Brain-Inspired Computing
Class 2
Flexible representations

Alessio Franci

University of Liege

February 11, 2023

Contents

1. Representation in simple negative and positive feedback systems
2. Organizing bifurcations and flexible representations
3. Flexible multi-category representations
4. Flexible decision making

Representation in simple negative and positive feedback systems

Negative feedback leads to analog/faithful representations

Let's start with the simple input-output linear dynamics

$$\tau \dot{x}(t) = -x(t) + I(t) \tag{1}$$

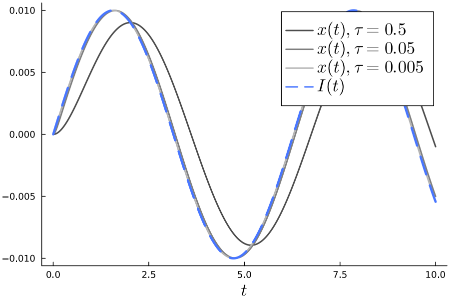
where $I(t)$ is the input and the state $x(t)$ is the output. The positive parameter τ fixes the characteristic timescale of dynamics (1). Equivalently, dynamics (1) can be interpreted as a low-pass filter with transfer function $\frac{1}{\tau s+1}$ and τ^{-1} is the filter cut-off frequency.

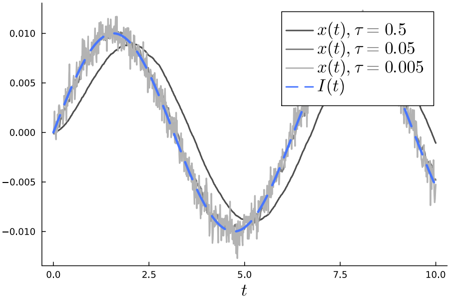
Dynamics (1) contain a single feedback loop of x on itself and this loop is negative because

$$\frac{\partial \dot{x}}{\partial x} = -\tau^{-1} < 0\,.$$

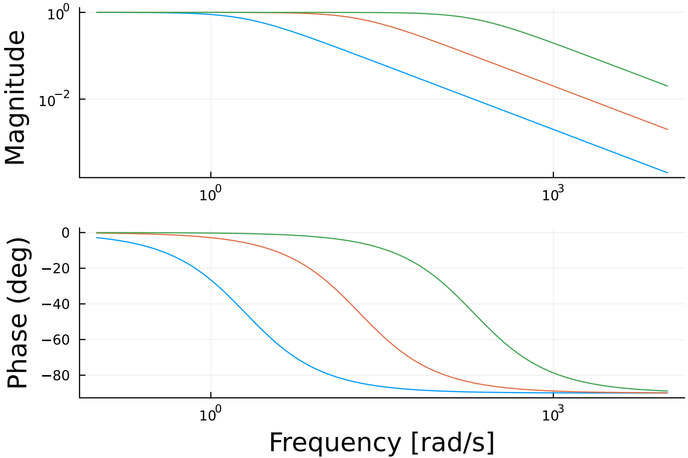
In other words, making the dynamics of x faster is equivalent to increasing the strength of the negative feedback loop.

Negative feedback leads to analog/faithful representations





Negative feedback leads to analog/faithful representations



Negative feedback leads to analog/faithful representations

In the absence of noise, as the negative feedback increases, the trajectory represent the input more and more closely. This is most naturally explained by computing $\lim_{\tau\to 0} \frac{1}{\tau s+1} = 1$, which means that in the limit of infinitely strong negative feedback dynamics (1) implement an all-pass filter. In other words, in an ideal (noise-free) setting, strong negative feedback implies faithful analog signal representation.

However, the all-pass filtering properties for small τ also let the high-frequency noise components pollute (be represented in) the representation. From a reference tracking perspective, strong negative feedback correspond to **high-gain control** because

$$\dot{x}(t) = \tau^{-1} \cdot (-x(t) + I(t))$$

For $\tau \to 0$, the state $x(t)$ perfectly tracks the input $I(t)$. As we just observed, perfect tracking is not necessarily a good way of representing an incoming signal, i.e., in all non-ideal cases in which noise pollutes the system. In practice, more complicated transfer functions (e.g., a Butterworth filter) could be used to obtain a faithful representation only in desired frequency ranges.

Remark: For simplicity, all the idea of the course we will be illustrated based on the simple dynamics (1). Generalizing (some of) them to more complicated linear dynamics can be the subject of final projects/TFEs.

Localized positive feedback leads to digital/categorical representations

Consider now the nonlinear dynamics

$$\tau\dot{x}(t) = -x(t) + \tanh(k\cdot x(t) + I(t))\,. \tag{2}$$

Observe that dynamics (1) are the small signal approximation (i.e., the linearization) of (2) for $k = 0$.

To uncover the feedback structure of dynamics (2), compute

$$\begin{aligned}\frac{\partial \dot{x}}{\partial x} &= \tau^{-1}\left(-1 + k\cdot\tanh'(k\cdot x + I)\right)\\ &= -\tau^{-1} + \tau^{-1}\cdot k\cdot(1 - \tanh(k\cdot x + I)^2)\,.\end{aligned}$$

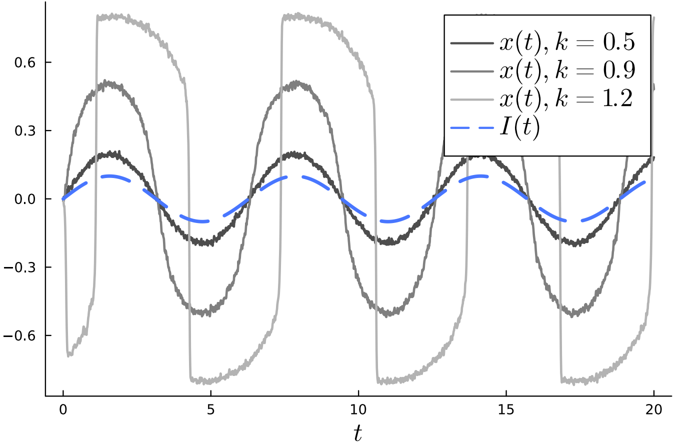
The first term in the last expression corresponds to the same negative feedback loop with gain τ^{-1} of dynamics (1). The second term corresponds to a **positive feedback loop** with **localized gain** $\tau^{-1}\cdot k$.

Localized positive feedback leads to digital/categorical representations

The **positive feedback is localized** both with respect to the state and the input strength because $(1 - \tanh(x + I)^2) \simeq 1$ for small x and I but $(1 - \tanh(x + I)^2) \to 0$ for large x or large I.

Localization of the positive feedback gain ensures global stability of the dynamics and, in particular, that $x(t)$ is bounded for all time. More precisely, it is easy to check that if $|x(0)| < 1$ then the solution $x(t)$ to (2) satisfies $|x(t)| < 1$ for all $t \geq 0$.
Exercise: Prove it.

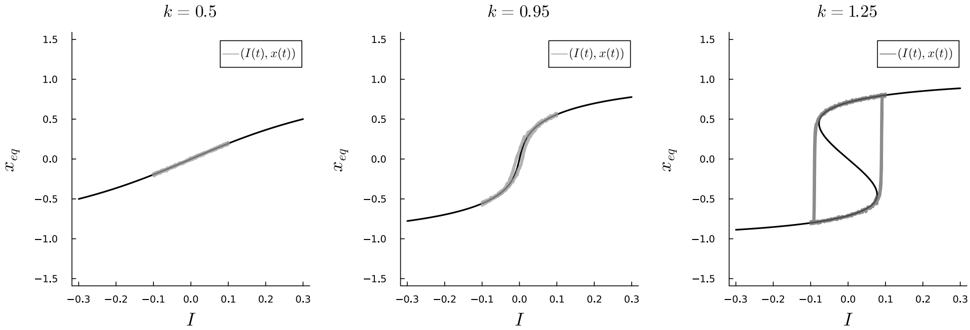
Localized positive feedback leads to digital/categorical representations



Localized positive feedback leads to digital/categorical representations

Increasing k leads to a smooth **transition from an analog and faithful to a digital and categorical representation** of the stimulus. For $k < 1$, positive feedback has the effect of amplifying the stimulus in a weakly non-linear fashion, due to saturation and positive feedback localization. For $k \gtrsim 1$ the state $x(t)$ exhibits sharp transitions from large positive to large negative values (or vice-versa) as the stimulus changes sign. In other words, $x(t)$ mostly represents the sign, i.e., a categorical feature, of the stimulus and only weakly its amplitude. As k grows further above 1, the categorical nature of the state becomes increasingly pronounced. Furthermore, **hysteresis** appears in the representation, meaning that a sufficiently large change in sign must happen for the represented category to change, which constitutes a basic form of memory.

Localized positive feedback leads to digital/categorical representations



Digital/categorical representations and bistability

From a dynamical systems perspective, the transition from faithful/analog to categorical/digital representations corresponds to the transition from a single stable equilibrium to multiple stable equilibria as k grows. For τ small (i.e., strong negative feedback/high-gain tracking) trajectories slides along the **equilibrium bifurcation diagram** defined by

$$\{(I,x)\,:\,-x+\tanh(k\cdot x+I)=0\}$$

Let's convince ourselves that this bifurcation diagram has the shape of a monotone increasing function for $k<1$ and that it is S-shaped for $k>1$.

Let's also study the stability of the branches of this bifurcation diagram and plot the associated vector field for fixed values of I. (If you feel rusty on 1D nonlinear dynamical systems analysis, please see [5, Chapter 2].)

Positive feedback leads to digital/categorical representations

Thus, for $k < 1$, there is a unique equilibrium x_{eq} for all values of I. And this equilibrium is stable. The dynamics

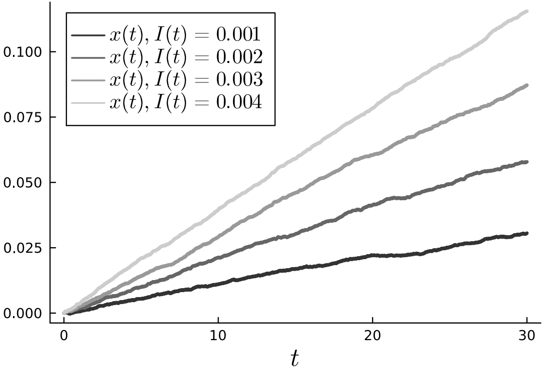
$$\dot{x}(t) = \tau^{-1} \cdot (-x(t) + \tanh(k \cdot x(t) + I(t)))$$

tracks this implicitly defined stable equilibrium through high-gain negative feedback control (for small τ) as $I(t)$ changes.

For $k > 1$, there is an interval of values of I for which two stable equilibria coexist. The system is **bistable**. Each equilibrium corresponds to a represented category. Initial conditions (i.e., the system memory) determine which category is represented. Because of the hysteresis associated to bistability, a sufficiently large change in inputs is needed to switch category.

Exercise: What happens to the representation when τ grows? Informally explain from a linear filtering perspective.

Almost perfect integration from the balance of positive and negative feedback

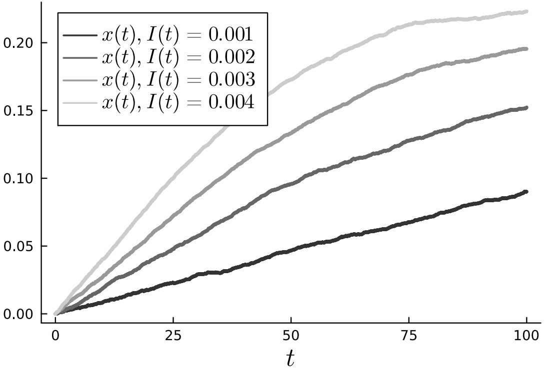


When $k = 1$ positive and negative feedback are perfectly balanced in dynamics (2). In this case, the linearization at the origin boils down the perfect integrator

$$\tau \dot{x}(t) = I(t)$$

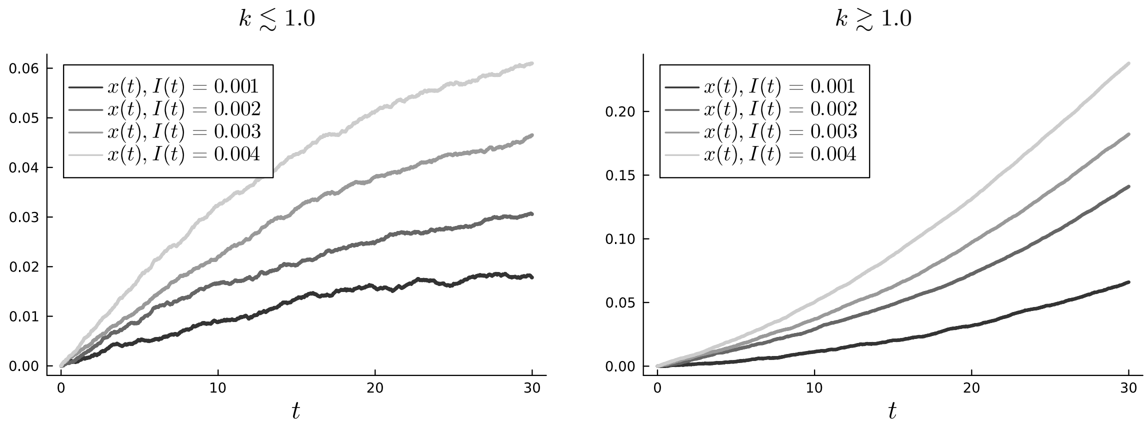
Perfect integration corresponds to a linear marginally stable behavior with infinite input-output gain, which underlies slow amplification of even tiny signals. In this sense, perfect integration constitutes a weak form categorical representation because positive (negative) signals are mapped to positive (negative) states with (local) infinite gain.

Almost perfect integration from the balance of positive and negative feedback



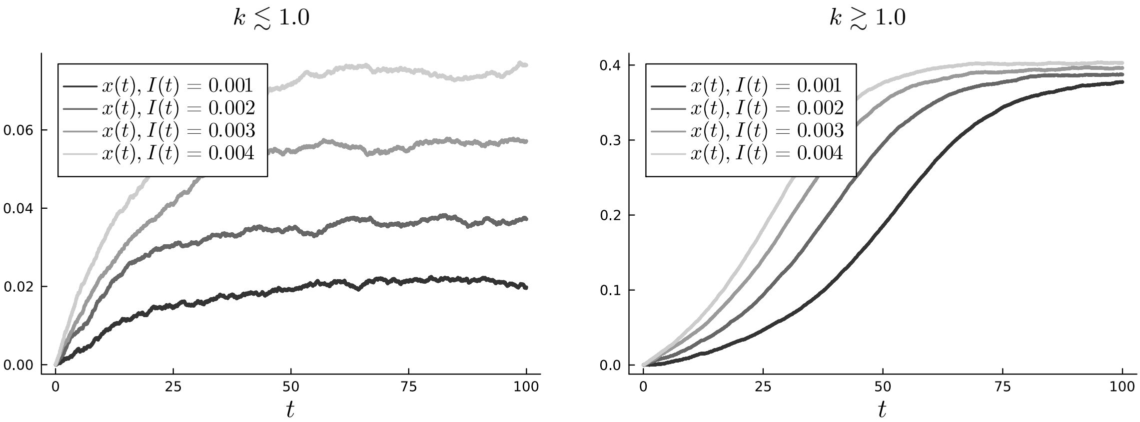
For longer times, the effects of the saturation non-linearity become evident.

Almost perfect integration from the balance of positive and negative feedback



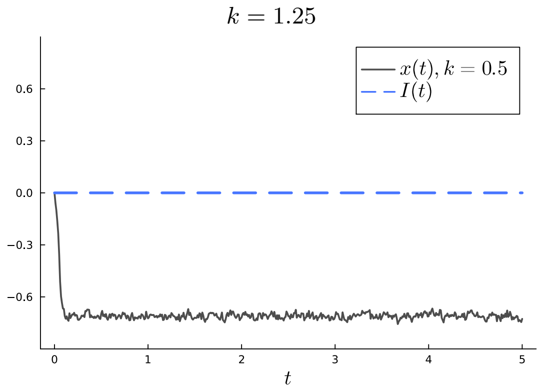
When positive and negative feedback are not perfectly balanced, integration becomes either weakly damped for $k \lesssim 1$ or weakly non-linearly amplified for $k \gtrsim 1$. This allows rapid transition from weakly sensitive signal integration to highly sensitive signal integration as a function, for instance, of context.

Almost perfect integration from the balance of positive and negative feedback



For larger time, damped integration leads to a faithful representation of the input strength, whereas amplified integration leads to a categorical representation of the input sign.

An issue with our categorical representation model



Because the only two stable equilibria for $I \approx 0$ correspond either to a high or a low category, even tiny inputs, or zero inputs (in the presence of noise) are mapped to either the high or the low category.

In general, this is not always desirable (unless one wants to throw a coin and pick a random category - we will see that this can sometime be useful).

The emergence of a stable neutral category from state-dependent positive feedback

A solution to the issue above is to modify the representing dynamics in such a way that a stable "neutral" category can coexist with the high and low category.

$\longrightarrow$ We want our representation to **represent faithfully weak inputs** (not clearly pointing to any category and possibly corrupted by noise) and to **represent categorically strong inputs** (clearly pointing to a category).

$\longrightarrow$ We can achieve this by letting the positive feedback depend on the state, i.e., letting

$$k = \tilde{k}(x) := k_0 + K_x \cdot x^2 \tag{3}$$

in dynamics (2), with $k_0, K_x \geq 0$, which leads to

$$\tau \dot{x}(t) = -x(t) + \tanh\left(\tilde{k}(x) \cdot x + I(t)\right) . \tag{4}$$

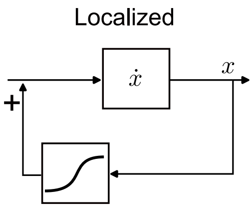
The emergence of a stable neutral category from state-dependent positive feedback

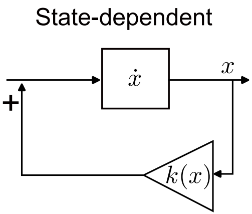
The resulting state-dependent feedback gain is

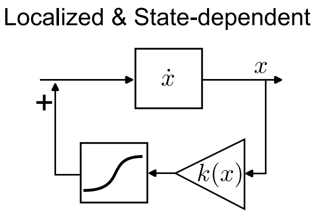
$$\begin{aligned}\frac{\partial \dot{x}}{\partial x} &= \tau^{-1}\cdot\left(-1+\left(\tilde{k}(x)+\tilde{k}'(x)\cdot x\right)\cdot\tanh'\left(\tilde{k}(x)\cdot x+I\right)\right)\\ &= -\tau^{-1}+\tau^{-1}\cdot\left(k_0+3K_x\cdot x^2\right)\cdot\left(1-\tanh\left(\tilde{k}(x)\cdot x+I\right)^2\right).\end{aligned}$$

The first term in the last expression corresponds to the same negative feedback loop with gain τ^{-1} of dynamics (1) and (2). The second term corresponds to a positive feedback loop with **localized and state-dependent** gain $\tau^{-1}\cdot\left(k_0+3K_x\cdot x^2\right)$.

Gain localization vs Gain state-dependence

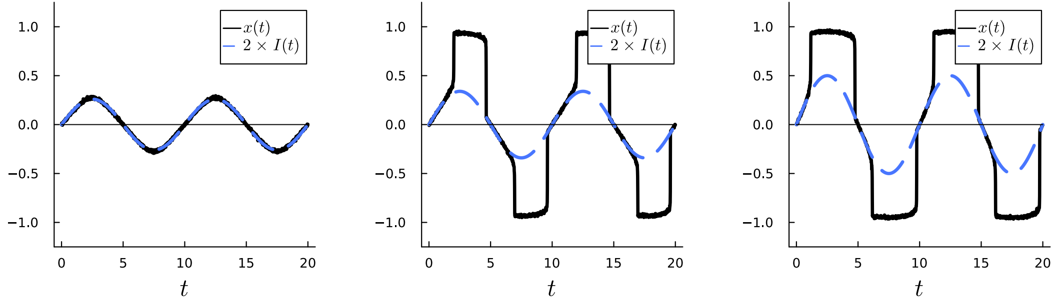






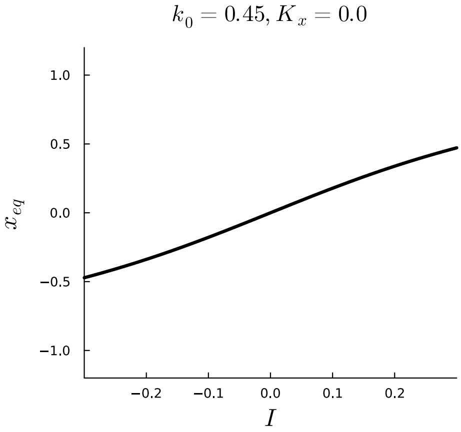
The emergence of a stable neutral category from state-dependent positive feedback

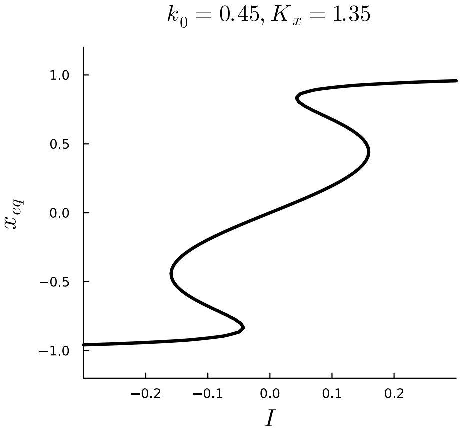
By picking $0 < k_0 < 1$ and $K_x > 0$ and large enough:



Rationale: for $0 < k_0 < 1$ and $K_x > 0$ and large enough the feedback gain $\frac{\partial \dot{x}}{\partial x}$ changes sign four times, which leads to a "double hysteresis" bifurcation diagram. Equivalently, the function $\tanh\left(\left(k_0 + K_x \cdot x^2\right) \cdot x + I\right)$ has four kinks.

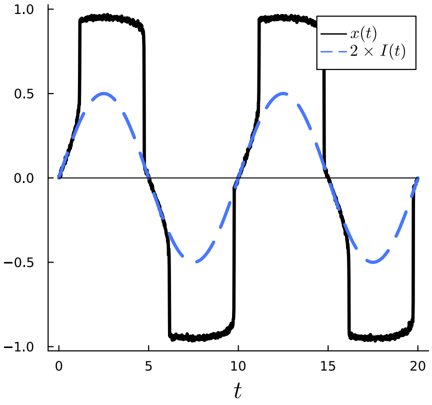
The emergence of a stable neutral category from state-dependent positive feedback

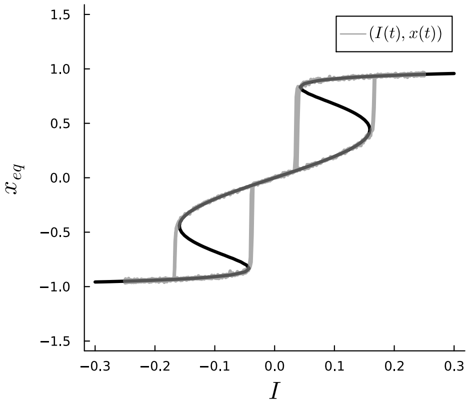




The emergence of a stable neutral category from state-dependent positive feedback

As, in the single hysteresis case, a time-varying input lets the state slide along the $\{(I, x)\}$-bifurcation diagram (provided, again, the characteristic timescale of input variation are slower than the state dynamics characteristic timescale τ)





A possible realization of state-dependent positive feedback

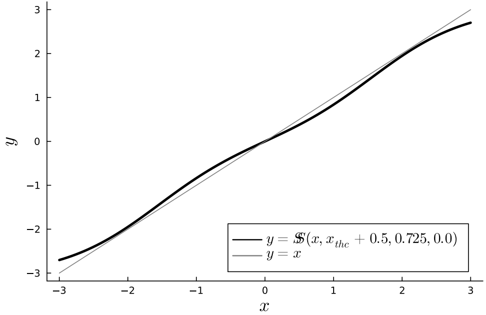
The way we introduced state-dependency in the positive feedback gain, i.e., letting $k = \tilde{k}(x) = k_0 + K_x \cdot x^2$ is mathematically useful but potentially hard to implement in an engineering context, like during the design of a neuromorphic chip.

How can we translate to a more realizable model?

Consider the sum-of-2-shifted-sigmoids function

$$\mathbb{S}(x, x_{th}, k, I) = \frac{\tanh(kx - x_{th} + u) + \tanh(kx + x_{th} + u)}{2\tanh'(x_{th})}.$$

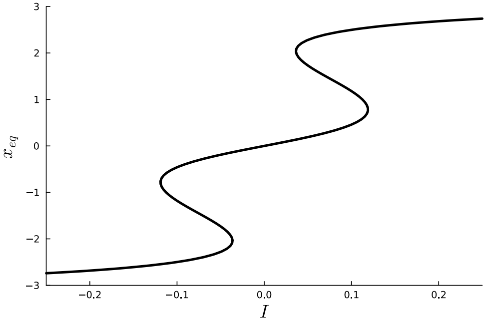
Its graph for $x_{th} = x_{th,c} + 0.5$, $k = 0.725$, $I = 0.0$, where $x_{th,c} = 0.5 \cdot \cosh^{-1}(2)$, is



The bifurcation diagram of

$$\tau\dot{x} = -x + \mathbb{S}(x, x_{th}, k, I)$$

for the same parameter values is



A possible realization of state-dependent positive feedback

We can overcome this limitation using the geometric rational underlying the appearance of the double hysteresis. The sum of two shifted $\tanh(\cdot)$ functions

$$\mathbb{S}(x, x_{th}, k, u) = \frac{\tanh(kx - x_{th} + u) + \tanh(kx + x_{th} + u)}{2\tanh'(x_{th})}$$

exhibit (for $|x_{th}|$ large enough) the same four kinks as $\tanh\left(\left(k_0 + K_x \cdot x^2\right) \cdot x + I\right)$. Bifurcation theory will help us formalizing this kind of ideas.

Using geometry and bifurcations jointly with feedback analysis is a powerful methodology for designing nonlinear input-output behaviors [2, 1, 4, 3].

[1] Guillaume Drion et al. "A novel phase portrait for neuronal excitability". In: *PLoS One* (2012).

[2] Alessio Franci, Guillaume Drion, and Rodolphe Sepulchre. "Modeling the modulation of neuronal bursting: a singularity theory approach". In: *SIAM Journal on Applied Dynamical Systems* 13.2 (2014), pp. 798–829.

[3] Alessio Franci and Rodolphe Sepulchre. "Realization of nonlinear behaviors from organizing centers". In: *53rd IEEE Conference on Decision and Control*. IEEE. 2014, pp. 56–61.

[4] Alessio Franci et al. "A balance equation determines a switch in neuronal excitability". In: *PLoS Computational Biology* 9.5 (2013), e1003040.

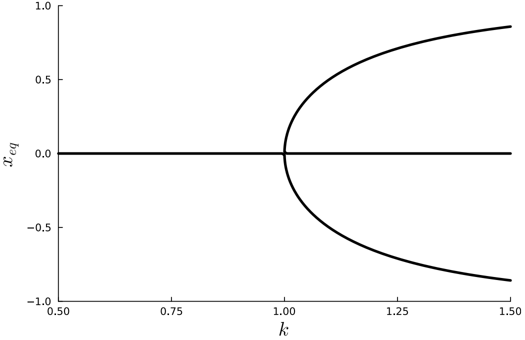
[5] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

Organizing bifurcations and flexible representations

The transition from analog to digital through a pitchfork

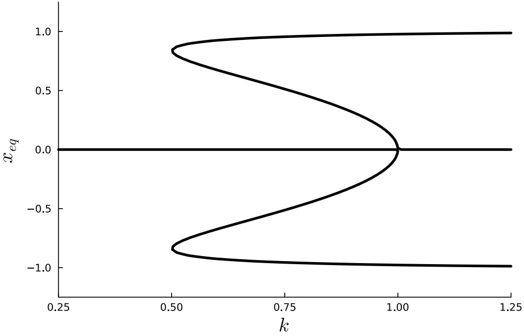
The $\{(I, x)\}$ bifurcation diagram of $\tau\dot{x} = -x + \tanh(k \cdot x + I)$ for different values of k is useful to develop a geometrical understanding of how input variations are mapped to state variations as a function of the positive feedback strength.

We now study its $\{(k, x)\}$ bifurcation diagram for $I = 0$ to visualize the transition from analog/faithful to digital/categorical representations as a bifurcation. The bifurcation is a pitchfork.

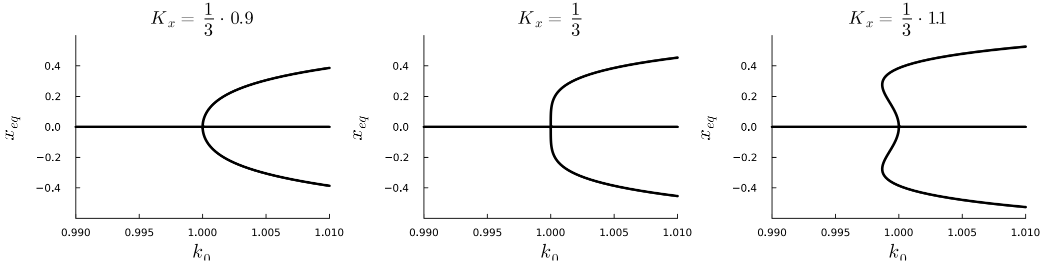


Existence of a stable neutral category: subcritical pitchfork

The $\{(k_0, x)\}$ bifurcation diagram of the state-dependent flexible representation $\tau\dot{x} = -x + \tanh\left((k_0 + K_x \cdot x^2)\cdot x + I\right)$ with a stable neutral category is also organized by a pitchfork. But this time the bifurcation is **subcritical**.



The organizing center of flexible representations

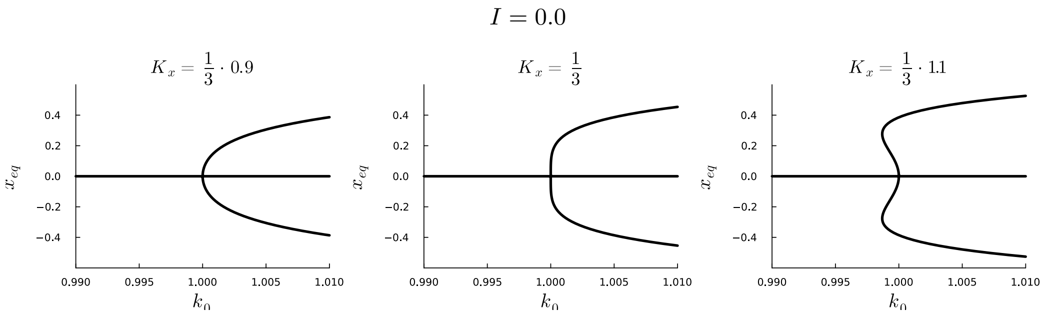


The organizing center of flexible representations

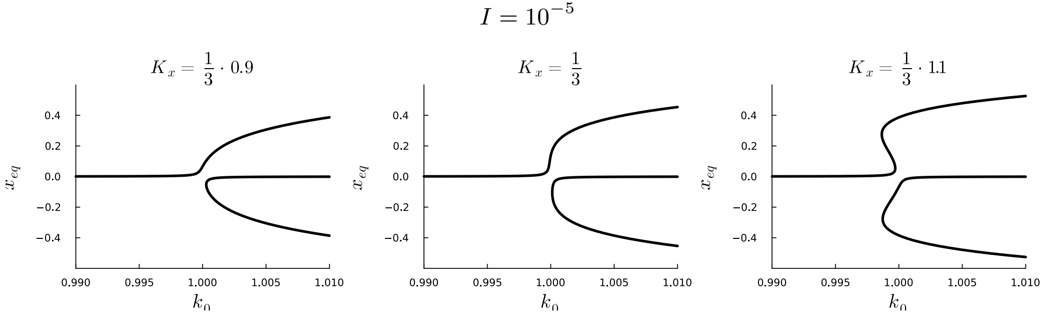
For flexibility, a given representation might need to be modulated between having or not a stable neutral category. For instance, in some cases it might be convenient that, in order to **break indecision**, the representation flips a coin to assign a category even to noisy/weak inputs, while in other cases it might be convenient to remain patient (and undecided) until better information is available. Interestingly, honeybee societies pass from the latter to the former behavior when deciding for a new nest site and multiple sites with similar qualities are available [2].

The degenerate situation corresponding to the transition from the coin-flipping representation to the patient representation is mathematically associated to a **higher-codimension bifurcation** (or "**organizing center**") [1]. In this case, the organizer is a quintic pitchfork.

Subcritical pitchfork leads to enhanced input sensitivity and discontinuous analog/faithful to digital/categorical transition



Subcritical pitchfork leads to enhanced input sensitivity and discontinuous analog/faithful to digital/categorical transition



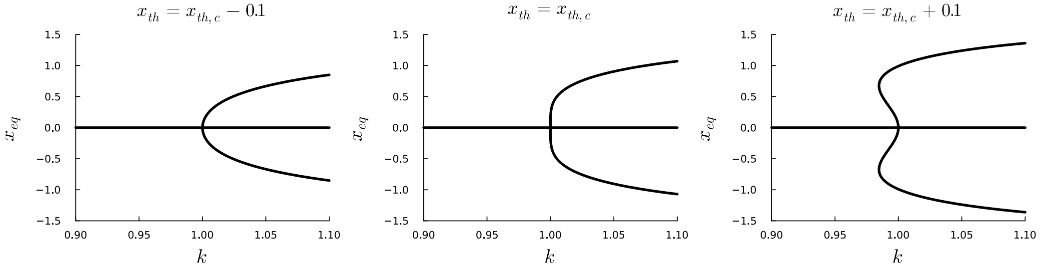
Subcritical pitchfork leads to enhanced input sensitivity and abrupt analog/faithful to digital/categorical transition

The co-existence of a stable neutral category and two stable representing categories on the one hand makes the system more patient because strong enough inputs are needed to switch to one of the two representing categories. But on the other hand it also makes the system respond with enhanced sensitivity to inputs when it is close to the pitchfork bifurcation.

Furthermore, whereas in the supercritical pitchfork case the transition from analog/faithful to digital/categorical representation is continuous (i.e., representing categories emerges and differentiate slowly), in the subcritical case it becomes more abrupt (i.e., representing categories are sharply separated already at bifurcation).

Organizing centers as a form of model equivalence

The $\{(k,x)\}$ bifurcation diagrams of the "neuromorphic realization" $\tau\dot{x} = -x + \mathbb{S}(x, x_{th}, k, I)$ are organized by the same organizing center as their mathematical counterpart



Organizing centers as a form of model equivalence

We can formally prove that the two model are equivalent by showing that they are organized by the same organizing center. That is, locally they both "look like" the quintic pitchfork model

$$\tau\dot{x} = \alpha x + \beta x^3 - x^5 + I$$

where

$$\alpha = k_0 - 1$$
$$\beta = K_x - 1/3$$

in model $\tau\dot{x} = -x + \tanh\left((k_0 + K_x \cdot x^2)\cdot x + I\right)$

$$\alpha = k - 1$$
$$\beta = x_{th} - x_{th,c},$$

in model $\tau\dot{x} = -x + \mathbb{S}(x, x_{th}, k, I)$, where $x_{th,c} = 0.5 \cdot \cosh^{-1}(2)$.

Organizing centers can be "recognized" in a given model using the algebraic techniques of the **Recognition Problem** [1]. They are a bifurcation-theory version of the **Taylor Expansion Theorem**.

[1] M. Golubitsky and D.G. Schaeffer. *Singularities and Groups in Bifurcation Theory*. Vol. 1. Springer-Verlag, 1985.

[2] Darren Pais et al. "A mechanism for value-sensitive decision-making". In: *PloS one* 8.9 (2013), e73216.

Flexible multi-category representations

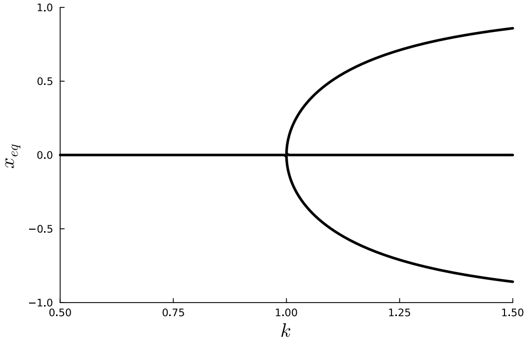
Flexible multi-category representation

The idea explored until here provide a framework for flexible **2-category** representation?

How can we generalize those ideas to n-**category** representations?

We will see that the notions of **symmetry** and **symmetric bifurcations** provide the right tools to achieve this generalization

Unbiased categorical representations and symmetric bifurcations



When there are no biases toward either category ($I = 0$), the emergence of the two representing categories in model $\tau\dot{x} = -x + \tanh(k\cdot x + I)$ has **symmetry**, which reflect that in the absence of biases **the two categories are indistinguishable**.

Formally, if $\sigma\, x = -x$ and we let $F(x, k, I) = -x + \tanh(k\cdot x + I)$, then $\sigma\, F(x, k, 0) = F(\sigma\, x, k, 0)$, that is, in the terminology of equivariant bifurcation theory [4, 3], the representing dynamics $\tau\dot{x} = F(x, k, 0)$ is **equivariant** (has symmetry) $\mathbb{Z}_2 = (1, \sigma)$.

"Having symmetry" means that if $x(t)$ is a solution to $\tau\dot{x} = F(x, k, 0)$, than also $\sigma, x(t)$ is a solution.

2-nodes 2-category representation

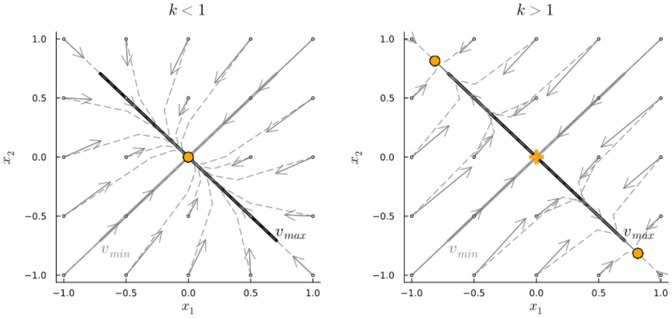
Consider model

$$\dot{x}_1 = -x_1 + \tanh(-k \cdot x_2 + I_1)$$
$$\dot{x}_2 = -x_2 + \tanh(-k \cdot x_1 + I_2)$$

of two representing nodes receiving two signals and mutually inhibiting each other. Let's represent its "motif":

2-nodes 2-category representation

The model undergoes a symmetric **pitchfork bifurcation along the subspace** $\{x_2 = -x_1\}$



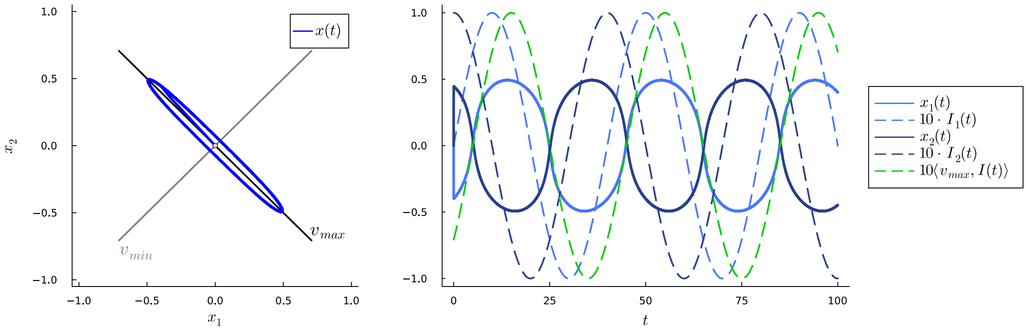
Let's convince ourselves this is the case: *i)* Use Lyapunov to show that $\{x_2 = -x_1\}$ is GES; *ii)* Study bifurcation for $x_2 = -x_1$.

For more complicated systems: Lyapunov-Schmidt reduction techniques [2, 4, 3].

2-nodes 2-category representation

Category A: $I_1 > I_2$. Category B: $I_2 > I_1$

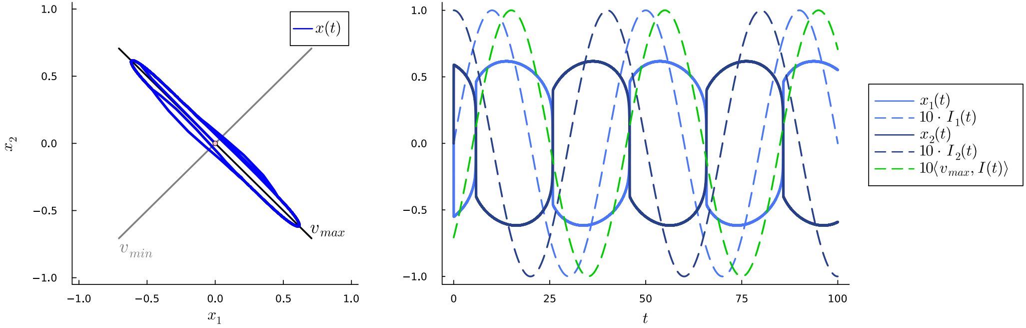
$$k < 1$$



2-nodes 2-category representation

Category A: $I_1 > I_2$. Category B: $I_2 > I_1$

$k > 1$



Generalization two a larger number of categories

Generalization two a larger number of categories can be derived by symmetry arguments:

Key observation: Model

$$\dot{x}_1 = -x_1 + \tanh(-k \cdot x_2 + I_1)$$
$$\dot{x}_2 = -x_2 + \tanh(-k \cdot x_1 + I_2)$$

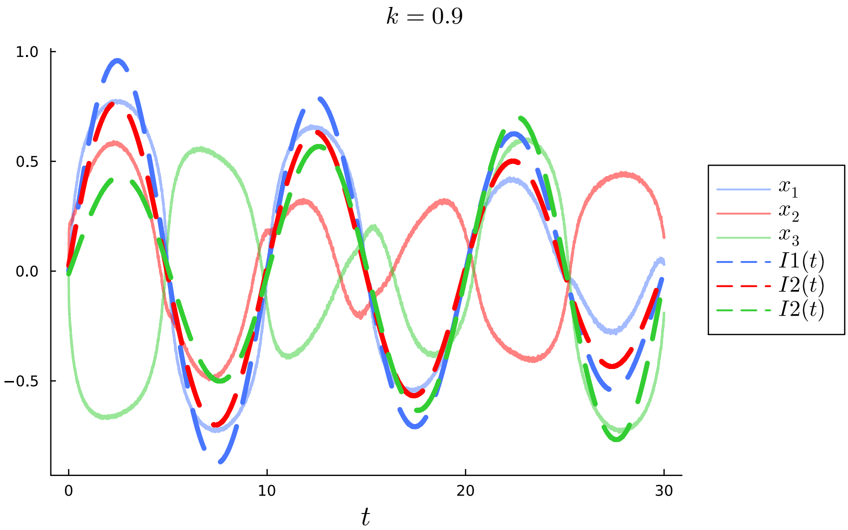
is symmetric with respect to simultaneously swapping $x_1 \leftrightarrow x_2$ and $I_1 \leftrightarrow I_2$.

3-nodes 6-category representation

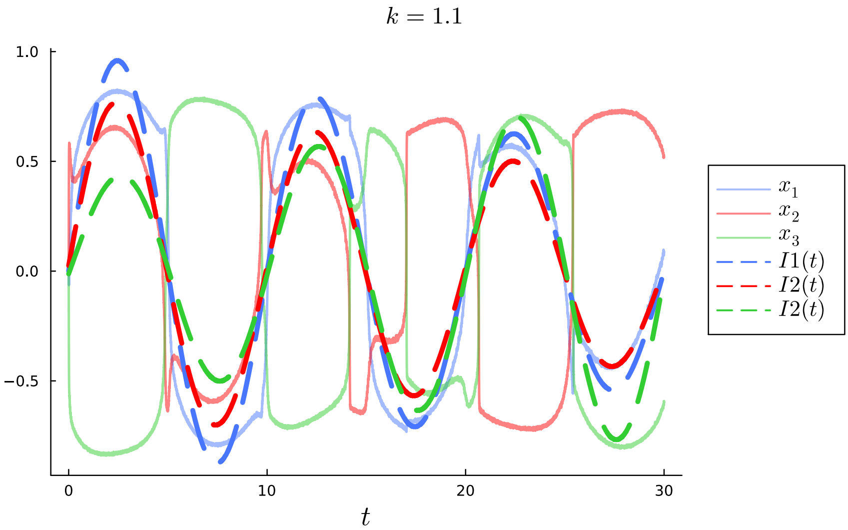
Then, a three-node, 6-category representation can be obtained in representing dynamics that are symmetric with respect to simultaneously permuting $\{x_1, x_2, x_3\}$ and $\{I_1, I_2, I_3\}$, for example

$$\dot{x}_1 = -x_1 + \tanh(-k \cdot x_2 - k \cdot x_3 + I_1)$$
$$\dot{x}_2 = -x_2 + \tanh(-k \cdot x_1 - k \cdot x_3 + I_2)$$
$$\dot{x}_3 = -x_3 + \tanh(-k \cdot x_1 - k \cdot x_2 + I_3)$$

3-nodes 6-category representation



3-nodes 6-category representation



3-nodes 6-category representation

How do we prove such a behavior?

The two main techniques are **center manifold reduction** [1] and **Lyapunov-Schmidt reduction** [2].

They both rely on the observation that the linearization J (the Jacobian) of model

$$\dot{x}_1 = -x_1 + \tanh(-k \cdot x_2 - k \cdot x_3 + I_1)$$
$$\dot{x}_2 = -x_2 + \tanh(-k \cdot x_1 - k \cdot x_3 + I_2)$$
$$\dot{x}_3 = -x_3 + \tanh(-k \cdot x_1 - k \cdot x_2 + I_3)$$

at the origin and for $I_i = 0$ has two zero eigenvalues and its kernel $\ker(J) = \{x_1 + x_2 + x_3 = 0\}$, in analogy with the 2-nodes 2-category case, where $\ker(J) = \{x_1 + x_2 = 0\}$.

Then we can "project" dynamics onto $\ker(J)$ to derive the representation properties. The process works in arbitrary dimension and for an arbitrary number of categories. Part of the symmetry can be broken to further shape the representation properties.

[1] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. 7th. Vol. 42. Applied Mathematical Sciences. New-York: Springer, 1983.

[2] M. Golubitsky and D.G. Schaeffer. *Singularities and Groups in Bifurcation Theory*. Vol. 1. Springer-Verlag, 1985.

[3] M. Golubitsky and I. Stewart. *The symmetry perspective: from equilibrium to chaos in phase space and physical space*. Vol. 200. Birkh auser Basel, 2002.

[4] M. Golubitsky, I. Stewart, and D.G. Schaeffer. *Singularities and Groups in Bifurcation Theory*. Vol. 2. Springer-Verlag, 1988.

Flexible decision making

Mostly connection with neuroscience. Roman also has a bunch of stuff on "categories".