Non-Euclidean Contraction and its Extensions with Application to Network Systems

Saber Jafarpour

Center for Control, Dynamical Systems & Computation
University of California at Santa Barbara

Georgia Institute of Technology

May 3, 2021

Acknowledgment

Francesco Bullo
UCSB

Alexander Davydov
UCSB

Pedro Cisneros-Velarde
UCSB

SJ and P. Cisneros-Velarde and F. Bullo. Weak and Semi-Contraction for Network Systems and Diffusively-Coupled Oscillators. IEEE Transactions on Automatic Control, accepted, Mar. 2021.

SJ and A. Davydov and F. Bullo. Non-Euclidean Contraction Theory for Monotone and Positive Systems. Submitted, Apr. 2021.

P. Cisneros-Velarde and SJ and F. Bullo. Distributed and time-varying primal-dual dynamics via contraction analysis. Submitted, May 2020.

Motivation: Large-scale Nonlinear Networks

Power grids

Brain neural network

Transportation network

Nonlinearity:

- Multiple equilibria
- Transient stability
- Cluster synchronization

Large-scale:

- Stochastic
- Distributed

- *"... As [power] systems become more heavily loaded, nonlinearities play an increasingly important role in power system behavior ... "* [I. Hiskens,1995]
- *"... in Oahu, Hawaii, at least 800,000 micro-inverters interconnect photovoltaic panels to the grid... "* [IEEE Spectrum, 2015]

Presentation outline

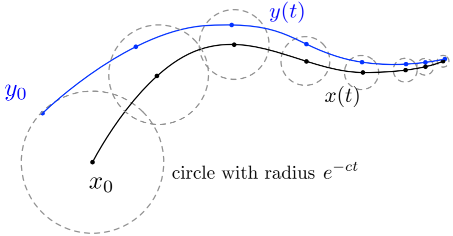
- non-Euclidean contraction theory
  - one-sided Lipschitz constant
  - characterization of contraction wrt non-Euclidean norms
  - contraction-based small gain theorem
- weakly-contracting systems
  - definition and examples
  - dichotomy in asymptotic behavior
  - example: distributed primal-dual
- semi-contracting systems
  - definition and examples
  - convergence to invariant subspaces
  - example: diffusively-coupled oscillators

Definition: Contracting systems

$\dot{x} = f(t, x)$ is contracting wrt to $\|\cdot\|$ with rate $c > 0$:

$$\|x(t) - y(t)\| \leq e^{-ct}\|x(0) - y(0)\|.$$

Contracting system: flow is a contracting map.



- B. P. Demidovich. Dissipativity of a nonlinear system of differential equations. *Uspekhi Matematicheskikh Nauk*, 16(3(99)):216, 1961
- **Application in control theory**: W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998
- **Differential framework**: F. Forni and R. Sepulchre. A differential Lyapunov framework for contraction analysis. *IEEE Trans. Autom. Control*, 59(3):614–628, 2014
- **Non-Euclidean contraction**: S. Coogan. A contractive approach to separable Lyapunov functions for monotone systems. *Automatica*, 106:349–357, 2019
- **Review**: 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: Dynamics, Controls and Applications*, pages 313–339. Springer, 2016
- **Review**: Z. Aminzare and E. D. Sontag. Contraction methods for nonlinear systems: A brief introduction and some open problems. In *Proc CDC*, pages 3835–3847, Dec. 2014

Highly ordered **asymptotic** and **transient** behaviors:

1. initial conditions are forgotten
2. no overshoot in distance between trajectories
3. time-invariant f: unique globally stable equilibrium
4. periodic f: unique globally stable periodic solution
5. robustness properties: input-to-state stability even in the presence of unmodeled dynamics.

Why *non-Euclidean* norms?

1. systematic and efficient stability analysis:
   - conservation law: $\mathbb{1}_n^\top \mathbf{x} = c$
   - geometric symmetry: $f(\mathbf{x} + \mathbb{1}_n) = f(\mathbf{x})$
2. ℓ_2-norm gives conservative estimates:
   - **contraction**: lack of symmetry in dynamics: $\frac{\partial f_i}{\partial x_j} \neq \frac{\partial f_j}{\partial x_i}$
   - **nonlinearity**: frequency instability in power grids, cluster synchronization in Brain neural networks
3. error analysis for large-scale networks:

$$x \in \mathcal{N}(0, \sigma I_n) \quad \Longrightarrow \quad \begin{aligned} &\mathbb{E}(\|x\|_2^2) = n\sigma^2, \\ &\mathbb{E}(\|x\|_\infty^2) \sim 2\ln(n)\sigma^2 \end{aligned}$$

Fundamental Question

How to check if a system is contracting?

- wrt ℓ_2-norm
  1. Differential condition: Linear matrix inequality
  2. Integral condition: one-sided Lipschitz constant
- wrt non-Euclidean norms
  1. Differential condition: matrix measure
- Differential conditions computationally intensive
- Differential conditions challenging for switching systems

One-sided Lipschitz constant: scalar vector fields

Let $\dot{x} = f(x)$, where $f : \mathbb{R} \to \mathbb{R}$:

Lipschitz constant $\ell \in \mathbb{R}$

$$|f(x) - f(y)| \leq \ell |x - y| \quad \Longrightarrow \quad -\ell \leq f'(x) \leq \ell$$

One-sided Lipschitz constant $b \in \mathbb{R}$

$$(x - y)(f(x) - f(y)) \leq b(x - y)^2 \quad \Longrightarrow \quad f'(x) \leq b$$
$$\langle f(x) - f(y), x - y \rangle \leq b|x - y|^2$$

- By the Gröwall-Bellman Lemma: $|x(t) - y(t)| \leq e^{bt}|y(0) - x(0)|$.
- Numerical analysis: E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer, 1993

Contraction for ℓ_2 norm

For $x \in \mathbb{R}^n$ and differentiable f:

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

For $P = P^\top \succ 0$, define $\|x\|_P^2 = x^\top P x$
equivalent properties:

1. **osL**: $(f(t,x) - f(t,y))^\top P(x - y) \leq -c\|x - y\|_P^2$, for all x, y, t
2. **LMI**: $PDf(t,x) + Df(t,x)^\top P \preceq -2cP$ for all x, t,
3. **dIS**: $D^+\|x(t) - y(t)\|_P \leq -c\|x(t) - y(t)\|_P$, for all soltns $x(\cdot), y(\cdot)$
4. **IS**: $\|x(t) - y(t)\|_P \leq e^{-c(t-t_0)}\|x(t_0) - y(t_0)\|_P$, for all soltns $x(\cdot), y(\cdot)$

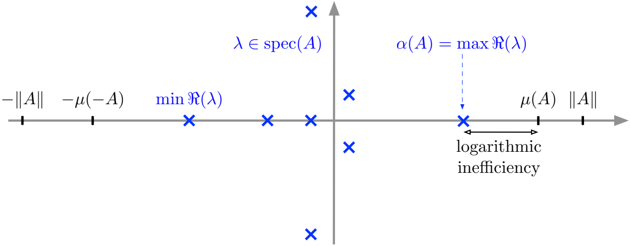
If f not differentiable, then **osL** $\iff$ **dIS** $\iff$ **IS**

Question: How to extend **osL** and **LMI** to non-Euclidean norms?

The **matrix measure** of $A \in \mathbb{R}^{n \times n}$ wrt to $\| \cdot \|$:

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

- Directional derivative of norm $\| \cdot \|$ in direction of A,
- Logarithmic norm: T. Ström. On logarithmic norms. *SIAM Journal on Numerical Analysis*, 12(5):741–753, 1975
- spectral property: $-\|A\| = \Re(\lambda) \leq \mu_{\| \cdot \|}(A) \leq \|A\|$, for every $\lambda \in \mathrm{spec}(A)$



A **weak semi-inner product** (WSIP) is $[\![\cdot,\cdot]\!] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying

1. $[\![x_1 + x_2, y]\!] \leq [\![x_1, y]\!] + [\![x_2, y]\!]$ and $x \mapsto [\![x, y]\!]$ is continuous,
2. $[\![\alpha x, y]\!] = [\![x, \alpha y]\!] = \alpha [\![x, y]\!]$ for $\alpha \geq 0$ and $[\![-x, -y]\!] = [\![x, y]\!]$,
3. $[\![x, x]\!] > 0$, for all $x \neq 0$,
4. $|\,[\![x, y]\!]\,| \leq [\![x, x]\!]^{1/2} [\![y, y]\!]^{1/2}$,
5. (compatibility) $[\![x, x]\!] = \|x\|^2$ for all x

- Extension of semi-inner product: G. Lumer. Semi-inner-product spaces. *Transactions of the American Mathematical Society*, 100:29–43, 1961

- Properties:

Matrix measure: $$\mu(A) = \sup_{\|x\|=1} [\![Ax, x]\!]\,,$$

Norm derivative formula: $$\|x(t)\| D^+ \|x(t)\| = [\![\dot{x}(t), x(t)]\!]\,.$$

Norm	WSIP	Matrix measure
$\|x\|_P = \sqrt{x^\top P x}$	$[\![x,y]\!]_{2,P^{1/2}} = x^\top P y$	$\mu_{2,P^{1/2}}(A) = \min\{b \in \mathbb{R} \mid A^\top P + PA \preceq 2bP\}$ $= \max_{\|x\|_P=1} x^\top P A x$
$\|x\|_p = \left(\sum_i \|x_i\|^p\right)^{1/p}$ $p \in]1,\infty[$	$[\![x,y]\!]_p = \|y\|_p^{2-p}(y \circ \|y\|^{p-2})^\top x$	$\mu_p(A) = \max_{\|x\|_p=1} (x \circ \|x\|^{p-2})^\top A x$
$\|x\|_1 = \sum_i \|x_i\|$	$[\![x,y]\!]_1 = \|y\|_1 \mathrm{sign}(y)^\top x$	$\mu_1(A) = \max_{j \in \{1,\dots,n\}} \left(a_{jj} + \sum_{i \neq j} \|a_{ij}\|\right)$ $= \sup_{\|x\|_1=1} \mathrm{sign}(x)^\top A x$
$\|x\|_\infty = \max_i \|x_i\|$	$[\![x,y]\!]_\infty = \max_{i \in I_\infty(y)} y_i x_i$	$\mu_\infty(A) = \max_{i \in \{1,\dots,n\}} \left(a_{ii} + \sum_{j \neq i} \|a_{ij}\|\right)$ $= \max_{\|x\|_\infty=1} \max_{i \in I_\infty(x)} x_i (Ax)_i$

Table of norms, WSIPs, and matrix measures for weighted ℓ_2, ℓ_p for $p \in (1,\infty)$, ℓ_1, and ℓ_∞ norms. Note: $I_\infty(x) = \{i \in \{1,\dots,n\} \mid |x_i| = \|x\|_\infty\}$.

Contraction for arbitrary norm

For $x \in \mathbb{R}^n$ and differentiable f:

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

For norm $\|\cdot\|$ with matrix measure $\mu(\cdot)$ and compatible WSIP $[\![\cdot,\cdot]\!]$, equivalent properties:

1. **osL**: $[\![f(t,x) - f(t,y), x - y]\!] \leq -c\|x-y\|^2$ for all $x, y,\ t \geq 0$,
2. **MM**: $\mu(Df(t,x)) \leq -c$, for all $x, t \geq 0$,
3. **dIS**: $D^+\|x(t) - y(t)\| \leq -c\|x(t) - y(t)\|$, for soltns $x(\cdot), y(\cdot)$,
4. **IS**: $\|x(t) - y(t)\| \leq e^{-c(t-t_0)}\|x(t_0) - y(t_0)\|$, for all soltns $x(\cdot), y(\cdot)$

Measure bound	Demidovich condition	One-sided Lipschitz condition
$\mu_{2,P}(Df(x)) \leq -c$	$PDf(x) + Df(x)^\top P \preceq -2cP$	$(x-y)^\top P\big(f(x)-f(y)\big) \leq -c\,\|x-y\|_P^2$
$\mu_p(Df(x)) \leq -c$	$(v \circ \|v\|^{p-2})^\top Df(x)v \leq -c\|v\|_p^p$	$((x-y)\circ\|x-y\|^{p-2})^\top(f(x)-f(y)) \leq -c\|x-y\|_p^p$
$\mu_1(Df(x)) \leq -c$	$\mathrm{sign}(v)^\top Df(x)v \leq -c\,\|v\|_1$	$\mathrm{sign}(x-y)^\top(f(x)-f(y)) \leq -c\,\|x-y\|_1$
$\mu_\infty(Df(x)) \leq -c$	$\max_{i\in I_\infty(v)} v_i\,(Df(x)v)_i \leq -c\,\|v\|_\infty^2$	$\max_{i\in I_\infty(x-y)}(x_i-y_i)(f_i(x)-f_i(y)) \leq -c\,\|x-y\|_\infty^2$

Table of equivalences between measure bounded Jacobians, differential Demidovich and one-sided Lipschitz conditions. Note: $I_\infty(v) = \{i \in \{1,\dots,n\} \mid |v_i| = \|v\|_\infty\}$.

- n interconnected systems $\dot{x}_i = f_i(x_i, x_{-i})$ where $x_i \in \mathbb{R}^{N_i}$
- Lyapunov functions $V_i : \mathbb{R}^{N_i} \to \mathbb{R}_{\geq 0}$:

$$\mathcal{L}_{f_i} V_i(x_i) = \frac{\partial V_i}{\partial x_i} f_i(x_i, x_{-i}) \leq g_i(V_i(x_i), V_{-i}(x_{-i}))$$

- **Classical small-gain**:

$$g_i(\mathbf{v}) = -\alpha_i(v_i) + \sum_{j \neq i} \gamma_{ij}(v_j)$$

for class $\mathcal{K}_\infty$ functions α_i, γ_{ij}.

Comparison system

Study properties of

$$\dot{\mathbf{v}} = g(\mathbf{v}), \qquad g = (g_1, \dots, g_n)^\top$$

for $\mathbf{v} \in \mathbb{R}^n_{\geq 0}$.

- $V(\mathbf{x}) = (V_1(x_1), \dots, V_n(x_n))^\top$
- $[\![\cdot,\cdot]\!]_{p,R}$ is associated with $\|R(\cdot)\|_p$, for $p \in [1,\infty]$, $R \in \mathbb{R}^{n\times n}$ non-negative.

Theorem: Contraction-based small-gain

If there exists $c > 0$ such that for all $\mathbf{v} \geq \mathbf{w} \geq \mathbb{0}_n$,

$$[\![g(\mathbf{v}) - g(\mathbf{w}), \mathbf{v} - \mathbf{w}]\!]_{p,R} \leq -c\|\mathbf{v} - \mathbf{w}\|^2_{p,R},$$

Then

$$\|V(x(t))\|_{p,R} \leq e^{-ct}\|V(x(0))\|_{p,R}$$

For $R = \mathrm{diag}(\eta)$ where $\eta \in \mathbb{R}^n_{>0}$

- if $p \in [1,\infty)$, sum-separable Lyapunov function: $\sum_{i=1}^n \eta_i^p V_i^p(x_i)$
- if $p = \infty$, max-separable Lyapunov function: $\max_i\{\eta_i V_i(x_i)\}$

Key Lemma

$[\![x, y]\!]_{p,R} \leq [\![z, y]\!]_{p,R}$, for every $x \leq z$ and $y \geq \mathbb{0}_n$.

$$\begin{aligned} \|V(x(t))\|_{p,R} D^+ \|V(x(t))\|_{p,R} &= \left[\!\!\left[\dot{V}(x(t)), V(x(t))\right]\!\!\right]_{p,R} && \text{Norm derivative formula} \\ &\leq [\![g(V(x(t))), V(x(t))]\!]_{p,R} && \text{Key Lemma} \\ &\leq -c\|V(x(t))\|_{p,R}^2 && \text{osL} \end{aligned}$$

- Unlike classical small-gain theorems we do not need g to be monotone.

Presentation outline

- non-Euclidean contraction theory
  - one-sided Lipschitz constant
  - characterization of contraction wrt non-Euclidean norms
  - contraction-based small gain theorem
- weakly-contracting systems
  - definition and examples
  - dichotomy in asymptotic behavior
  - example: distributed primal-dual
- semi-contracting systems
  - definition and examples
  - convergence to invariant subspaces
  - example: diffusively-coupled oscillators

Contraction theory for networks

Challenge: many real-world networks are not contracting.


Network flow system $\dot{x} = f(x)$ preserving commodity x:

$$\begin{aligned} \text{constant} &= \mathbb{1}_n^\top x(t) \\ &\implies 0 = \mathbb{1}_n^\top \dot{x}(t) = \mathbb{1}_n^\top f(x(t)) \\ &\implies \mathbb{0}_n = \mathbb{1}_n^\top Df(x(t)) \end{aligned}$$

If additionally f has Metzler Jacobian, then $\mu_1(Df(x)) = 0$.

Definition: Weakly-contracting systems

$\dot{x} = f(t,x)$ with f continuously differentiable in x is weakly-contracting wrt $\|\cdot\|$:

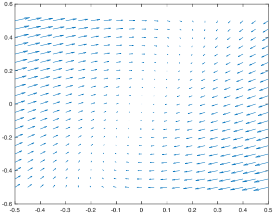
$$\mu_{\|\cdot\|}(Df(t,x)) \leq 0$$

1. Lotka-Volterra population dynamics (Lotka, 1920; Volterra, 1928) (ℓ_1-norm)
2. Kuramoto oscillators (Kuramoto, 1975) and coupled swing equations (Bergen and Hill, 1981) (ℓ_1-norm and ℓ_∞-norm)
3. Daganzo's cell transmission model for traffic networks (Daganzo, 1994), (ℓ_1-norm)
4. compartmental systems in biology, medicine, and ecology (Sandberg, 1978; Maeda et al., 1978). (ℓ_1-norm)
5. saddle-point dynamics for optimization of weakly-convex functions (Arrow et al., 1958). (ℓ_2-norm)

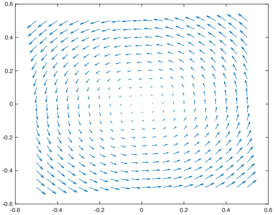
Theorem: Dichotomy for weakly-contracting systems

For a weakly-contracting system $\dot{x} = f(x)$, either

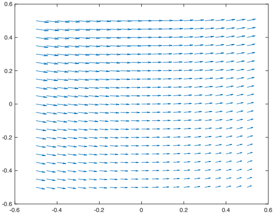
1. f has no equilibrium and every trajectory is unbounded, or
2. f has at least one equilibrium x^* and every trajectory is bounded.



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



$$\dot{x}_1 = -x_2$$
$$\dot{x}_2 = x_1$$



$$\dot{x}_1 = -x_1 + \sin(x_2) + 3$$
$$\dot{x}_2 = x_1$$

Theorem

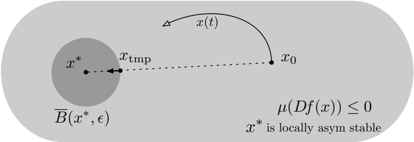
If $\dot{x} = f(x)$ is weakly-contracting and f has at least one equilibrium x^* then:

(i) each equilibrium x^{**} is stable with weak Lyapunov function $x \mapsto \|x - x^{**}\|$,

(ii) if the norm $\|\cdot\|$ is a (p, R)-norm, $p \in \{1, \infty\}$ and f is piecewise real analytic, then every trajectory converges to the set of equilibria,

(iii) x^* is locally asy stable $\implies$ x^* is globally asy stable.

Idea of the proof



Optimization problem

$$\min_{x\in\mathbb{R}^k} f(x) = \min_{x\in\mathbb{R}^k} \sum_{i=1}^{n} f_i(x)$$

Distributed implementation

- n agents communicate over a undirected weighted graph G,
- agent i have access to function f_i and can exchange x_i with its neighbors.

$$\min_{x\in\mathbb{R}^k} \quad \sum_{i=1}^{n} f_i(x_i)$$

$$x_1 = x_2 = \ldots = x_n$$

In matrix form by assuming $x = (x_1^\top, \ldots, x_n^\top)^\top \in \mathbb{R}^{nk}$:

$$\min_{x\in\mathbb{R}^k} \quad \sum_{i=1}^{n} f_i(x_i)$$

$$(L \otimes I_k)x = \mathbb{0}_{nN}$$

If each f_i is continuously differentiable in x_i:

Lagrangian

$\mathcal{L}(x,\nu) = \sum_{i=1}^n f_i(x_i) + \nu^\top (L \otimes I_k)x$

Distributed primal-dual algorithm (component form):

$$\dot{x}_i = -\frac{\partial \mathcal{L}}{\partial x_i} = -\nabla f_i(x_i) - \sum_{j=1}^n a_{ij}(\nu_i - \nu_j),$$

$$\dot{\nu}_i = \frac{\partial \mathcal{L}}{\partial \nu_i} = \sum_{j=1}^n a_{ij}(x_i - x_j)$$

Distributed primal-dual algorithm (vector form):

$$\dot{x} = -\nabla f(x) - (L \otimes I_k)\nu,$$
$$\dot{\nu} = (L \otimes I_k)x$$

Assume

1. f has a minimum $x^* \in \mathbb{R}^k$,
2. for each $i \in \{1, \dots, n\}$, f_i is twice differentiable, $\nabla^2 f_i(x) \succeq 0$ for all x, and $\nabla^2 f_i(x^*) \succ 0$, and
3. the undirected weighted graph G is connected with Laplacian L.

Theorem: Distributed primal-dual dynamics

The distributed primal-dual algorithm

1. is weakly-contracting wrt ℓ_2-norm,
2. $(x(t), \nu(t)) \to (\mathbb{1}_n \otimes x^*, \mathbb{1}_n \otimes \nu^*)$, with $\nu^* = \sum_{i=1}^n \nu_i(0)$,
3. exponential convergence rate is $-\alpha_{\text{ess}}\left(\begin{bmatrix} -\nabla^2 f(x^*) & -L \otimes I_k \\ L \otimes I_k & \mathbb{0} \end{bmatrix}\right)$ where

$$\alpha_{\mathrm{ess}}(A) := \max\{\Re(\lambda) \mid \lambda \in \mathrm{spec}(A) \setminus \{0\}\}.$$

- Idea: flows converges to each other only in **certain directions**.

Definition: Semi-norms

$|||\cdot|||$ is a *semi-norm* if

1. $|||cv||| = |c||||v|||$, for every $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$;
2. $|||v + w||| \leq |||v||| + |||w|||$, for every $v, w \in \mathbb{R}^n$.

- Define the subspace $\mathrm{Ker}\, |||\cdot||| = \{v \in \mathbb{R}^n \mid |||v||| = 0\}$.
- Example: for $k < n$, $R \in \mathbb{R}^{k \times n}$, and norm $\|\cdot\|$, we get $|||x|||_R = \|Rx\|$.
- Example: for a network G with edge set $\mathcal{E}$ and incidence matrix B:

$$|||x|||_{\mathcal{E}} := \max_{(i,j) \in \mathcal{E}} |x_i - x_j| = \|B^\top x\|_\infty$$

For strongly connected graphs $\mathrm{Ker}\, |||\cdot|||_{\mathcal{E}} = \mathrm{span}\{\mathbb{1}_n\}$

Definition: Matrix semi-measures

The **matrix semi-measure** of $A \in \mathbb{R}^{n\times n}$ wrt $|||\cdot|||$:

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

- Directional derivative of $|||\cdot|||$ in direction of A.
- if Ker $|||\cdot|||$ is invariant under A then $\Re(\lambda) \leq \mu_{|||\cdot|||}(A)$, for every $\lambda \in \mathrm{spec}_{\mathrm{Ker}\,|||\cdot|||^\perp}(A^\top)$.

Definition: Semi-contracting systems

$\dot{x} = f(t,x)$ with f continuously differentiable in x is semi-contracting wrt the semi-norm $|||\cdot|||$ with rate $c > 0$:

$$\mu_{|||\cdot|||}(Df(t,x)) \leq -c$$

1. Kuramoto oscillators (Kuramoto, 1975) and coupled swing equations (Bergen and Hill, 1981), (ℓ_1-norm)
2. Chua's diffusively-coupled circuits (Wu and Chua, 1995), (ℓ_2-norm)
3. morphogenesis in developmental biology (Turing, 1952), (ℓ_1-norm)
4. Goodwin model for oscillating auto-regulated gene (Goodwin, 1965). (ℓ_1-norm)

Theorem: Semi-contracting systems

Consider $\dot{x} = f(t, x)$ with f continuously differentiable in x and assume

- f is semi-contracting wrt the semi-norm $|||\cdot|||$ with rate $c > 0$, and
- **(Affine invariance)**: $f(t, x^* + \mathrm{Ker}\,|||\cdot|||) \subseteq \mathrm{Ker}\,|||\cdot|||$ for every t

Then,

1. for every trajectory $x(t)$,

$$|||x(t) - x^*||| \leq e^{-ct}|||x(0) - x^*|||, \qquad \text{for every } t \geq 0.$$

2. every trajectory converges to $x^* + \mathrm{Ker}\,|||\cdot|||$.

- partial contraction and horizontal contraction

Example: Diffusively-coupled oscillators

- n agents connected by a weighted undirected graph G,
- identical internal dynamics $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^k \to \mathbb{R}^k$

$$\dot{x}_i = f(t, x_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j), \quad i \in \{1, \dots, n\}$$

- **synchronization**:

$$\lim_{t\to\infty} \|x_i - x_j\| = 0 \qquad \text{for every } i, j$$

- synchronization of diffusively-coupled oscillators:

  1. contractivity of the internal dynamics
  2. strength of the diffusive coupling

Example: Diffusively-coupled oscillators

Introduce a **local-global mixed norm:** $(2,p)$-tensor norm on $\mathbb{R}^{nk} = \mathbb{R}^n \otimes \mathbb{R}^k$

$$\|u\|_{(2,p)} = \inf\left\{ \Big(\sum_{i=1}^{r} \|v^i\|_2^2 \|w^i\|_p^2 \Big)^{\frac{1}{2}} \;\Big|\; u = \sum_{i=1}^{r} v^i \otimes w^i \right\}.$$

- closely related to, but different from, the projective tensor product norm

  R. A. Ryan. *Introduction to Tensor Products of Banach Spaces.*
  Springer, 2002

- different from the mixed global norm

  G. Russo, M. Di Bernardo, and E. D. Sontag. A contraction approach to the hierarchical analysis and design of networked systems.
  IEEE Trans. Autom. Control, 58(5):1328–1331, 2013

- **Global norm**: ℓ_2-norm for the interactions between agents
- **Local norm**: ℓ_p-norm for internal dynamics of each agent

Example: Diffusively-coupled oscillators

The orthogonal projection $\mathcal{P} \in \mathbb{R}^{n\times n}$

$$\mathcal{P} = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top = \begin{bmatrix} \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} \end{bmatrix}$$

- $(\mathcal{P} \otimes I_k)x$ measures **dissimilarity** of the states x_i

$$x = \mathbb{1}_n \otimes x^* \implies$$

$$(\mathcal{P} \otimes I_k)x = (\mathcal{P} \otimes I_k)(\mathbb{1}_n \otimes x^*) = \mathcal{P}\mathbb{1}_n \otimes x^* = \mathbb{0}_{(n-1)\times k}.$$

Example: Diffusively-coupled oscillators

- G is an undirected weighted graph with Laplacian L,
- $p \in [1, \infty]$, $Q \in \mathbb{R}^{k \times k}$

$$\dot{x}_i = f(t, x_i) - \sum_{j=1}^{n} a_{ij}(x_i - x_j), \quad i \in \{1, \dots, n\}$$

Theorem: diffusively-coupled oscillators are semi-contracting

Suppose that

$$\mu_{p,Q}(Df(t,x)) \leq \lambda_2(L) - c, \qquad \text{for every } t, x$$

then

1. the dynamics is semi-contracting wrt $\| \cdot \|_{(2,p),(\mathcal{P} \otimes Q)}$;
2. for every trajectory $x(t)$,

$$\|x(t) - \mathbb{1}_n \otimes x_{\mathrm{ave}}(t)\|_{(2,p),(\mathcal{P} \otimes Q)} \leq e^{-ct} \|x(0) - \mathbb{1}_n \otimes x_{\mathrm{ave}}(0)\|_{(2,p),(\mathcal{P} \otimes Q)}.$$

3. the system achieves synchronization: $\lim_{t \to \infty} x(t) = \mathbb{1}_n \otimes x_{\mathrm{ave}}(t)$

where $x_{\mathrm{ave}}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t)$

Example: Diffusively-coupled oscillators

$$\mu_{p,Q}(Df(t,x)) \leq \lambda_2(L) - c, \qquad \text{for every } t, x$$

- trade off between **internal dynamics** and **coupling strength**
- f time-invariant: every trajectory converges to the unique equilibrium point.
- f periodic: every trajectory converges to the unique periodic orbit.
- Unstable dynamics f, sufficiently strong coupling $\implies$ $\lambda_2(L)$ large $\implies$ the network synchronizes.

Summary

- notion of one-sided Lipschitz constant and weak semi-inner product
- characterization of contraction wrt non-Euclidean norms
- contraction-based small-gain theorem
- two extensions of classical contraction:
  - weak contraction
  - semi-contraction
- properties of weakly-contracting and semi-contracting systems

Future research

- contraction-based compositional analysis of interconnected systems
  - scalable stability certificates using non-Euclidean contraction.
- computing equilibra of contracting and weakly-contracting systems
  - explicit and implicit integration algorithms
  - accelerated convergence.
- optimization algorithms using contraction theory
  - extension to gradient descent algorithms and time-varying algorithms.
  - connection with discrete-time algorithms for optimization.
- implicit deep learning using contraction theory
  - use contraction condition for well-posedness in the optimization problem.

Consider the following system on $\mathbb{R}^2$:

$$\dot{x}_1 = -x_1 + 3x_2^4 x_1 - x_1^3$$
$$\dot{x}_2 = -x_2 - 3x_1^4 x_2 - x_2^5$$

and pick $V_i(x_i) = x_i^2$ for $i \in \{1, 2\}$. Then

$$\dot{V}_1 = -2V_1 + 6V_1 V_2^2 - 2V_1^2 \leq -2V_1 + 6V_1 V_2^2 := g_1(V_1, V_2)$$
$$\dot{V}_2 = -2V_2 - 6V_1^2 V_2 - 2V_2^3 \leq -2V_2 - 6V_1^2 V_2 := g_2(V_1, V_2)$$

but $g = (g_1, g_2)^\top$ is not monotone. However, for every $\mathbf{v} \geq \mathbf{w} \geq \mathbb{0}_n$,

$$\llbracket g(\mathbf{v}) - g(\mathbf{w}), \mathbf{v} - \mathbf{w} \rrbracket_2$$
$$= \begin{bmatrix} v_1 - w_1 & v_2 - w_2 \end{bmatrix} \begin{bmatrix} -2v_1 + 2w_1 + 6v_1 v_2^2 - 6w_1 w_2^2 \\ -2v_2 + 2w_2 - 6v_1^2 v_2 + 6w_1^2 w_2 \end{bmatrix}$$
$$\leq -2|v_1 - w_1|^2 - 2|v_2 - w_2|^2 = -2\|\mathbf{v} - \mathbf{w}\|_2^2$$