \left\{ \frac{a_{\min}}{\mu_{\max}}, \frac{a_{\min}}{a_{\max}} \mu_{\min} \right\}$$

$$\text{For each } \mu_{\min} I_n \preceq Q \preceq \mu_{\max} I_n, \quad \begin{bmatrix} -Q & -A^\top \\ A & 0 \end{bmatrix}^\top P + P \begin{bmatrix} -Q & -A^\top \\ A & 0 \end{bmatrix} \preceq -2cP$$

Example: Distributed optimization from primal-dual gradient descent

Consider a tree (undirected acyclic connected graph) with n nodes and $m = n - 1$ edges with oriented incidence matrix A^\top . Then:

$$0 \prec \lambda_2 I_{n-1} \preceq AA^\top \preceq \lambda_n I_{n-1}$$

decomposable optimization: Rewrite $\min_{x \in \mathbb{R}^n} f(x)$ when $f(x) = \sum_i f_i(x)$ as

$$\begin{aligned} \min_{x_i \in \mathbb{R}^n} \quad & \sum_{i=1}^n f_i(x_i) \\ \text{s.t.} \quad & x_i = x_j \quad \text{for each edge } e = (i, j) \end{aligned}$$

distributed optimization via primal-dual gradient dynamics:

$$\begin{cases} \dot{x}_i &= -\nabla_i f_i(x_i) - \sum_{e : i \text{ is the head of } e} \lambda_e + \sum_{e : i \text{ is the tail of } e} \lambda_e \\ \dot{\lambda}_e &= x_i - x_j \quad \text{for each edge } e = (i, j) \end{cases}$$

When dual dynamics is fast and each f_i is μ_i -strongly convex, $c = \frac{5}{18} \frac{\lambda_2}{\lambda_n} \min_i \mu_i$

Composite minimization and proximal gradient

For strongly convex + strongly smooth f , convex, closed, proper $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$,

$$x^* = \operatorname{argmin}_{x \in \mathbb{R}^n} f(x) + g(x) \iff x^* = \operatorname{prox}_{\gamma g}(x^* - \gamma \nabla f(x))$$

$$\text{where } \operatorname{prox}_{\gamma g}(z) = \operatorname{argmin}_{x \in \mathbb{R}^n} g(x) + \frac{1}{2\gamma} \|x - z\|_2^2.$$

- ① minimization problem

$$\min_{x \in \mathbb{R}^n} f(x) + g(x)$$

- ② is transcribed into strongly infinitesimally contracting *proximal gradient* dynamics

$$\dot{x} = F_{\text{ProxG}}(x) := -x + \operatorname{prox}_{\gamma g}(x - \gamma \nabla f(x))$$

Example #4: Proximal gradient dynamics

proximal gradient dynamics:

$$\dot{x} = \mathsf{F}_{\text{ProxG}}(x) := -x + \text{prox}_{\gamma g}(x - \gamma \nabla f(x))$$

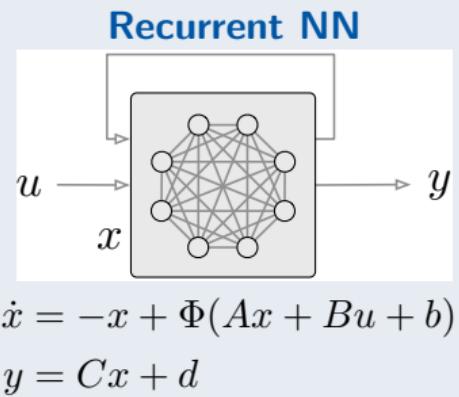
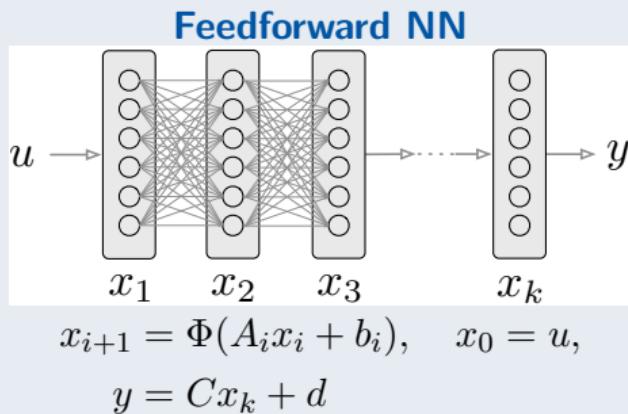
f m -strongly convex and ℓ -strongly smooth

- ① if $0 < \gamma < \frac{2}{\ell}$, then **F_{PDG} is infinitesimally contracting w.r.t. $\|\cdot\|_2$ with rate c**

$$c = 1 - \max\{|1 - \gamma m|, |1 - \gamma \ell|\}$$

and maximal rate at $\gamma^* = \frac{2}{m+\ell}$

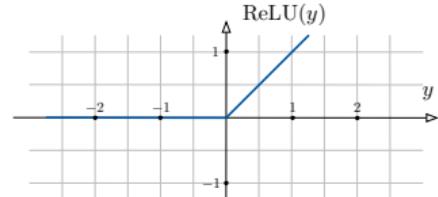
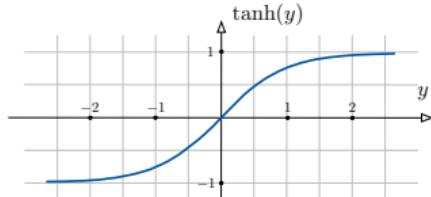
- ② if $f(x) = \frac{1}{2}x^\top Ax + b^\top x$ with $A \succ 0$ and $\gamma > 1/\lambda_{\min}(A)$,
then **F_{PDG} is infinitesimally contracting w.r.t. $\|\cdot\|_{2,(\gamma A - I_n)^{1/2}}$ with rate $c = 1$**



Example #5: Firing-rate recurrent neural network

$$\dot{x} = \mathsf{F}_{\text{FR}}(x) := -x + \Phi(Wx + Bu)$$

sigmoid, hyperbolic tangent
 $\text{ReLU} = \max\{x, 0\} = (x)_+$
 $0 \leq \Phi'_i(y) \leq 1$



F_{FR} is infinitesimally contracting wrt $\|\cdot\|_\infty$ with rate $1 - \mu_\infty(W)_+$ if

$$\mu_\infty(W) < 1 \quad (\text{i.e., } w_{ii} + \sum_j |w_{ij}| < 1 \text{ for all } i)$$

$$\begin{aligned} \text{osLip}_\infty(\mathsf{F}_{\text{FR}}) &= \sup_{x,u} \mu_\infty(-I_n + (D\Phi(Wx + Bu))W) = -1 + \sup_{x,u} \mu_\infty(D\Phi(Wx + Bu)W) \\ &= -1 + \max_{d \in [0,1]^n} \mu_\infty(\text{diag}(d)W) \quad (\text{max convex polytope, } 2^n \text{ vertices}) \\ &= -1 + \max \{\mu_\infty(0), \mu_\infty(W)\} = -1 + \mu_\infty(W)_+ \end{aligned}$$

Example #6: Firing-rate network with symmetric synapses

$$\dot{x} = \mathsf{F}_{\text{FR}}(x) := -x + \Phi(Wx + Bu)$$

$$0 \leq \Phi'_i(y) \leq 1 \quad \text{and} \quad W = W^\top \text{ with } \lambda_W = \lambda_{\max}(W)$$

F_{FR} is infinitesimally contracting:

(for $\lambda_W < 0$)

with rate 1 wrt $\|\cdot\|_{2,(-W)^{1/2}}$

(for $\lambda_W = 0$)

with rate $\|\cdot\|_{2,Q_{\text{FR},\epsilon}}$, **for each** $\epsilon > 0$

(for $0 < \lambda_W < 1$)

with rate $1 - \lambda_W$ **wrt** $\|\cdot\|_{2,Q_{\text{FR},\lambda_W}}$

For $\lambda_W = 1$, F_{FR} is weakly infinitesimally contracting wrt $\|\cdot\|_{2,Q_{\text{FR},\lambda_W}}$

- $Q_{\text{FR},a} := Uh_a(\Lambda)U^\top \succ 0$, where $W = U\Lambda U^\top$ and $h_a(z) := 2a(1 + \sqrt{1 - z/a})$
- optimal rates
- proof based upon LMI calculations and Sylvester's law of inertia

Outline

1 History and resources

2 Basic definitions

- Discrete- and continuous-time dynamics on vector spaces
- Dynamics on Riemannian manifolds

3 Examples

- Optimization-based dynamics
- Recurrent neural network dynamics

4 Properties of contracting dynamics

- iISS
- Periodic systems
- Composite norms and interconnected systems
- Contractivity of delay dynamics
- Forward Euler theorem

5 Generalizations

6 Conclusions and future research

7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality

- Systems with invariance/conservation properties
- Induced seminorms and duality

8 Advanced Topics: Time-varying convex optimization via contracting dynamics

- Tracking equilibrium trajectories

Equilibrium and Lyapunov functions

Equilibria of contracting vector fields:

For a time-invariant F , c -strongly contracting wrt $\|\cdot\|$

- ① for each $t > 0$, t -flow of F is a contraction,
i.e., distance between solutions exponentially decreases with rate c
- ② there exists an equilibrium x^* , that is unique, globally exponentially stable with global Lyapunov functions

$$x \mapsto V_1(x) = \|x - x^*\|^2 \quad \text{and} \quad x \mapsto V_2(x) = \|\mathsf{F}(x)\|^2$$

For a time-invariant F ,

- ① $\text{osLip}(\mathsf{F}) = -c$ wrt ℓ_2 and $D\mathsf{F}(x) = D\mathsf{F}(x)^\top$ for all x ,
- ② for each scalar w ,

$$V_3(x) = - \int_0^1 x^\top \mathsf{F}(tx) dt + w$$

is c -strongly convex, is global Lyapunov, and $\text{grad}V_3(x) = -\mathsf{F}(x)$ for all x .

For time and input-dependent vector \mathbf{F} ,

$$\dot{x} = \mathbf{F}(t, x, u(t)), \quad x(0) = x_0 \in \mathcal{X}, \quad u(t) \in \mathcal{U} \quad (1)$$

Given norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{U}}$, assume constants $c, \ell > 0$ s.t.

- **osLip wrt x :** $\text{osLip}_x(\mathbf{F}) \leq -c < 0$, uniformly in t, u
- **Lip wrt u :** $\text{Lip}_u(\mathbf{F}) \leq \ell$, uniformly in t, x

Then

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

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

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

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

- ③ F has **incremental $\mathcal{L}_{\mathcal{X}, \mathcal{U}}^q$ gain equal to ℓ/c , for $q \in [1, \infty]$,**

$$\|x(\cdot) - y(\cdot)\|_{\mathcal{X}, q} \leq \frac{\ell}{c} \|u_x(\cdot) - u_y(\cdot)\|_{\mathcal{U}, q} \quad (\text{for } x_0 = y_0)$$

Given norm $\|\cdot\|_{\mathcal{X}}$ on \mathbb{R}^n (or $\|\cdot\|_{\mathcal{U}}$ on \mathbb{R}^k),

- $\mathcal{L}_{\mathcal{X}}^q$, $q \in [1, \infty]$, is vector space of continuous signals, $x : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$, with well-defined bounded norm

$$\|x(\cdot)\|_{\mathcal{X},q} = \begin{cases} \left(\int_0^\infty \|x(t)\|_{\mathcal{X}}^q dt \right)^{1/q} & \text{if } q \in [1, \infty[\\ \sup_{t \geq 0} \|x(t)\|_{\mathcal{X}} & \text{if } q = \infty \end{cases} \quad (2)$$

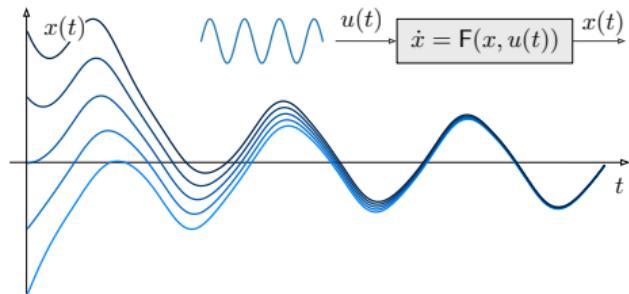
- Input-state system has $\mathcal{L}_{\mathcal{X},\mathcal{U}}^q$ -induced gain upper bounded by $\gamma > 0$ if, for all $u \in \mathcal{L}_{\mathcal{U}}^q$, the state x from zero initial state satisfies

$$\|x(\cdot)\|_{\mathcal{X},q} \leq \gamma \|u(\cdot)\|_{\mathcal{U},q} \quad (3)$$

From time-invariant to periodic systems

For time-varying vector field F and norm $\|\cdot\|$

- ① $\text{osLip}_x(F) \leq -c < 0$
- ② F is T -periodic

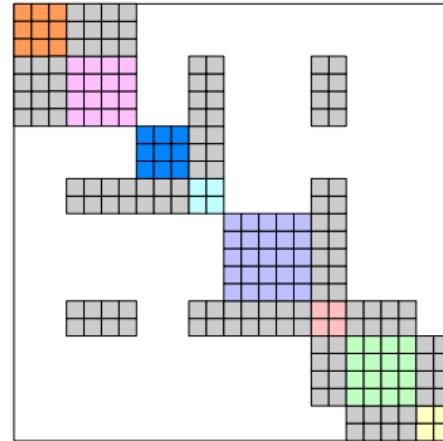
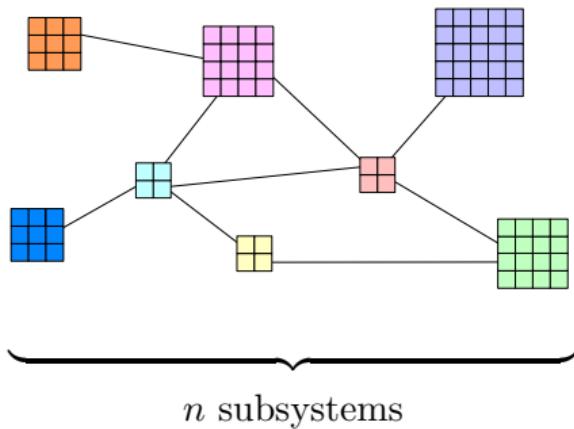


Then

- ① there exists a unique periodic solution $x^* : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$ with period T
- ② for every initial condition x_0 ,

$$\|x(t, x_0) - x^*(t)\| \leq e^{-ct} \|x_0 - x^*(0)\| \quad (4)$$

Composite norms



- ① n local norms $\|\cdot\|_i$ on \mathbb{R}^{N_i}
- ② an aggregating norm $\|\cdot\|_{\text{agg}}$ on \mathbb{R}^n
- ③ composite norm

G. Russo, M. Di Bernardo, and E. D. Sontag. A contraction approach to the hierarchical analysis and design of networked systems. *IEEE Transactions on Automatic Control*, 58(5):1328–1331, 2013. [doi](#)

Networks of contracting systems

Interconnected subsystems: $x_i \in \mathbb{R}^{N_i}$ and $x_{-i} \in \mathbb{R}^{N-N_i}$:

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

Network contraction theorem

- **osLip wrt x_i :** $\text{osLip}_{x_i}(F_i) \leq -c_i$, uniformly in x_{-i}
- **Lip wrt to x_j :** $\text{Lip}_{x_j}(F_i) \leq \ell_{ij}$, uniformly in x_{-j}

- 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

The network science of Metzler Hurwitz matrices

$\begin{bmatrix} -c_1 & \dots & \ell_{1n} \\ \vdots & & \vdots \\ \ell_{n1} & \dots & -c_n \end{bmatrix}$ is **Metzler** (so that Perron-Frobenius Theorem applies)

Hurwitzness depends upon both topology and edge weights

- directed acyclic interconnections of contracting systems are strongly contracting
- For $n = 2$, Hurwitz if and only if **small gain condition**

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

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

Hurwitz Metzler Theorem

- ① M is Hurwitz,
- ② there exists $\eta \in \mathbb{R}_{>0}^n$ such that $\eta^\top M < 0_n^\top$ or, equivalently, $\mu_{1,[\eta]}(M) < 0$,
- ③ there exists $\xi \in \mathbb{R}_{>0}^n$ such that $M\xi < 0_n$ or, equivalently, $\mu_{\infty,[\xi]^{-1}}(M) < 0$, and
- ④ there exists a diagonal $P = P^\top \succ 0$ satisfying $M^\top P + PM \prec 0$ or, equivalently,
 $\mu_{2,P^{1/2}}(M) < 0$.

Input: a Metzler matrix $M \in \mathbb{R}^{n \times n}$

Output: polynomials $\{\gamma_{\mathcal{C}_2}, \dots, \gamma_{\mathcal{C}_n}\}$ in entries of M

- 1: $\mathcal{C} :=$ set of simple cycles of digraph associated to M
- 2: $\gamma_\phi :=$ gain of cycle $\phi \in \mathcal{C}$
- 3: **for** i from 2 to n
- 4: $\mathcal{C}_i :=$ cycles in \mathcal{C} passing through only nodes $1, \dots, i$
- 5: $\gamma_{\mathcal{C}_i} := \sum_{\substack{\phi \in \mathcal{C}_i \\ \phi \perp \psi}} \gamma_\phi - \sum_{\substack{\phi, \psi \in \mathcal{C}_i \\ \phi \perp \psi}} \gamma_\phi \gamma_\psi + \sum_{\substack{\phi, \psi, \rho \in \mathcal{C}_i \\ \phi \perp \psi, \phi \perp \rho, \psi \perp \rho}} \gamma_\phi \gamma_\psi \gamma_\rho - \dots$

Network small-gain theorem for Metzler matrices

$$\text{Metzler } M \text{ is Hurwitz} \iff \gamma_{\mathcal{C}_2} < 1, \dots, \gamma_{\mathcal{C}_n} < 1$$

- not unique: distinct/equivalent conditions after renumbering, redundancy
- computational efficiency: after precomputation of simple cycles

$$M = \begin{bmatrix} -c_1 & 0 & 0 & \ell_{14} \\ 0 & -c_2 & \ell_{23} & \ell_{24} \\ 0 & \ell_{32} & -c_3 & \ell_{34} \\ \ell_{41} & \ell_{42} & \ell_{43} & -c_4 \end{bmatrix}$$

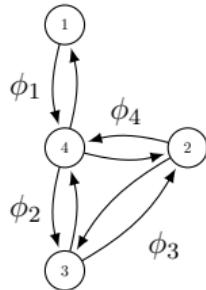


Figure: associated digraph and simple cycles

- $\gamma_{\phi_1} = \frac{\ell_{14}\ell_{41}}{c_1 c_4}$, $\gamma_{\phi_2} = \frac{\ell_{34}\ell_{43}}{c_3 c_4}$, $\gamma_{\phi_3} = \frac{\ell_{23}\ell_{32}}{c_2 c_3}$, and $\gamma_{\phi_4} = \frac{\ell_{24}\ell_{42}}{c_2 c_4}$
- $\mathcal{C}_2 = \emptyset$
- $\mathcal{C}_3 = \{\phi_3\}$: $\gamma_{\mathcal{C}_3} = \gamma_{\phi_3} < 1$ (redundant)
- $\mathcal{C}_4 = \{\phi_1, \dots, \phi_4\}$: $\gamma_{\mathcal{C}_4} = \sum_{i=1}^4 \gamma_{\phi_i} - \gamma_{\phi_1} \gamma_{\phi_3} < 1$

$$\begin{bmatrix} -c_1 & 0 & 0 & 0 & \ell_{15} & \ell_{16} \\ 0 & -c_2 & 0 & \ell_{24} & \ell_{25} & 0 \\ 0 & 0 & -c_3 & \ell_{34} & 0 & \ell_{36} \\ 0 & \ell_{42} & \ell_{43} & -c_4 & 0 & 0 \\ \ell_{51} & \ell_{52} & 0 & 0 & -c_5 & 0 \\ \ell_{61} & 0 & \ell_{63} & 0 & 0 & -c_6 \end{bmatrix}$$

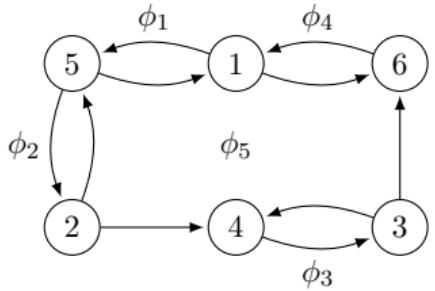


Figure: associated digraph and simple cycles

- $\mathcal{C}_2, \mathcal{C}_3$ empty
- $\mathcal{C}_4 = \{\phi_3\}$: $\gamma_3 < 1$ (redundant)
- $\mathcal{C}_5 = \{\phi_1, \phi_2, \phi_3\}$: $\gamma_{\mathcal{C}_5} = \gamma_1 + \gamma_2 + \gamma_3 - \gamma_1\gamma_3 - \gamma_2\gamma_3 < 1$
- $\mathcal{C}_6 = \{\phi_1, \dots, \phi_5\}$: $\gamma_{\mathcal{C}_6} = \sum_{i=1}^5 \gamma_i - \gamma_1\gamma_3 - \gamma_2\gamma_3 - \gamma_3\gamma_4 - \gamma_2\gamma_4 + \gamma_2\gamma_3\gamma_4 < 1$

Incremental ISS for strongly contracting delay ODEs

$$\dot{x}(t) = F(x(t), x(t-s), u(t)), 0 \leq s \leq S, \quad \|\cdot\|_{\mathcal{X}}, \|\cdot\|_{\mathcal{U}} \quad (5)$$

assume there exist positive constants $c, \ell_{\mathcal{U}}, \ell_{\mathcal{X}}$ such that, for all variables,

$$\text{osL } x : \quad \|F(x, d, u) - F(y, d, u), x - y\|_{\mathcal{X}} \leq -c\|x - y\|_{\mathcal{X}}^2 \quad (6)$$

$$\text{Lip } x(t-s) : \quad \|F(x, x_1, u) - F(x, x_2, u)\|_{\mathcal{X}} \leq \ell_{\mathcal{X}}\|x_1 - x_2\|_{\mathcal{X}} \quad (7)$$

$$\text{Lip } u : \quad \|F(x, d, u) - F(x, d, v)\|_{\mathcal{X}} \leq \ell_{\mathcal{U}}\|u - v\|_{\mathcal{U}} \quad (8)$$

By the curve norm derivative formula, subadditivity, and Cauchy-Schwarz inequality,

$$\begin{aligned} \|x(t) - y(t)\|_{\mathcal{X}} D^+ \|x(t) - y(t)\|_{\mathcal{X}} &= [\![F(x(t), x(t-s), u_x(t)) - F(y(t), y(t-s), u_y(t)), x(t) - y(t)]\!]_{\mathcal{X}} \\ &\leq [\![F(x(t), x(t-s), u_x(t)) - F(y(t), x(t-s), u_x(t)), x(t) - y(t)]\!]_{\mathcal{X}} \\ &\quad + [\![F(y(t), x(t-s), u_x(t)) - F(y(t), y(t-s), u_x(t)), x(t) - y(t)]\!]_{\mathcal{X}} \\ &\quad + [\![F(y(t), y(t-s), u_x(t)) - F(y(t), y(t-s), u_y(t)), x(t) - y(t)]\!]_{\mathcal{X}} \\ &\leq -c\|x(t) - y(t)\|_{\mathcal{X}}^2 + \ell_{\mathcal{X}}\|x(t-s) - y(t-s)\|_{\mathcal{U}}\|x(t) - y(t)\|_{\mathcal{X}}, \\ &\quad + \ell_{\mathcal{U}}\|u_x(t) - u_y(t)\|_{\mathcal{U}}\|x(t) - y(t)\|_{\mathcal{X}}. \end{aligned}$$

Thus, with $m(t) = \|x(t) - y(t)\|_{\mathcal{X}}$, delay differential inequality:

$$D^+ m(t) \leq -cm(t) + \ell_{\mathcal{X}} \sup_{0 \leq s \leq S} m(t-s) + \ell_{\mathcal{U}}\|u_x(t) - u_y(t)\|_{\mathcal{U}}, \quad (9)$$

Halanay inequality is applicable. If $c > \ell_{\mathcal{X}}$, then

$$m(t) \leq m_0 e^{-\rho(t-t_0)} + \ell_{\mathcal{U}} \int_{t_0}^t e^{-\rho(t-\tau)} \|u_x(\tau) - u_y(\tau)\|_{\mathcal{U}} d\tau, \quad (10)$$

where $\rho > 0$ is the unique positive root of $\rho = c - \ell_{\mathcal{X}} e^{\rho S}$ and $m_0 = \sup_{0 \leq s \leq S} m(t_0 - s)$.

Networks of contracting systems with time delays

Interconnected subsystems $i \in \{1, \dots, n\}$

$$\dot{x}_i = F_i(x_i, x_{-i}, x_{-i}(t-s), u_i), \quad 0 \leq s \leq S, \quad \|\cdot\|_i, \|\cdot\|_{i,\mathcal{U}} \quad (11)$$

Assume there exist positive constants st

osL x_i : $\llbracket F_i(x_i, \dots) - F_i(y_i, \dots), x_i - y_i \rrbracket_i \leq -c_i \|x_i - y_i\|_i^2$

Lip x_{-i} : $\|F_i(\dots, x_{-i}, \dots) - F_i(\dots, y_{-i}, \dots)\|_i \leq \sum_{j=1, j \neq i}^n \gamma_{ij} \|x_j - y_j\|_j$

Lip x_{-1}^{-s} : $\|F_i(\dots, x_{-i}^{-s}, \dots) - F_i(\dots, y_{-i}^{-s}, \dots)\|_i \leq \sum_{j=1, j \neq i}^n \hat{\gamma}_{ij} \|x_j^{-s} - y_j^{-s}\|_j$

Lip u_i : $\|F_i(\dots, u_i) - F_i(\dots, v_i)\|_i \leq \ell_{i,\mathcal{U}} \|u_i - v_i\|_{i,\mathcal{U}}$

With $m_i(t) = \|x_i(t) - y_i(t)\|_i$, delay differential inequality:

$$D^+ m(t) \leq -Cm(t) + \Gamma m(t) + \widehat{\Gamma} \sup_{0 \leq s \leq S} m(t-s) + \ell_{i,\mathcal{U}} \|u_x(t) - u_y(t)\|_{\mathcal{U}}$$

and, if the Metzler matrix $-C + \Gamma + \widehat{\Gamma}$ is Hurwitz, then (11) is incremental ISS

Forward Euler theorem

Forward Euler theorem for contracting dynamics

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

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

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

- ① $\text{Id} + \alpha F$ is contracting for

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

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

$$\begin{aligned}\alpha^* &= \frac{1}{c} \left(\frac{1}{2\kappa^2} - \frac{3}{8\kappa^3} + \mathcal{O}\left(\frac{1}{\kappa^4}\right) \right) \\ \ell^* &= 1 - \frac{1}{4\kappa^2} + \frac{1}{8\kappa^3} + \mathcal{O}\left(\frac{1}{\kappa^4}\right)\end{aligned}$$

Improved bounds for inner-product norms

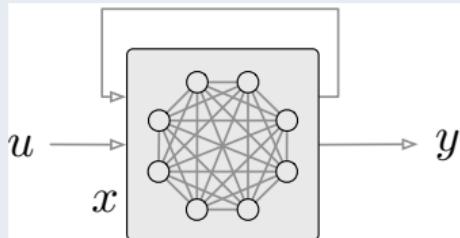
- ① the map $\text{Id} + \alpha F$ is a contraction map wrt $\|\cdot\|_{2,P^{1/2}}$ for

$$0 < \alpha < \frac{2}{c\kappa^2}$$

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

$$\alpha_E^* = \frac{1}{c\kappa^2} \quad \ell_E^* = 1 - \frac{1}{2\kappa^2} + \mathcal{O}\left(\frac{1}{\kappa^4}\right)$$

Application: ℓ_∞ -contracting neural networks



$$\dot{x} = -x + \Phi(Ax + Bu + b)$$

(*recurrent NN*)

$$x = \Phi(Ax + Bu + b)$$

(*implicit NN*)

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

(*forward Euler*)

If

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

- recurrent NN is contracting with rate $1 - \mu_\infty(A)_+$
- implicit NN is well posed
- forward Euler is contracting with factor $1 - \frac{1 - \mu_\infty(A)_+}{1 - \min_i(a_{ii})_-}$
- input-state Lipschitz constant $\text{Lip}_{u \rightarrow x} = \frac{\|B\|_\infty}{1 - \mu_\infty(A)_+}$

$$\text{at } \alpha^* = \frac{1}{1 - \min_i(a_{ii})_-}$$

Outline

1 History and resources

2 Basic definitions

- Discrete- and continuous-time dynamics on vector spaces
- Dynamics on Riemannian manifolds

3 Examples

- Optimization-based dynamics
- Recurrent neural network dynamics

4 Properties of contracting dynamics

- iISS
- Periodic systems
- Composite norms and interconnected systems
- Contractivity of delay dynamics
- Forward Euler theorem

5 Generalizations

6 Conclusions and future research

7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality

- Systems with invariance/conservation properties
- Induced seminorms and duality

8 Advanced Topics: Time-varying convex optimization via contracting dynamics

- Tracking equilibrium trajectories

From nominal to uncertain systems

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

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

Assume:

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

Then

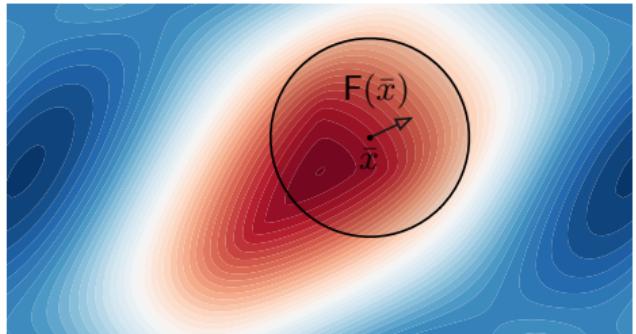
- ① $F + \Delta$ is strongly contracting with rate $c - d$
- ② the unique equilibria x_F^* of F and $x_{F+\Delta}^*$ of $F + \Delta$ satisfy

$$\|x_F^* - x_{F+\Delta}^*\| \leq \frac{\|\Delta(x_F^*)\|}{c - d}$$

From global to local contractivity

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

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



Assume:

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

Then

- ① B is forward invariant
- ② $F|_B$ is strongly infinitesimally contracting

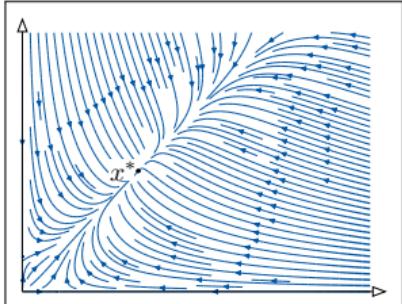
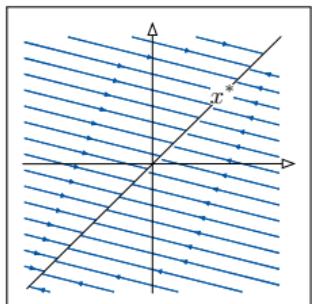
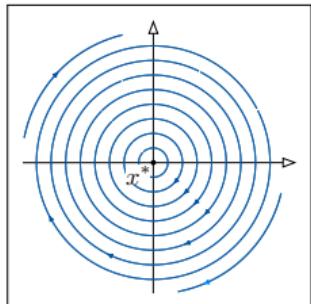
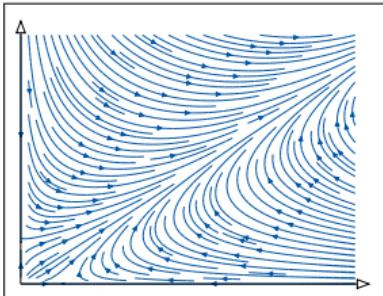
From strongly to weakly contracting systems

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

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

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



- ① **Lotka-Volterra population dynamics** (Lotka, 1920; Volterra, 1928):
 ℓ_1 -weakly contracting (after a rescaling change of coordinates)
- ② **Matrosov-Bellman interconnected stable systems** (Bellman, 1962; Matrosov, 1962):
strongly contracting wrt composite norm
- ③ **Kuramoto coupled oscillators** (Kuramoto, 1975):
strongly semicontracting wrt (ℓ_2, Π_n) norm, in neighb'd of each phase-cohesive equilibrium
- ④ **Yorke multigroup SIS epidemic model** (Lajmanovich and Yorke, 1976):
equilibrium contracting wrt weighted ℓ_1/ℓ_∞ norms (at disease-free and endemic eq.)
- ⑤ **Hopfield and cellular neural networks** (Hopfield, 1982):
 ℓ_1/ℓ_∞ -strongly contracting
- ⑥ **Daganzo cell transmission model for traffic networks** (Daganzo, 1994):
 ℓ_1 -weakly contracting, when the dynamics is monotone
- ⑦ **Chua's diffusively-coupled dynamical systems** (Wu and Chua, 1995):
strongly semi-contracting wrt $(2, p)$ tensor norm on $\mathbb{R}^n \otimes \mathbb{R}^k$
- ⑧ ...

Outline

1 History and resources

2 Basic definitions

- Discrete- and continuous-time dynamics on vector spaces
- Dynamics on Riemannian manifolds

3 Examples

- Optimization-based dynamics
- Recurrent neural network dynamics

4 Properties of contracting dynamics

- iISS
- Periodic systems
- Composite norms and interconnected systems
- Contractivity of delay dynamics
- Forward Euler theorem

5 Generalizations

6 Conclusions and future research

7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality

- Systems with invariance/conservation properties
- Induced seminorms and duality

8 Advanced Topics: Time-varying convex optimization via contracting dynamics

- Tracking equilibrium trajectories

contractivity = robust computationally-friendly stability

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

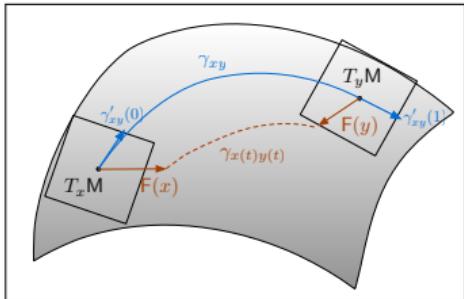


	Lyapunov Theory	Contraction Theory for Dynamical Systems
existence of equilibrium Lyapunov function inputs	F admits global Lyapunov function assumed arbitrary ISS via \mathcal{KL} and \mathcal{L} functions	F is strongly contracting implied + computational methods $\ x - x^*\ $ and $\ F(x)\ $ iISS via explicit constants

search for contraction properties
design engineering systems to be contracting

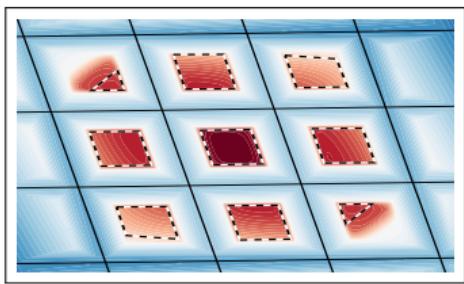
Theoretical frontiers

- higher order contraction
- relationship with monotone operator theory
- metric spaces
- computational methods



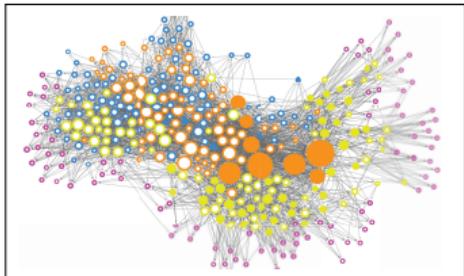
Limitations: not all stable systems are contractive:

- Lyapunov-diagonally-stable networks
- multistable and locally contracting systems
- biochemical networks
- control contraction design



Application to control and learning

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



Outline

- 1 History and resources
- 2 Basic definitions
 - Discrete- and continuous-time dynamics on vector spaces
 - Dynamics on Riemannian manifolds
- 3 Examples
 - Optimization-based dynamics
 - Recurrent neural network dynamics
- 4 Properties of contracting dynamics
 - iISS
 - Periodic systems
 - Composite norms and interconnected systems
 - Contractivity of delay dynamics
 - Forward Euler theorem
- 5 Generalizations
- 6 Conclusions and future research
- 7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality
 - Systems with invariance/conservation properties
 - Induced seminorms and duality
- 8 Advanced Topics: Time-varying convex optimization via contracting dynamics
 - Tracking equilibrium trajectories



Consider a vector field $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and let $\xi, \eta \in \mathbb{R}^n$.

- **Invariance property:** for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$,

$$\mathbf{F}(x + \alpha\xi) = \mathbf{F}(x) \quad \text{or equivalently} \quad D\mathbf{F}(x)\xi = \mathbb{0}_n$$

- **Conservation property:** for all $x, y \in \mathbb{R}^n$,

$$\eta^\top \mathbf{F}(x) = \eta^\top \mathbf{F}(y) \quad \text{or equivalently} \quad \eta^\top D\mathbf{F}(x) = \mathbb{0}_n^\top$$

Let $A \in \mathbb{R}^{n \times n}$ be row-stochastic: $A\mathbb{1}_n = \mathbb{1}_n$ and $A \geq 0$

Averaging Systems

$$x_{k+1} = Ax_k$$

Invariance: dynamics unaffected by translations in $\text{span}\{\mathbb{1}_n\}$

Examples: distributed optimization, robotic coordination, frequency synchronization, ...

Dynamical Flow Systems

$$x_{k+1} = A^\top x_k$$

Conservation: quantity $\mathbb{1}_n^\top x$ is constant

Examples: compartmental models, Markov chains

Historical starting point

Given row-stochastic $A \in \mathbb{R}^{n \times n}$,

Markov-Dobrushin ergodic coefficient

$$\tau_1(A) = \max_{\|z\|_1=1, \mathbf{1}_n^\top z=0} \|A^\top z\|_1$$

$\tau_1(A) < 1$ under mild connectivity conditions

$\tau_p(A)$ also defined for general $p \in [1, \infty]$

How is τ_1 an induced norm?



A. A. Markov. Extensions of the law of large numbers to dependent quantities. *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 15, 1906. (in Russian)

R. L. Dobrushin. Central limit theorem for nonstationary Markov chains. I. *Theory of Probability & Its Applications*, 1(1):65–80, 1956. 

$$A \in \mathbb{R}^{n \times n} \text{ row-stochastic}$$

Classical Property of Averaging Systems $x_{k+1} = Ax_k$

Given $x \in \mathbb{R}^n$, max-min disagreement:

$$s(Ax) \leq \tau_1(A) s(x), \quad \text{where } s(x) = \max_i \{x_i\} - \min_j \{x_j\}$$

Classical Property of Markov Chains $x_{k+1} = A^\top x_k$

Given π, σ in the simplex Δ_n , total variation distance:

$$d_{\text{TV}}(A^\top \pi, A^\top \sigma) \leq \tau_1(A) d_{\text{TV}}(\pi, \sigma), \quad \text{where } d_{\text{TV}}(\pi, \sigma) = \frac{1}{2} \sum_i |\pi_i - \sigma_i|$$

Why is the same τ_1 relevant in both cases?

A **seminorm** is a function $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ s.t., $\forall a \in \mathbb{R}$ and $\forall x, y \in \mathbb{R}^n$:

- ① (*homogeneity*): $\|ax\| = |a|\|x\|$
- ② (*subadditivity*): $\|x + y\| \leq \|x\| + \|y\|$

The *kernel* is the vector space:

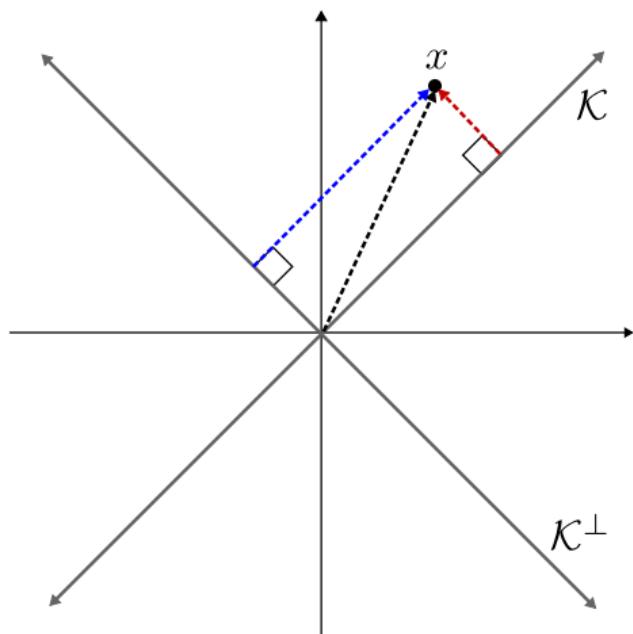
$$\mathcal{K} = \{x \in \mathbb{R}^n : \|x\| = 0\}$$

We focus on *consensus seminorms*, where $\mathcal{K} = \text{span}\{\mathbf{1}_n\}$.

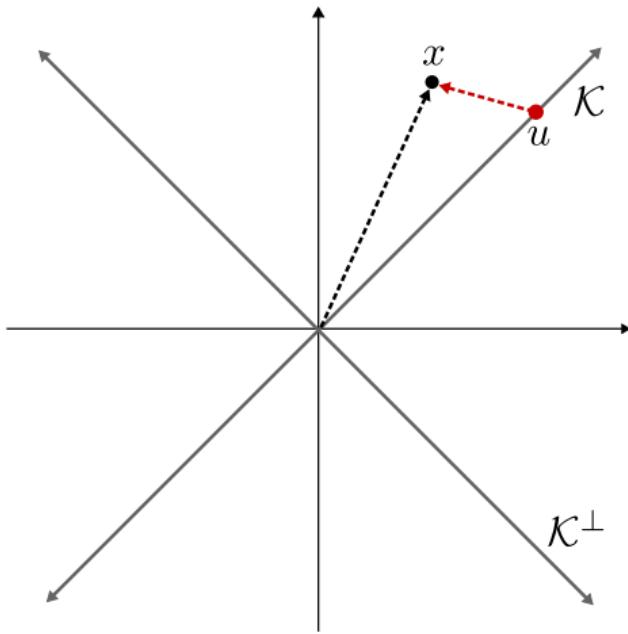
Note: $\|\cdot\|$ is invariant under translations in \mathcal{K}

Projection and distance-based seminorms: graphical definitions

Projection seminorms



Distance seminorms



$$\|x\|_{\text{proj},p} \triangleq \|\Pi_\perp x\|_p, \quad \Pi_\perp = \Pi_\perp^\top$$

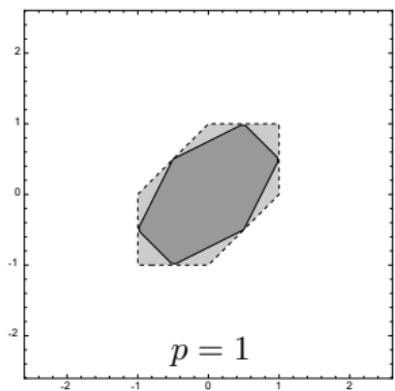
$$\|x\|_{\text{dist},p} \triangleq \min_{u \in \mathcal{K}} \|x - u\|_p$$

Consensus seminorms

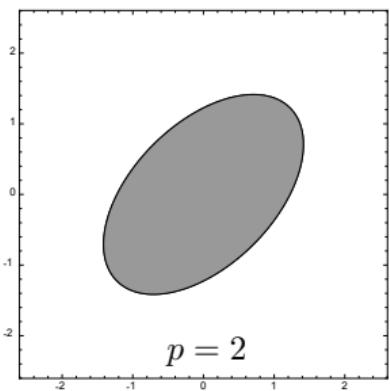
When $\mathcal{K} = \text{span}\{\mathbb{1}_n\}$, **consensus seminorms**

	$\ x\ _{\text{proj},p}$	$\ x\ _{\text{dist},p}$
ℓ_1	$\sum_{i=1}^n x_i - x_{\text{avg}} $	$\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} x_{(i)} - \sum_{j=\lceil \frac{n}{2} \rceil + 1}^n x_{(j)}$
ℓ_2	$\sqrt{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$	$\sqrt{\frac{1}{n} \sum_{i,j} (x_i - x_j)^2}$
ℓ_∞	$\max_i x_i - x_{\text{avg}} $	$\frac{1}{2} (x_{(1)} - x_{(n)})$

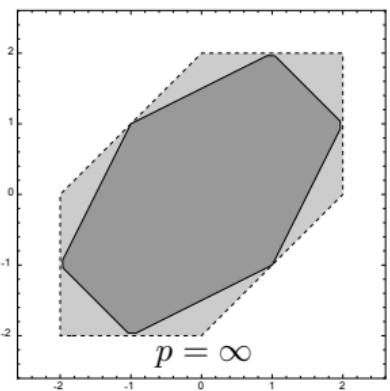
where we have sorted $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$



$$p = 1$$



$$p = 2$$



$$p = \infty$$

Figure: Two-dimensional sections of three-dimensional unit disks of projection (solid contours) and distance (dashed contours) consensus seminorms. We plot the sections corresponding to $(x_1, x_2, x_3 = 0)$.

Induced matrix seminorms

Consider a seminorm $\|\cdot\|$ on \mathbb{R}^n with kernel \mathcal{K} .

Induced matrix seminorm: function $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}_{\geq 0}$ where

$$\|A\| = \max_{\substack{\|x\| \leq 1 \\ x \perp \mathcal{K}}} \|Ax\|, \quad \forall A \in \mathbb{R}^{n \times n}$$



In general, $\|Ax\| \not\leq \|A\|\|x\|$
Inequality is true if $x \in \mathcal{K}^\perp$ or $A\mathcal{K} \subseteq \mathcal{K}$

Properties of dual and induced norms

- ① ℓ_p and ℓ_q norms are dual, for $1/p + 1/q = 1$

$$\|\cdot\|_p = (\|\cdot\|_q)_\star \quad \|\cdot\|_q = (\|\cdot\|_p)_\star$$

- ② dual norm satisfies (sharp) *Hölder inequality*: $x^\top y \leq \|x\|_p \|y\|_q$
- ③ equality between dual induced norms: $\|A\|_p = \|A^\top\|_q$
- ④ induced norm is submultiplicative: $\|AB\| \leq \|A\| \|B\|$

Key facts about dual and induced seminorms

Properties of dual and induced seminorms

- ① ℓ_p -distance and ℓ_q -projection seminorms are **dual**, for $1/p + 1/q = 1$

$$\|\cdot\|_{\text{dist},p} = (\|\cdot\|_{\text{proj},q})_* \quad \|\cdot\|_{\text{proj},q} = (\|\cdot\|_{\text{dist},p})_*$$

- ② dual seminorm satisfies (sharp) **Markov inequality**: $x^\top \Pi_\perp y \leq \|x\|_{\text{dist},p} \|y\|_{\text{proj},q}$
- ③ equality between dual induced seminorms: $\|A\|_{\text{dist},p} = \|A^\top\|_{\text{proj},q}$
- ④ induced seminorm is submultiplicative: $\|AB\| \leq \|A\| \|B\|$ if $A\mathcal{K} \subseteq \mathcal{K}$ or $B\mathcal{K}^\top \subseteq \mathcal{K}^\top$

Ergodic coefficients are induced seminorms

$$\|A\|_{\text{dist},p} = \|A^\top\|_{\text{proj},q} = \tau_q(A) := \max_{\|z\|_q=1, z \perp \mathbb{1}_n} \|A^\top z\|_q$$

How Markov and Banach's results meet

Classical Property of Averaging Systems

Given row-stochastic $A \in \mathbb{R}^{n \times n}$ and $x, y \in \mathbb{R}^n$:

$$\begin{aligned}\|A(x - y)\|_{\text{dist},\infty} &\leq \tau_1(A) \|x - y\|_{\text{dist},\infty} \\ &= \|A\|_{\text{dist},\infty} \|x - y\|_{\text{dist},\infty}\end{aligned}$$

Classical Property of Markov Chains

Given row-stochastic $A \in \mathbb{R}^{n \times n}$ and π, σ in the simplex Δ_n :

$$\begin{aligned}\|A^\top(\pi - \sigma)\|_{\text{proj},1} &\leq \tau_1(A) \|\pi - \sigma\|_{\text{proj},1} \\ &= \|A^\top\|_{\text{proj},1} \|\pi - \sigma\|_{\text{proj},1}\end{aligned}$$

Summary and future work

- ① ergodic coefficients are contraction factors
- ② duality explains their roles in both averaging and flow systems
- ③ nonEuclidean norms play a key role
- ④ **semicontraction theory**
 - ① discrete/continuous-time Markov chains
 - ② discrete/continuous-time nonlinear consensus algorithms
 - ③ local contractivity of Kuramoto and Kuramoto-Sakaguchi models

Future work

consider the set of undirected, unweighted connected graphs + selfloops
for each adjacency A_i , define row-stochastic $\mathcal{A}_i = \text{diag}(A_i \mathbb{1}_n)^{-1} A_i$ (equal neighbor)
find a consensus seminorm $\|\cdot\|$ such that, for each i ,

$$\|\mathcal{A}_i\| < 1$$

or **prove** that it does not exist

Continuous-time semicontraction theory

The *induced log seminorm* of $A \in \mathbb{R}^{n \times n}$ is

$$\mu_{\|\cdot\|}(A) \triangleq \lim_{h \rightarrow 0^+} \frac{\|I_n + hA\| - 1}{h}$$

Laplacian L , corresponding to weighted digraph with adj. matrix A :

$$\mu_{\text{dist},1}(-L) = -\min_j \left\{ (d_{\text{out}})_j - \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor - 1} a_{(i),j} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n-1} a_{(i),j} \right\}, \quad d_{\text{out}} = A\mathbb{1}_n$$

$$\mu_{\text{dist},2}(-L) = \min \left\{ b : \Pi_{\perp} L + L^T \Pi_{\perp} \succeq -2b\Pi_{\perp} \right\}, \quad \Pi_{\perp} = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^T$$

$$\mu_{\text{dist},\infty}(-L) = -\min_{i \neq j} \left\{ a_{ij} + a_{ji} + \sum_{k \neq i,j} \min\{a_{ik}, a_{jk}\} \right\}$$

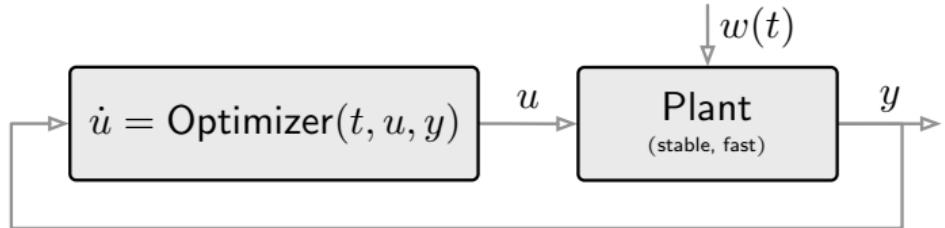
Let $p, q \in [1, \infty]$ such that $p^{-1} + q^{-1} = 1$. For any matrix $M \in \mathbb{R}^{n \times n}$, and any kernel \mathcal{K} ,

$$\mu_{\text{dist},p}(M) = \mu_{\text{proj},q}(M^T)$$

Outline

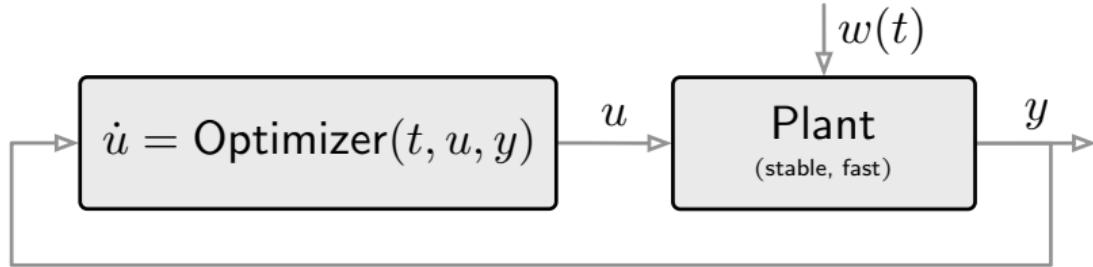
- 1 History and resources
- 2 Basic definitions
 - Discrete- and continuous-time dynamics on vector spaces
 - Dynamics on Riemannian manifolds
- 3 Examples
 - Optimization-based dynamics
 - Recurrent neural network dynamics
- 4 Properties of contracting dynamics
 - iISS
 - Periodic systems
 - Composite norms and interconnected systems
 - Contractivity of delay dynamics
 - Forward Euler theorem
- 5 Generalizations
- 6 Conclusions and future research
- 7 Advanced Topics: Semicontractivity, ergodic coefficients, and duality
 - Systems with invariance/conservation properties
 - Induced seminorms and duality
- 8 Advanced Topics: Time-varying convex optimization via contracting dynamics
 - Tracking equilibrium trajectories

Solving optimization problems via dynamical systems



- studies in linear and nonlinear programming (Arrow, Hurwicz, and Uzawa 1958)
- neural networks (Hopfield and Tank 1985) and analog circuits (Kennedy and Chua 1988)
- optimization on manifolds (Brockett 1991)
- ...
- power grids (Bolognani, Carli, Cavraro, Zampieri 2013)
- online and dynamic feedback optimization (Dall'Anese, Dörfler, Simonetto, ...)

Example: Time-varying optimization algorithms



optimization via dynamical systems

online time-varying optimization, optimization-based feedback control, ...

$$\begin{cases} \min & \text{cost}_1(u) + \text{cost}_2(y) \\ \text{s.t.} & y = \text{Plant}(u, w(t)) \end{cases} \implies \begin{cases} \dot{u} = \text{Optimizer}(t, u, y) \\ y = \text{Plant}(u, w(t)) \end{cases}$$

From convex optimization to contracting dynamics – time-varying

Many convex optimization problems can be solved with contracting dynamics

$$\dot{x} = F(x, \theta)$$

	Convex Optimization	Contracting Dynamics
Unconstrained	$\min_{x \in \mathbb{R}^n} f(x, \theta)$	$\dot{x} = -\nabla_x f(x, \theta)$
Constrained	$\min_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \quad x \in \mathcal{X}(\theta)}} f(x, \theta)$	$\dot{x} = -x + \text{Proj}_{\mathcal{X}(\theta)}(x - \gamma \nabla_x f(x, \theta))$
Composite	$\min_{x \in \mathbb{R}^n} f(x, \theta) + g(x, \theta)$	$\dot{x} = -x + \text{prox}_{\gamma g_\theta}(x - \gamma \nabla_x f(x, \theta))$
Equality	$\min_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \quad Ax = b(\theta)}} f(x, \theta)$	$\begin{aligned} \dot{x} &= -\nabla_x f(x, \theta) - A^\top \lambda, \\ \dot{\lambda} &= Ax - b(\theta) \end{aligned}$
Inequality	$\min_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \quad Ax \leq b(\theta)}} f(x, \theta)$	$\begin{aligned} \dot{x} &= -\nabla f(x, \theta) - A^\top \nabla M_{\gamma, b(\theta)}(Ax + \gamma \lambda), \\ \dot{\lambda} &= \gamma(-\lambda + \nabla M_{\gamma, b(\theta)}(Ax + \gamma \lambda)) \end{aligned}$

Tracking equilibrium trajectories

For parameter-dependent vector field $\mathbf{F} : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$ and differentiable $\theta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$

$$\dot{x}(t) = \mathbf{F}(x(t), \theta(t))$$

Assume there exist norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\Theta}$ s.t.

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

Theorem: Incremental ISS any two soltns: $x(t)$ with input u_x and $y(t)$ with input u_y

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

Tracking equilibrium trajectories

For parameter-dependent vector field $\mathbf{F} : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$ and differentiable $\theta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$

$$\dot{x}(t) = \mathbf{F}(x(t), \theta(t))$$

Assume there exist norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\Theta}$ s.t.

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

Theorem: Equilibrium tracking for contracting dynamics

- ① for each fixed θ , there exists a unique equilibrium $x^*(\theta)$
- ② the equilibrium map $x^*(\cdot)$ is Lipschitz with constant $\frac{\ell}{c}$
- ③ $D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$

Consequences for tracking error

$$D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq -c \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \|\dot{\theta}(t)\|_{\Theta}$$

- bounded input, bounded error
with asymptotic bound:

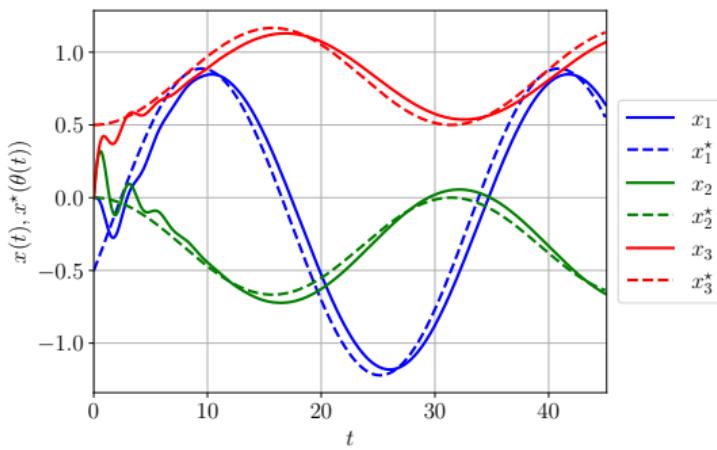
$$\limsup_{t \rightarrow \infty} \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} \leq \frac{\ell}{c^2} \limsup_{t \rightarrow \infty} \|\dot{\theta}(t)\|_{\Theta}$$

- bounded energy input, bounded energy error
- vanishing input, vanishing error
- exponentially vanishing input, exponentially vanishing error
- periodic input, periodic error

Numerical simulations

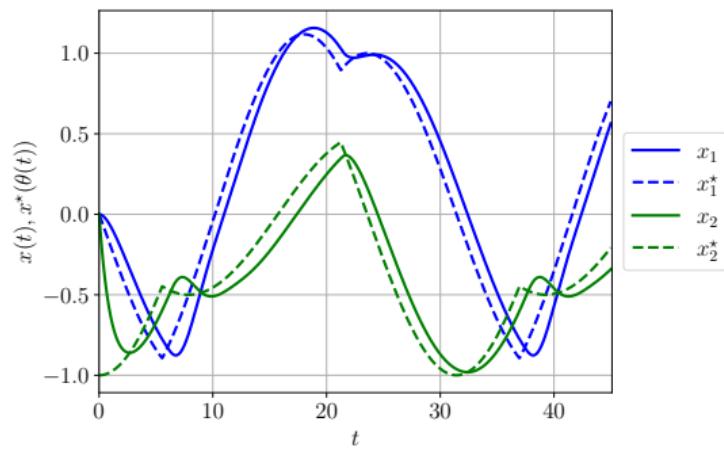
$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & \frac{1}{2} \|x - r(t)\|_2^2 \\ \text{s.t.} \quad & x_1 + 2x_2 + x_3 = \sin(\omega t), \end{aligned}$$

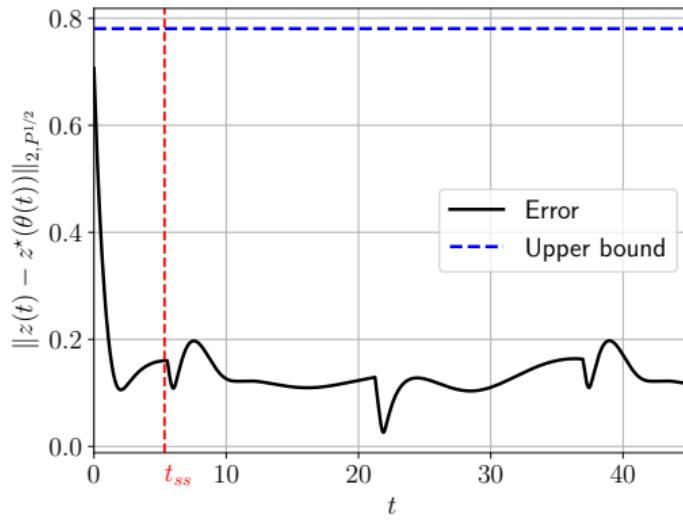
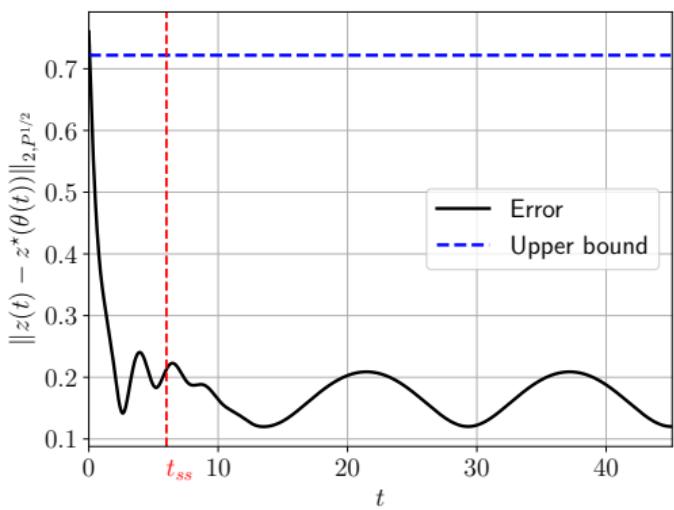
$$r(t) = (\sin(\omega t), \cos(\omega t), 1), \omega = 0.2$$



$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & \frac{1}{2} \|x + r(t)\|_2^2 \\ \text{s.t.} \quad & -x_1 + x_2 \leq \cos(\omega t), \end{aligned}$$

$$r(t) = (\sin(\omega t), \cos(\omega t)), \omega = 0.2$$





Proof sketch for equilibrium tracking

Given $\dot{x} = F(x, \theta(t))$ with $\text{osLip}_x(F) \leq -c$ and $\text{Lip}_u(F) \leq \ell$

Task: compare **trajectory** $x(t)$ with **equilibrium trajectory** $x^*(\theta(t))$

Consider **auxiliary dynamics** with two trajectories:

$$\dot{x} = F(x, \theta(t)) + v(t) =: F_{\text{aux}}(x, \theta, v)$$

① $v(t) = 0 \implies \text{trajectory } x(t)$

② $v(t) = \dot{x}^*(\theta(t)) \implies \text{equilibrium trajectory } x^*(\theta(t))$

F_{aux} is contracting with $\text{osLip}_x(F_{\text{aux}}) \leq -c$ and $\text{Lip}_v(F_{\text{aux}}) = 1$. Hence, iISS:

$$\begin{aligned} D^+ \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} &\leq -c \cdot \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + 1 \cdot \|0 - \dot{x}^*(\theta(t))\|_{\mathcal{X}} \\ &\leq -c \cdot \|x(t) - x^*(\theta(t))\|_{\mathcal{X}} + \frac{\ell}{c} \cdot \|\dot{\theta}(t)\|_{\Theta} \quad \left(\text{since } \text{Lip}(x^*) = \frac{\ell}{c} \right) \end{aligned}$$

Summary:

- ① from convex optimization to contracting dynamics
- ② tracking-bounds for time-varying contracting systems
- ③ applications to standard convex optimization problems

Ongoing work and open problems:

- ① contracting predictor-corrector methods
- ② tracking bounds in time-varying norms
- ③ convex but not strongly convex problems

Thank you for reading so far!

For any questions, please do not hesitate to email me