# Non-Euclidean Contractivity of Recurrent Neural Networks


Alexander Davydov

Center for Control,
Dynamical Systems & Computation
University of California at Santa Barbara
**davydovalexander.github.io**


2022 American Control Conference
Atlanta, Georgia, June 8, 2022

# Acknowledgments

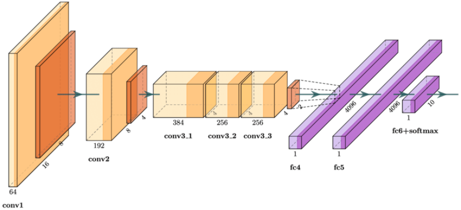
Francesco Bullo
UC Santa Barbara

Saber Jafarpour
Postdoc
Georgia Tech

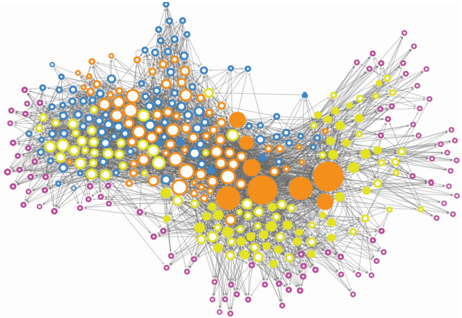
Anton Proskurnikov
Politecnico Torino & Russian Academy of Sciences


- A. Davydov, S. Jafarpour, and F. Bullo. Non-Euclidean contraction theory for robust nonlinear stability. *IEEE Transactions on Automatic Control*, July 2021. URL **https://arxiv.org/abs/2103.12263**. To appear
- 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. URL **http://arxiv.org/abs/2106.03194**
- 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: 10.1109/CDC45484.2021.9682883

# Artificial and Biological Neural Networks




**artificial neural network AlexNet '12**



**C. elegans connectome '17**

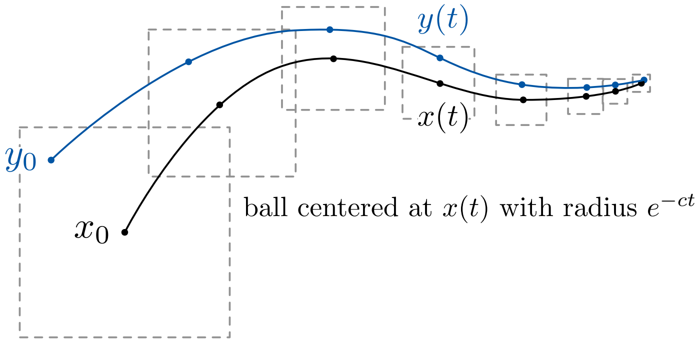
**Aim:** understand the dynamics of neural networks, so that

- **reproducible behavior, i.e., equilibrium response as function of stimuli**
- robust behavior in face of uncertain stimuli and dynamics
- learning models, efficient computational tools, periodic behaviors ...

A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 2012
G. Yan, P. E. Vértes, E. K. Towlson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási. Network control principles predict neuron function in the Caenorhabditis elegans connectome. *Nature*, 550(7677):519–523, 2017. doi: 10.1038/nature24056

# Contraction theory: definition

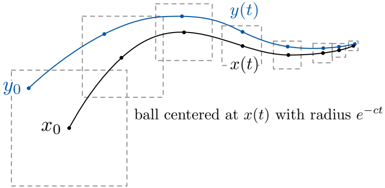
Given $\dot{x} = \mathsf{F}(t, x)$, vector field F is contractive if its flow is a contraction map





W. Lohmiller and J.-J. E. Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998. doi: 10.1016/S0005-1098(98)00019-3

# Properties of contracting dynamical systems

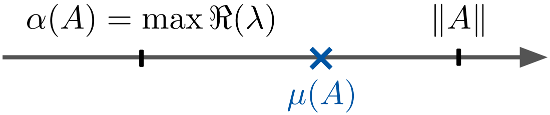



Highly ordered **transient** and **asymptotic** behavior:

1. time-invariant F: unique globally exponential stable equilibrium
   two natural Lyapunov functions
2. contractivity rate is natural measure/indicator of robust stability
3. modularity and interconnection properties,
4. ...

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

$$\mu(A) := \min\{b \in \mathbb{R} \mid \|e^{At}\| \leq e^{bt},\ \forall t \geq 0\}$$




**Weighted $\ell_1/\ell_\infty$ norms:**

$$\|x\|_{1,\eta} = \sum_{i=1}^{n} \eta_i |x_i|, \qquad \|x\|_{\infty,\eta} = \max_{i\in\{1,\dots,n\}} \frac{1}{\eta_i}|x_i|$$

# Contraction equivalence on normed vector spaces

For $x \in \mathbb{R}^n$ and locally Lipschitz

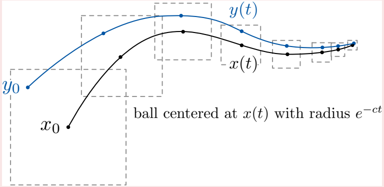
$$\dot{x} = \mathsf{F}(x),$$

norm $\|\cdot\|$ with log norm $\mu(\cdot)$

$$\mathsf{osLip}(\mathsf{F}) := \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \mu(D\mathsf{F}(x))$$

**Main equivalence:** for $c > 0$

$$\mathsf{osLip}(\mathsf{F}) \leq -c \qquad \iff$$



# Optimizing non-Euclidean log norms

The **Metzler majorant** of $A \in \mathbb{R}^{n\times n}$ is $\lceil A \rceil_{\mathrm{Mzr}}$ and is given by:

$$(\lceil A \rceil_{\mathrm{Mzr}})_{ij} := \begin{cases} a_{ii}, & \text{if } i = j, \\ |a_{ij}|, & \text{if } i \neq j. \end{cases}$$

**Convexity in $A$, quasiconvexity in $\eta$:**

$$\mu_{1,\eta}(A) = \min\{b \in \mathbb{R} \mid \lceil A \rceil_{\mathrm{Mzr}}^{\top} \eta \leq b\eta\},$$
$$\mu_{\infty,\eta}(A) = \min\{b \in \mathbb{R} \mid \lceil A \rceil_{\mathrm{Mzr}} \eta \leq b\eta\}$$

**Optimal weights:**

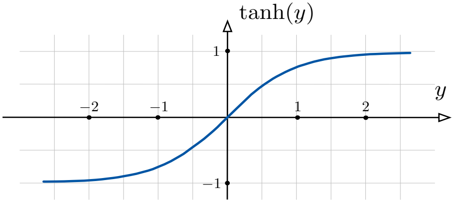
$$\inf_{\eta \in \mathbb{R}^n_{>0}} \mu_{1,\eta}(A) = \inf_{\eta \in \mathbb{R}^n_{>0}} \mu_{\infty,\eta}(A) = \alpha(\lceil A \rceil_{\mathrm{Mzr}})$$

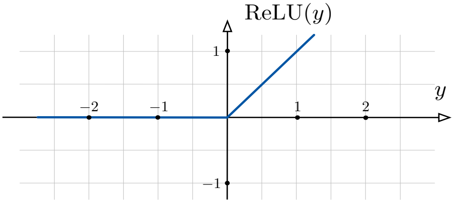
# Applications to recurrent neural networks

Continuous-time recurrent neural networks:

$$\dot{x} = -x + A\Phi(x) + u \qquad \text{(\textit{Hopfield NN})}$$
$$\dot{x} = -x + \Phi(Ax + u) \qquad \text{(\textit{Firing rate} $\sim$ \textit{Implicit NNs})}$$
$$\dot{x} = A\Phi(x) \qquad \text{(\textit{Persidskii-type})}$$
$$\dot{x} = Ax - B\Phi(Cx) \qquad \text{(\textit{Lur'e-type})}$$

activation functions are Lipschitz and monotone



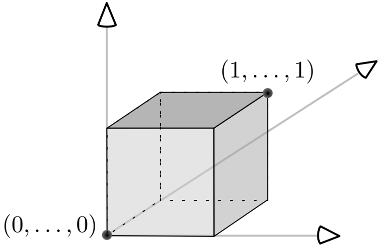
$$\dot{x} = -x + A\Phi(x) + u =: f_{\mathsf{H}}(x)$$

**Tight transcription.**

$$\mathsf{osLip}_1(f_{\mathsf{H}}) = -1 + \max_{d \in [d_{\mathsf{min}}, d_{\mathsf{max}}]^n} \mu_1(A \operatorname{diag}(d))$$

**Max log norms over hypercubes.**

$$\max_{d \in [d_{\mathsf{min}}, d_{\mathsf{max}}]^n} \mu_1(A \operatorname{diag}(d)) = \max \left\{ \mu_1(d_{\mathsf{min}} A), \mu_1(d_{\mathsf{max}} A) \right\}$$



# Non-Euclidean contractivity of Hopfield NN

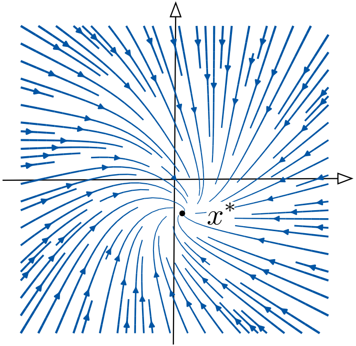
$$\dot{x} = -x + A\Phi(x) + u =: f_{\mathsf{H}}(x)$$

1. *closed-form formula*

$$\mathsf{osLip}_{1,\eta}(f_{\mathsf{H}}) = -1 + \max\{\mu_{1,\eta}(d_{\mathsf{min}}A), \mu_{1,\eta}(d_{\mathsf{max}}A)\}$$

2. *optimizing contraction rate*:

$$\begin{aligned}
\inf_{b\in\mathbb{R},\eta\in\mathbb{R}^n_{>0}} \quad & b \\
\text{s.t.} \quad & (-I_n + d_{\mathsf{min}}\lceil A\rceil_{\mathrm{Mzr}})^\top \eta \leq b\eta \\
& (-I_n + d_{\mathsf{max}}\lceil A\rceil_{\mathrm{Mzr}})^\top \eta \leq b\eta
\end{aligned}$$



# Comparison to Euclidean tests

$$\dot{x} = -x + A\Phi(x) + u := f_{\mathsf{H}}(x)$$

Suppose $d_{\mathsf{min}} = 0, d_{\mathsf{max}} = 1, u$ constant

$$\alpha(-I_n + \lceil A \rceil_{\mathrm{Mzr}}) < 0 \quad \implies \quad f_{\mathsf{H}} \text{ is } \ell_1 \text{ contracting with rate } 1 - \alpha(\lceil A \rceil_{\mathrm{Mzr}})_+.$$

$$\|A\|_2 < 1 \quad \implies \quad f_{\mathsf{H}} \text{ is } \ell_2 \text{ contracting with rate } 1 - \|A\|_2.$$

$$\mathrm{diag}(\eta)(-I_n + A) + (-I_n + A)^\top \mathrm{diag}(\eta) \prec 0 \quad \implies \quad \exists \text{ G.A.S. equilibrium.}$$

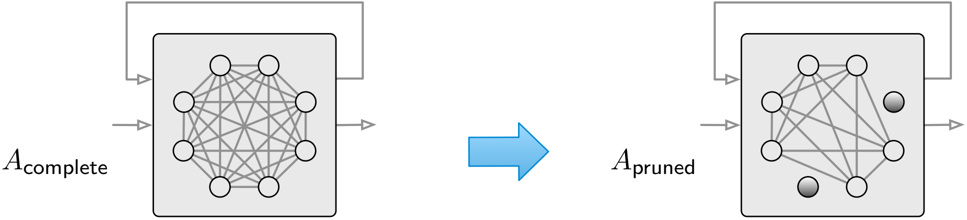
(E. Kaszkurewicz and A. Bhaya 1994, M. Forti and A. Tesi 1995, T. Chen and S. I. Amari 2001, ...)

# Additional advantages of non-Euclidean contraction

## Advantages of non-Euclidean approach

1. *computational advantages*: $\ell_1/\ell_\infty$ log-norm constraints lead to LPs, whereas $\ell_2$ constraints lead to LMIs

2. *guaranteed robustness to structural perturbations*: $\ell_1/\ell_\infty$ contractivity ensures:
    1. with respect to a class of activation functions
    2. remove any node and all its incident connections
    3. remove any set of edges



# Conclusions

## Summary:

1. contraction theory for artificial and biological neural networks
2. advantages of non-Euclidean norms and connection to Metzler matrices
3. application to continuous-time recurrent neural network
    - exact one-sided Lipschitz constant
    - tests correspond to LPs or checking Hurwitzness of a Metzler matrix

## Extensions and open problems:

1. bio-inspired Hebbian learning
2. From Metzler Hurwitz to LDS
3. Studying other neural network architectures

V. Centorrino, F. Bullo, and G. Russo. Contraction analysis of Hopfield neural networks with Hebbian learning. In *IEEE Conf. on Decision and Control*, Dec. 2022. URL **http://arxiv.org/abs/2204.05382**. Submitted