Exploring recurrent neural network dynamics: A spectral approach based on Koopman operator theory

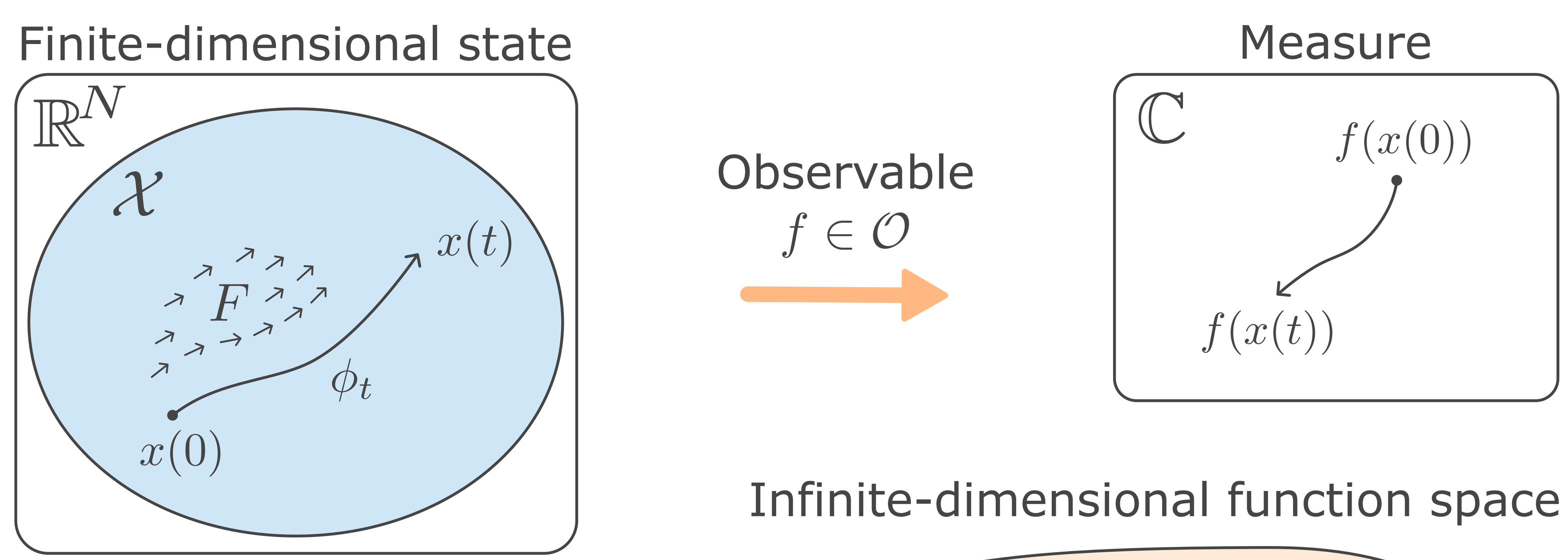
Benjamin Claveau[1, 2], Vincent Thibeault[1, 2], Antoine Allard[1, 2], Patrick Desrosiers[1, 2, 3]

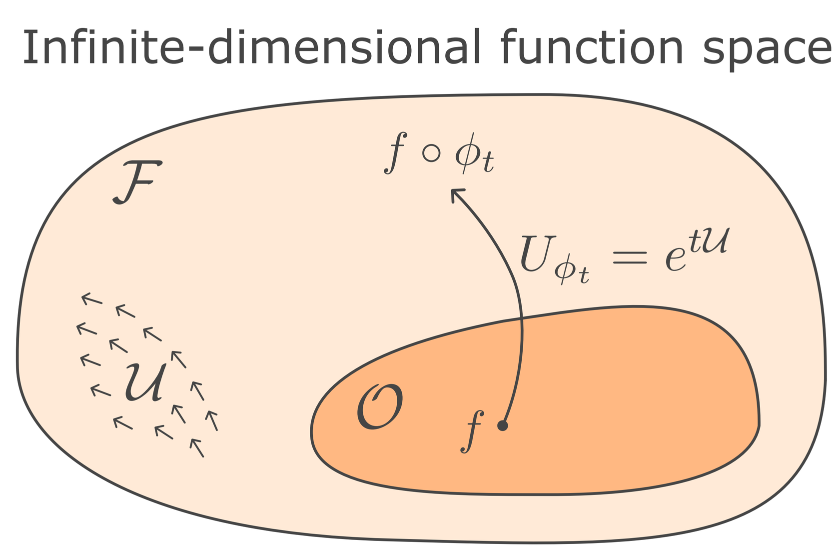
1. Département de physique, génie physique et d'optique, Université Laval, Québec, Canada
2. Centre Interdisciplinaire en Modélisation Mathématique de l'Université Laval, Québec, Canada
3. Centre de recherche CERVO, Québec, Canada

Dynamica · Université Laval · benjamin.claveau.1@ulaval.ca

WHAT IS KOOPMAN OPERATOR THEORY ?

- Nonlinear dynamical systems can be represented as linear systems of nonlinear observables f.
- The linear evolution operator in the functional space is the Koopman operator
$$U_{\phi_t} = e^{t\mathcal{U}},$$
where $\mathcal{U}$ is the Koopman operator's generator (KOG).
- The eigenfunctions ψ_λ with eigenvalue λ of the KOG are particular observables with uncoupled behaviour:
$$\mathcal{U}\,\psi_\lambda = \lambda\,\psi_\lambda \qquad \Rightarrow \qquad U_{\phi_t}\,\psi_\lambda = e^{\lambda t}\psi_\lambda$$
The eigenfunctions encode global and local dynamical features of dynamical systems, including equilibrium points (EP).



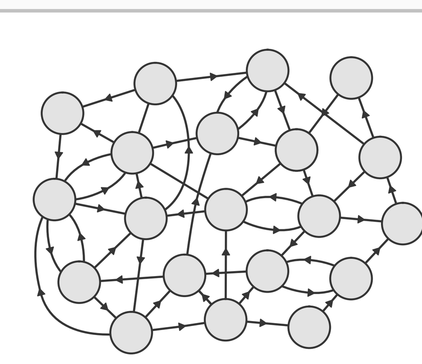


[1] Brunton et al., SIAM Rev. (2022)
[2] Mauroy and Mezić, IEEE Trans. Automat. Contr. (2016)
[3] Salova et al., Chaos (2019)

RESEARCH PROBLEM

Find the KOG spectral properties of recurrent neural networks described by
$$\frac{\mathrm{d}x_i}{\mathrm{d}t}(t) = -x_i(t) + \sum_{j=1}^{N} w_{ij}\,\sigma(x_j(t)) + \theta_i$$
where x_i : activity of neuron i
w_{ij} : weight of the connection from j to i
σ : activation function
θ_i : external input of neuron i
N : number of neurons

LONG-TERM GOALS
- Identify effective dimensionality
- Assess network resilience
- Characterize learning processes



LINEAR RECURRENT NETWORKS

1. Activation function: $\sigma(x_j) = \dfrac{1}{2} + \dfrac{x_j}{4}$
2. Weights eigendecomposition: $W = P\Lambda P^{-1}$
3. Linear change of variable: $z = P^{-1}\left(\dfrac{1}{4}W - I_N\right)x + P^{-1}\left(\dfrac{1}{2}W1_N + \theta\right)$
4. Eigenfunctions: $\psi(z) = \prod_{i=1}^{N} z_i^{\mu_i}$ with $\lambda = \sum_{i=1}^{N}\left(\dfrac{1}{4}\Lambda_{ii} - 1\right)\mu_i$ for $\mu_i \in \mathbb{C}$

The KOG eigenvalues and eigenfunctions are directly determined by the spectral properties of the weight matrix W.

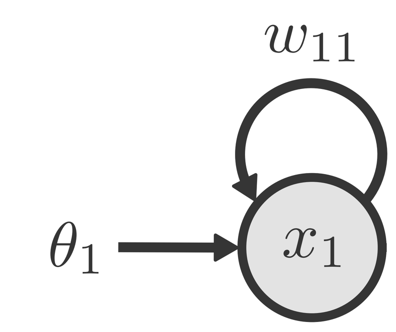
CIMMUL Centre Interdisciplinaire en Modélisation Mathématique de l'Université Laval · CRSNG NSERC · Fonds de recherche Québec · Sentinelle Nord Sentinel North · Digital Research Alliance of Canada · Alliance de recherche numérique du Canada

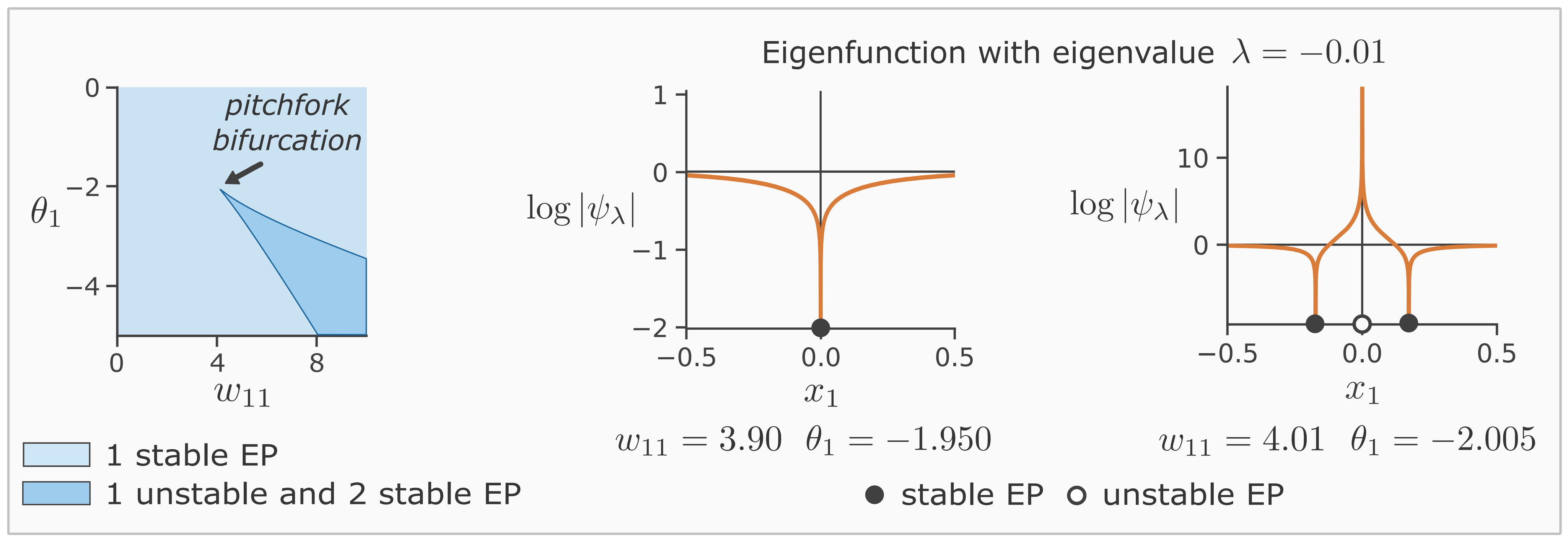
A SINGLE NONLINEAR NEURON

Activation function : $\sigma(x_1) = \dfrac{1}{2} + \dfrac{x_1}{4} - \dfrac{x_1^3}{48}$

Eigenfunctions : $\psi_\lambda(x_1) = (x_1 - z_1)^{\lambda c_1}(x_1 - z_2)^{\lambda c_2}(x_1 - z_3)^{\lambda c_3}$

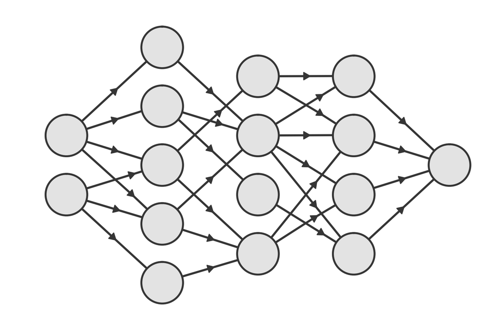
where z_1, z_2 and z_3 are the roots of the polynomial vector field.





The equilibrium points are observed through the zeros and singularities of the KOG eigenfunctions.

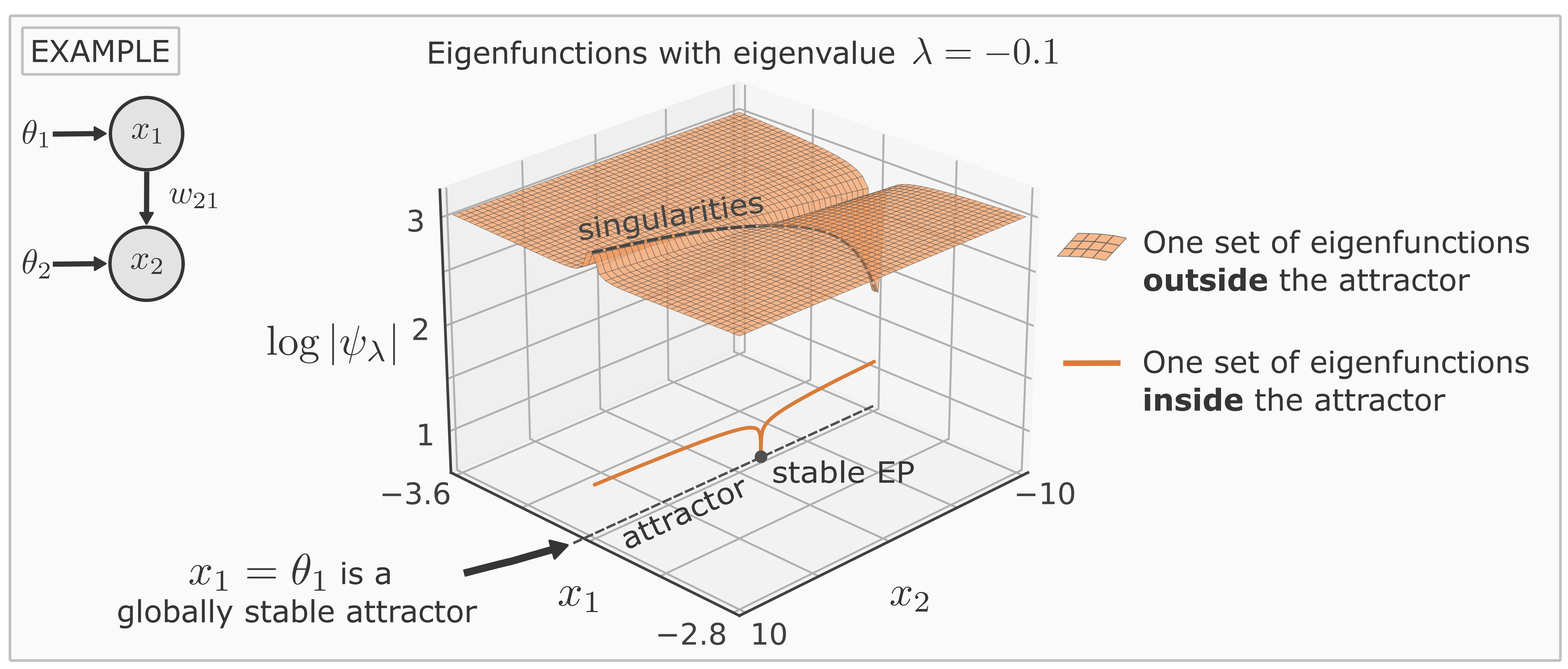
FEEDFORWARD NONLINEAR NETWORKS



Activation function : $\sigma(x_j) = \dfrac{1}{1 + e^{-x_j}}$

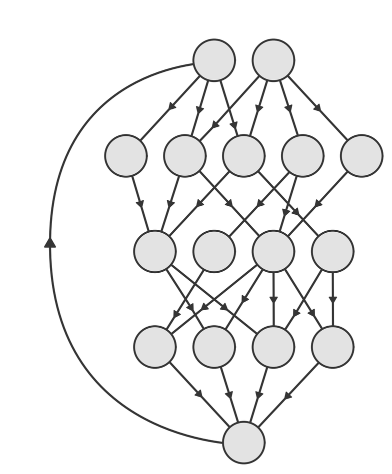
General eigenfunctions : $\psi_\lambda(x) = (x_1 - \theta_1)^{-\lambda}F(\Phi_2, \Phi_3, ..., \Phi_N)$ for $x_1 \neq \theta_1$

where $\Phi_i = \dfrac{x_i - \theta_i}{x_1 - \theta_1} + \sum_{j=1}^{i-1} w_{ij}\gamma_j(x_1)$ and $\gamma_j(x_1)$ is an integral-defined function.



There are distinct families of eigenfunctions defined inside and outside the attractors of the system.

TOWARDS NONLINEAR RECURRENCE



- The linear, single neuron and feedforward cases are all integrable systems, but nonlinear recurrent networks are not.
- Therefore, we aim to compute approximate eigenfunctions for nonlinear recurrent cases by applying perturbative methods on feedforward systems.
- Those eigenfunctions should give us insight on the relationship between the structure of the system and its attractors and bifurcations.