A dreaming machine with ghost memory states

Alexander Gorban and Ivan Tyukin*
University of Leicester

University of Leicester

"To sleep: perchance to dream: ay, there's the rub..."

William Shakespeare - To be, or not to be (from Hamlet 3/1)

<u>Brain</u>

Dynamical system

High dimensional

Reacts to stimulation

Memory is our imagination about past

1. Introduction. Framework

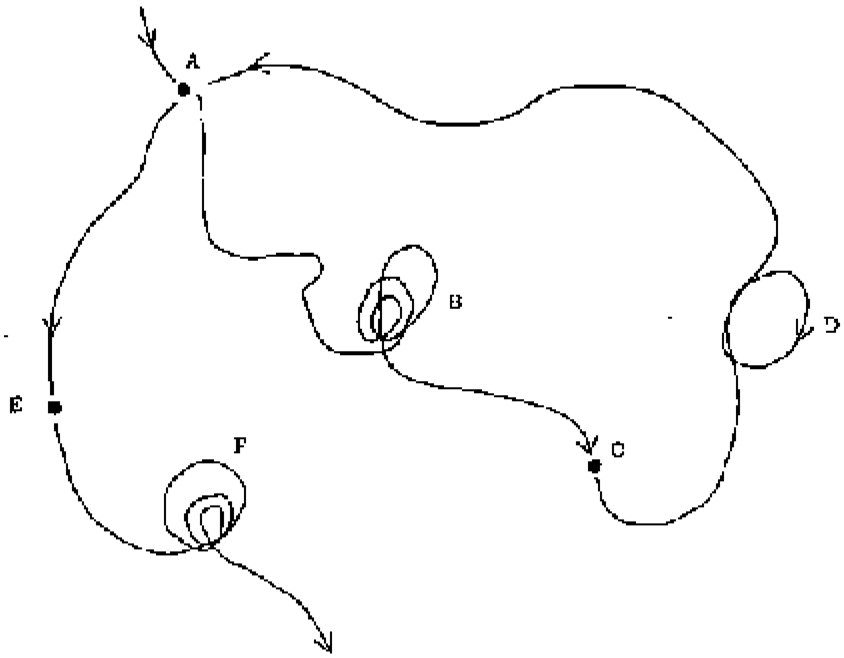
M. Hirsch, B. Baird, "Computing with dynamic attractors in neural networks", Biosystems **34, 173-195 (1995).**

"We view a computational medium as a set of structurally stable input-output subsystems which can be coupled in various ways into a larger system. By 'structurally stable' we mean that the dynamical behavior of each subsystem is largely immune to small perturbations due to noise or parameter changes. We assume that the dynamics of each subsystems is organized into attractor basins; the attractors can be stationary, periodic or chaotic. As the overall system evolves in time, each subsystem passes through a sequence of attractors, some function of which is presented to the observer as the 'output' of the system. These sequences of attractors are the 'computation' of the system".

Computing with Trajectories (I. Tsuda)

*From the mathematical point of view, ... **orbits linking attractors are important**. But there is no reason to stop at this point. The attractors are, strictly speaking, never reached and must be unstable in certain directions, so it is equally justified to speak of orbits that link, or connect, other orbits. What is the role of attractors at all? **A probable answer is that a high-dimensional system can only perform effective computations if it behaves like a lower dimensional system**. Chaotic itinerancy achieves this by permitting the orbits to enter the vicinity of attractors, thereby significantly reducing the dimensions. However, generalizing the same idea, there is no reason why it should not be possible to obtain other kinds of 'piecewise low-dimensional systems' which are based entirely on transients, and correspond to computations with trajectories distant from attractors.*

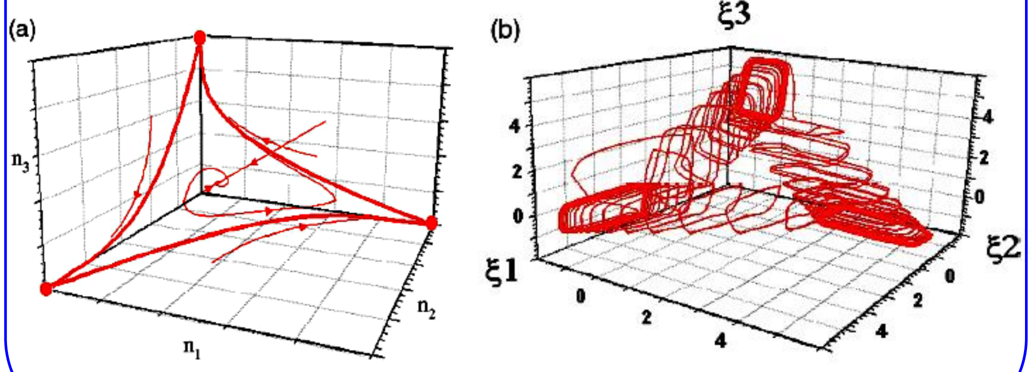
1. Introduction. Framework



Informal description of such processes by I.Tsuda: “attractor ruins” and “chaotic itinerancy”

1. Introduction. Framework

Winner-less competition in neuroscience
(Rabinovich, 2006)



heteroclinic connections

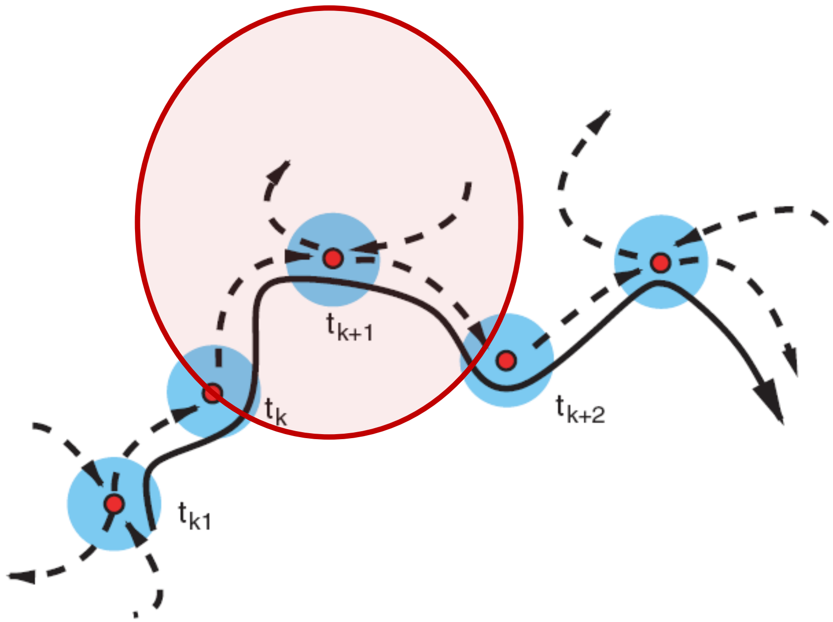


Figure 1. Schematic representation of a stable heteroclinic channel. The SHC is built with trajectories that condense in the vicinity of the saddle chain and their unstable separatrices (dashed lines) connecting the surrounding saddles (circles). The thick line represents an example of a trajectory in the SHC. The interval $t_{k+1}-t_k$ is the characteristic time that the system needs to move from the metastable state k to the k+1.

M. Rabinovich et al. 2008, PLOS Comp. Biology

1. Introduction. Problem

Phenomenon (postulates)

- A high-dimensional dynamical system evolving on a compact set

- Trajectories do not settle on any attracting sets of lower dimension (“dreaming”)

- Yet, it interacts with environment; Sensitive to small perturbations (“reduces” its dimension, “ghosts”)

- Computes something or makes decisions

Problems

- Mathematical definition of “attractor ruins” ?

- Machinery of analysis for these objects ?

- Constructive approaches for modelling ?

- Analysis of computational power of such machines ?

2. Concepts. Definitions

Let $\mathcal{A}$ be a subset of $\mathbb{R}^n$, and $\mathcal{V}(\mathcal{A}, \varepsilon)$ be its ε-neighborhood

1. Original system $\quad \Sigma: \ \dot{x} = f(t, x, u)$

$f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ - continuous

$u \in \mathbb{R}^m$ is the vector of inputs

---

2. A companion system (perturbed) $\quad \Sigma_\Delta: \ \dot{x} = f(t, x, u) + \delta(t), \ \|\delta\|_\infty \leq \Delta$

Without loss of generality we can assume that the state of the system evolves on a compact (and that the systems are forward-complete)

2. Concepts. Definitions

Definition 1 (Delay time) *Consider a solution* $x(t, x_0)$ *of system* Σ *passing through* $x' \in \mathcal{V}(\mathcal{A}, \varepsilon)$ *at* t'.

1) The function $T^+(x', \mathcal{V}(\mathcal{A}, \varepsilon))$

$$T^+(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) = \sup_t \{t - t',\ t \geq t' \mid x(\tau, x') \in \mathcal{V}(\mathcal{A}, \varepsilon)\ \forall\ \tau \in [t', t]\}$$

is the delay time of $x(t, x_0)$ *in forward time in* $\mathcal{V}(\mathcal{A}, \varepsilon)$ *at* (x', t').

2) The function $T^-(x', \mathcal{V}(\mathcal{A}, \varepsilon))$

$$T^-(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) = \sup_t \{t' - t,\ t \leq t' \mid x(\tau, x') \in \mathcal{V}(\mathcal{A}, \varepsilon)\ \forall\ \tau \in [t, t']\}$$

is the delay time of $x(t, x_0)$ *in backward time in* $\mathcal{V}(\mathcal{A}, \varepsilon)$ *at* (x', t').

---

maximal delay times

$$\sup_{x', t'} T^+(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)),\quad \inf_{x', t'} T^+(x', t', \mathcal{V}(\mathcal{A}, \varepsilon))$$

$$\sup_{x', t'} T^-(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)),\quad \inf_{x', t'} T^-(x', t', \mathcal{V}(\mathcal{A}, \varepsilon))$$

minimal delay times

2. Concepts. Definitions

Definition 2 (Inducible Delay time) *Let $x(t, x_0, \delta)$ be a solution of system Σ_Δ.*

1) The function $T^+_\Delta(x', \mathcal{V}(\mathcal{A}, \varepsilon))$:

$$T^+_\Delta(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) = \sup_{\delta,\ \|\delta(t)\|_\infty \leq \Delta} \sup_t \{t - t',\ t \geq t' \mid x(\tau, x', \delta) \in \mathcal{V}(\mathcal{A}, \varepsilon)\ \forall\ \tau \in [t', t]\}$$

is the inducible delay time of $x(t, x_0)$ in forward time in $\mathcal{V}(\mathcal{A}, \varepsilon)$ at (x', t').

2) The function $T^-_\Delta(x', \mathcal{V}(\mathcal{A}, \varepsilon))$

$$T^-_\Delta(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) = \sup_{\delta,\ \|\delta(t)\|_\infty \leq \Delta} \sup_t \{t' - t,\ t \leq t' \mid x(\tau, x', \delta) \in \mathcal{V}(\mathcal{A}, \varepsilon)\ \forall\ \tau \in [t, t']\}$$

is the inducible delay time of $x(t, x_0)$ in backward time in $\mathcal{V}(\mathcal{A}, \varepsilon)$ at (x', t')

$$T^+_\Delta(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) \geq T^+(x', t', \mathcal{V}(\mathcal{A}, \varepsilon))$$
$$T^-_\Delta(x', t', \mathcal{V}(\mathcal{A}, \varepsilon)) \geq T^-(x', t', \mathcal{V}(\mathcal{A}, \varepsilon))$$

Definition 3 (Δ-invariance) *Let $\Delta \in \mathbb{R}_{\geq 0} \in \mathbb{R}_{\geq 0}$ be given.*

1) A set $\mathcal{A}$ is Δ-positively invariant if

$$T_{\Delta}^{+}(x', t', \mathcal{V}(\mathcal{A}, 0)) = \infty, \ \forall \ x' \in \mathcal{A}.$$

2) A set $\mathcal{A}$ is Δ-backward invariant if

$$T_{\Delta}^{-}(x', t', \mathcal{V}(\mathcal{A}, 0)) = \infty, \ \forall \ x' \in \mathcal{A}.$$

3) A set $\mathcal{A}$ is Δ-invariant if it is both Δ-positively and backward invariant

Definition 4 (Pulling sets) *A set $\mathcal{A}$ is called pulling with respect to a set $\mathcal{U}$ iff for all $x' \in \mathcal{U}$ there exists $t \geq t'$:*

$$x(t, x') \in \mathcal{A}$$

Definition 5 *[Ghost attracting set] A Δ-positively invariant set $\mathcal{A}$ is called an (ε, Δ) ghost attracting (or simply ghost attracting) iff*

1) there exists a set $\mathcal{U}$ of positive measure and $\varepsilon \in \mathbb{R}_{\geq 0}$ such that $\mathcal{V}(\mathcal{A}, \varepsilon)$ is pulling with respect to $\mathcal{U}$;

2) $\mathcal{V}(\mathcal{A}, \varepsilon)$ is not positively invariant (admits τ_3-slow relaxations);

3) for every $x' \in \mathcal{U}$ there exists a function $\delta_{x'} : \mathbb{R} \to \mathbb{R}^n$, $\|\delta_{x'}(t)\|_\infty \leq \Delta$ such that

$$\lim_{t \to \infty} \mathrm{dist}(x(t, x', \delta_{x'}), \mathcal{A}) = 0.$$

Set $\mathcal{U}$ is the basin of attraction of $\mathcal{A}$.

An (ε, Δ) ghost attracting set is an (ε, Δ) *ghost attractor* iff the omega-limit set of $x(t, x', \delta'_x)$ coincides with $\mathcal{A}$ for all $x' \in \mathcal{U}$.

2. Concepts. Definitions

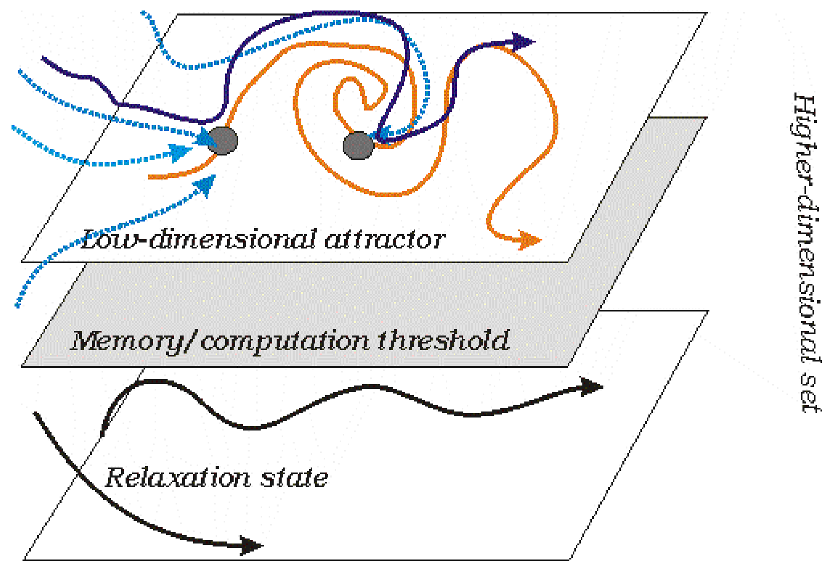
let $\{\varepsilon_i\}_{i=1}^{\infty}$ be a converging sequence of non-negative real numbers: $\lim_{i\to\infty} \varepsilon_i = 0$. A set $\mathcal{A}$ which is $(\varepsilon_i, 0)$ ghost attracting for all ε_i is weakly attracting.

Let $\mathcal{A}$ be a (weakly) attracting set. We say that $\mathcal{A}$ is *ε-persisting* if for some $\varepsilon \in \mathbb{R}_{>0}$ and sufficiently small $\Delta \in \mathbb{R}_{>0}$ the set $\mathcal{V}(\mathcal{A}, \varepsilon)$ is not a (ε, Δ) ghost.

A simple prototype

Space of all memory sets linked together by a dense trajectory

Sub-threshold dynamics



3. A minimal problem and first results

Problem: to be able to detect and describe mathematically creation of an attractor located on a dense trajectory …

3. A minimal problem and first results

If the model is a system of ODE's …

motions in the higher-dimensional set	$\dot{x} = f(x, \lambda, t)$	$f : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n,$
motions in the low-dimensional set	$\dot{\lambda} = g(x, \lambda, t),$	$g : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

Higher-dimensional, contracting

Assumption 1 *The function $f(\cdot,\cdot,\cdot)$ in (6) is locally Lipschitz in x and λ uniformly in t, and there exists $V : \mathbb{R}^n \to \mathbb{R}$, $V \in \mathcal{C}^1$ such that*

$$\underline{\alpha}(\|x\|) \le V(x) \le \bar{\alpha}(\|x\|),\ \underline{\alpha},\ \bar{\alpha} \in \mathcal{K}_\infty$$

$$\frac{\partial V}{\partial x} f(x, \lambda, t) \le \alpha(V(x)) + \beta(V(x))\varphi(|\lambda|),$$

$$\alpha, \beta \in \mathcal{C}^0[0, \infty),\ \alpha(0) = 0, \beta(0) = 0,\ \varphi \in \mathcal{K}.$$

Low-dimensional, exploring

Assumption 2 *The function $g(\cdot,\cdot,\cdot)$ in (6) is locally Lipschitz in x and λ uniformly in t, and there exist $\delta, \xi \in \mathcal{K}$ such that the following inequality holds for all $x \in \mathbb{R}^n$, $t \in \mathbb{R}$:*

$$-\xi(|\lambda|) - \delta(\|x\|) \le g(x, \lambda, t) \le 0\ \forall\ \lambda \ge 0.$$

3. A minimal problem and first results

From: Gorban, Tyukin, Steur, and Nijmeijer (submitted)

Lemma 1 (Boundedness 1) *Let system (6) be given and satisfy Assumptions 1, 2. Suppose that there exist a function*

$$\psi : \ \psi \in \mathcal{K} \cap \mathcal{C}^1(0,\infty)$$

and a positive constant $a \in \mathbb{R}_{>0}$ *such that*

$$\frac{\partial \psi(V)}{\partial V}\left[\alpha(V) + \beta(V)\varphi(\psi(V))\right] + \delta\left(\underline{\alpha}^{-1}(V)\right) + \xi\left(\psi(V)\right) \leq 0, \ \forall \ V \in [0,a].$$

Then the domain

$$\Omega_a = \{(x,\lambda) \mid x \in \mathbb{R}^n, \ \lambda \in \mathbb{R}_{\geq 0}, \ \psi(a) \geq \lambda \geq \psi(V(x)), \ V(x) \in [0,a]\}$$

is forward invariant with respect to (6), and furthermore

$$\exists \ \lambda' \in [0, \psi(a)] : \ \lim_{t\to\infty} \lambda(t) = \lambda',$$

and

$$\lim_{t\to\infty} g(x(t), \lambda(t), t) = 0.$$

3. A minimal problem and first results

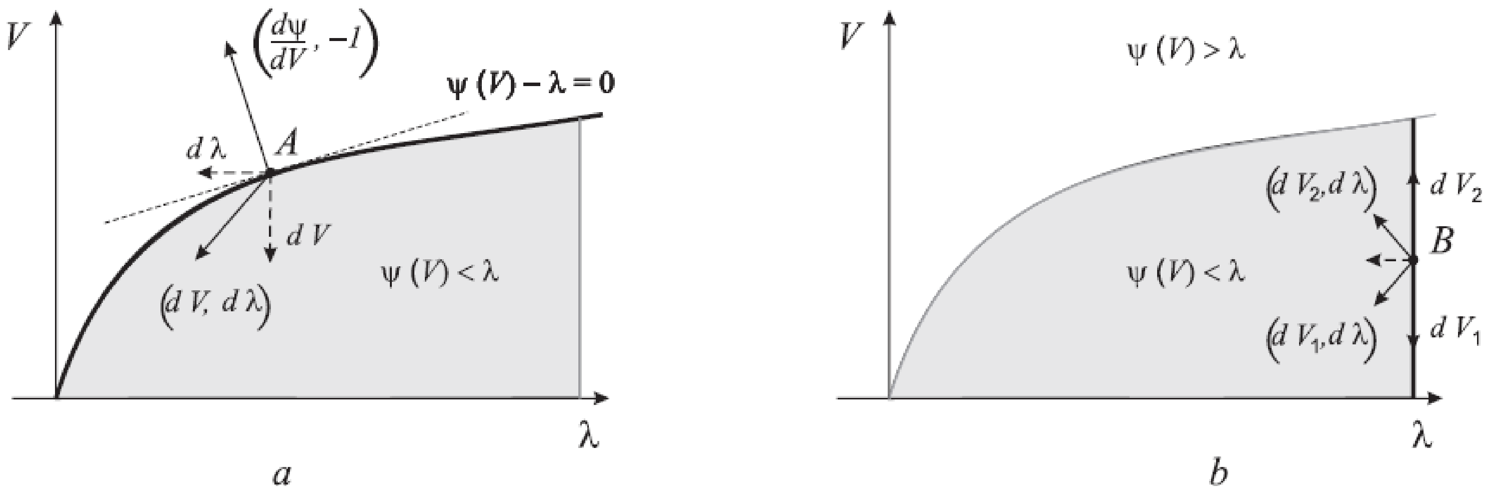


Figure 2: Illustration of the proof of Lemma 1. Panel a: boundary $\lambda = \psi(V)$. The vector $(\partial\psi/\partial V, -1)$ is a normal vector to the curve $\lambda = \psi(V)$ at the point A. Because $\partial\psi/\partial V \geq 0$ it is always pointing in the direction of $\lambda < \psi(V)$. Panel b: boundary $\lambda = \text{const}$. Because $\dot{\lambda} \leq 0$ for all $\lambda \geq 0$ the vector $(dV, d\lambda)$ is pointing in the direction of $\lambda \geq \psi(V)$.

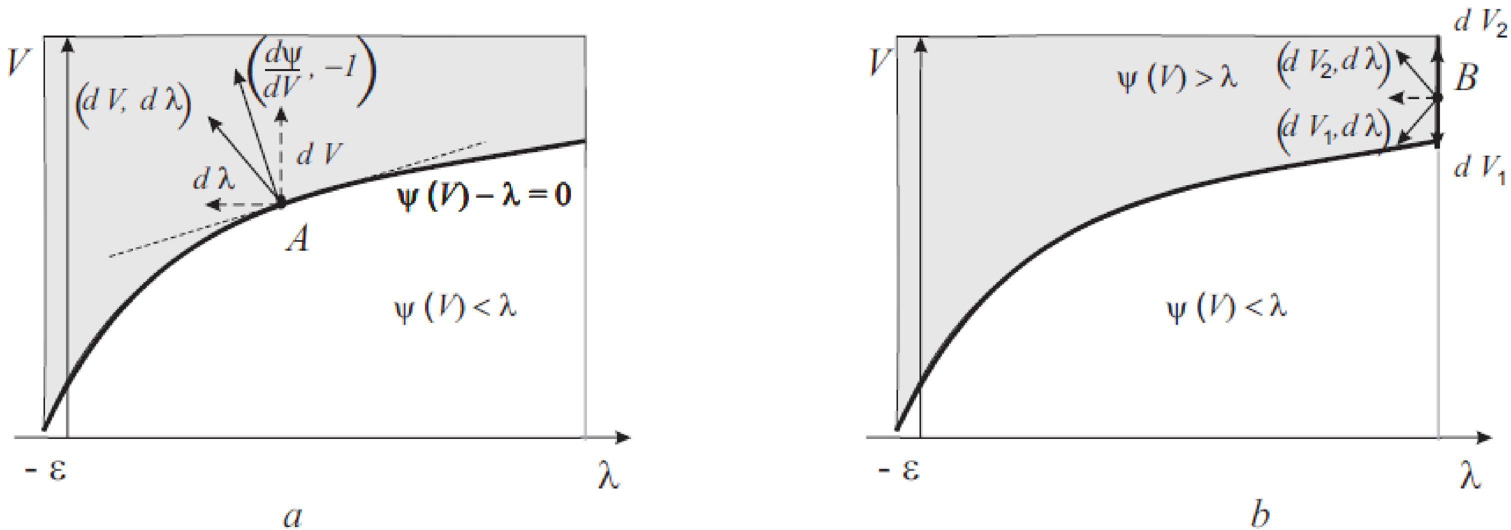
$$\dot{V} = \frac{\partial V}{\partial x} f(x, \lambda, t)$$
$$\dot{\lambda} = g(x, \lambda, t)$$

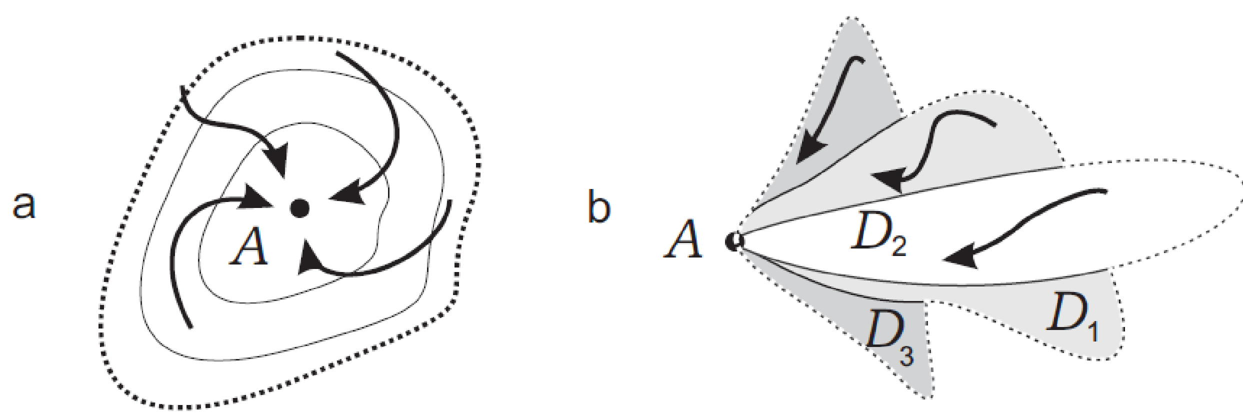
$$\left.\frac{\partial \psi}{\partial V}\dot{V} - \dot{\lambda}\right|_{\lambda=\psi(V)} \leq 0, \ \forall\, V \in [0, a],$$

$$\leq \left.\frac{\partial \psi}{\partial V}\left[\alpha(V(x)) + \beta(V(x))\varphi(|\lambda|)\right] + \delta(\|x\|) + \xi(|\lambda|)\right|_{\lambda=\psi(V(x))}$$

3. A minimal problem and first results

The same approach can be used to specify domains from with the trajectories necessarily escape…

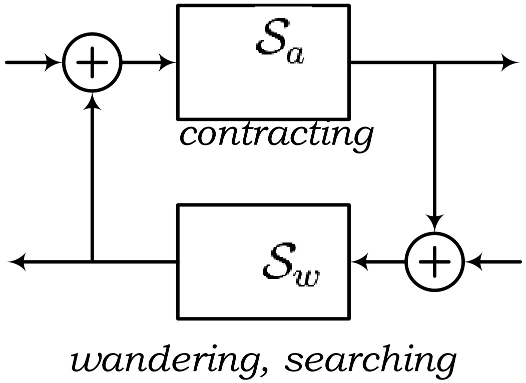


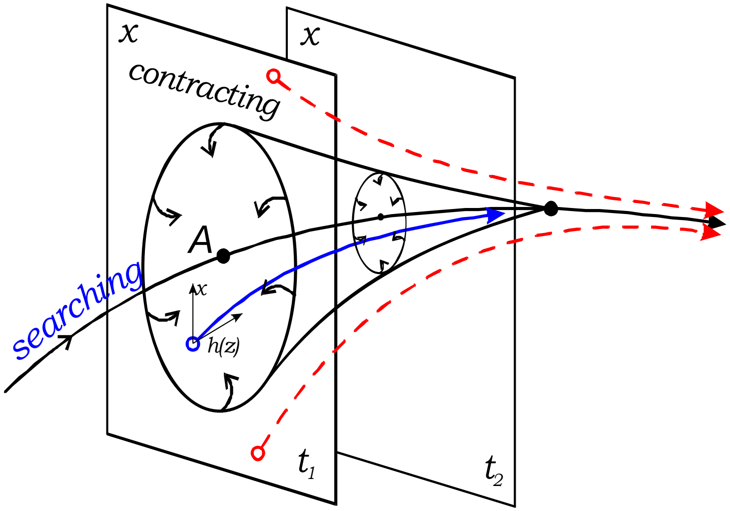


3. A minimal problem and first results

From: Tyukin, Steur, Nijmeijer, and van Leeuwen (SIAM Journal on Control and Optimization)

If the model is NOT a system of ODE's …





Contracting : $\mathcal{S}_a: \quad \|\mathbf{x}(t)\|_{\mathcal{A}} \leq \beta(\|\mathbf{x}(t_0)\|_{\mathcal{A}}, t - t_0) + c\|u_a(t)\|_{\infty,[t_0,t]}$

"Searching" :

$$\mathcal{S}_w: \int_{t_0}^{t} \gamma_1(u_w(\tau))d\tau \leq h(\mathbf{z}(t_0)) - h(\mathbf{z}(t)) \leq \int_{t_0}^{t} \gamma_0(u_w(\tau))d\tau$$

$$\gamma_0(a \cdot b) \leq \gamma_{0,1}(a) \cdot \gamma_{0,2}(b)$$

Interconnected as :

$$\int_{t_0}^{t} \gamma_1(\|\mathbf{x}(\tau)\|_{\mathcal{A}})d\tau \leq h(\mathbf{z}(t_0)) - h(\mathbf{z}(t)) \leq \int_{t_0}^{t} \gamma_0(\|\mathbf{x}(\tau)\|_{\mathcal{A}})d\tau$$

3. A minimal problem and first results

Separable contracting dynamics

$$\|\mathbf{x}(t)\|_{\mathcal{A}} \leq \|\mathbf{x}(t_0)\|_{\mathcal{A}} \cdot \beta_t(t - t_0) + c \cdot \|h(\mathbf{z}(\tau, \mathbf{z}_0))\|_{\infty,[t_0,t]}$$

With Lipschitz nonlinearity in the searching part

$$|\gamma_0(s)| \leq D_{\gamma,0} \cdot |s|$$

$$\int_{t_0}^{t} \gamma_1(\|\mathbf{x}(\tau)\|_{\mathcal{A}})d\tau \leq h(\mathbf{z}(t_0)) - h(\mathbf{z}(t)) \leq \int_{t_0}^{t} \gamma_0(\|\mathbf{x}(\tau)\|_{\mathcal{A}})d\tau,$$

Lemma 2 *(Non-uniform Small-Gain) There is a trapping region if the following holds*

$$\boxed{D_{\gamma,0} \cdot c \cdot \mathcal{G} < 1,}$$

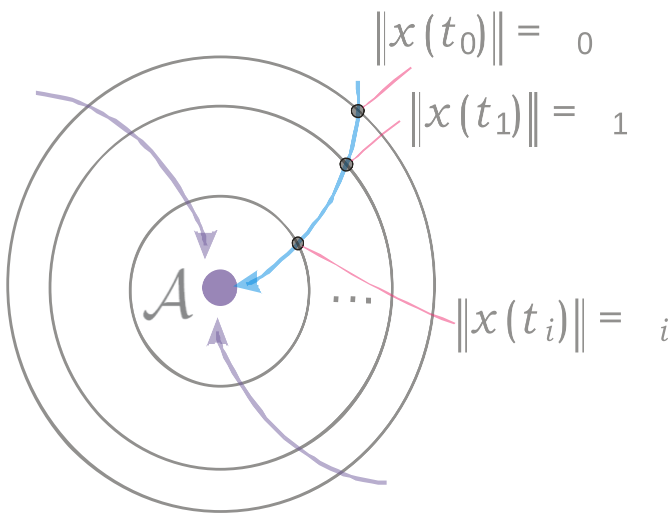
with

$$\mathcal{G} = \beta_t^{-1}\left(\frac{d}{\kappa}\right)\frac{k}{k-1}\left(\beta_t(0)\left(1 + \frac{\kappa}{1-d}\right) + 1\right)$$

for some $d \in (0,1),\ \kappa \in (1,\infty)$

Standard approaches

1) domains of attraction – neighborhoods
2) for autonomous systems implies stability



Given: sequence of time instances t_i

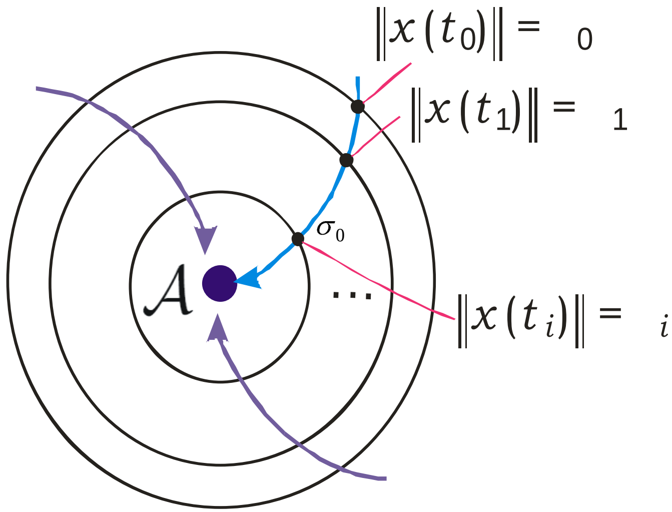
Prove: sequence of distances $_i$ does not increase (i.e. converges)

Mathematical framework

- contraction mapping theorems
- method of Lyapunov functions
- small-gain theorems

Standard approaches

1) domains of attraction – neighborhoods
2) for autonomous systems implies stability



Given: sequence of time instances t_i

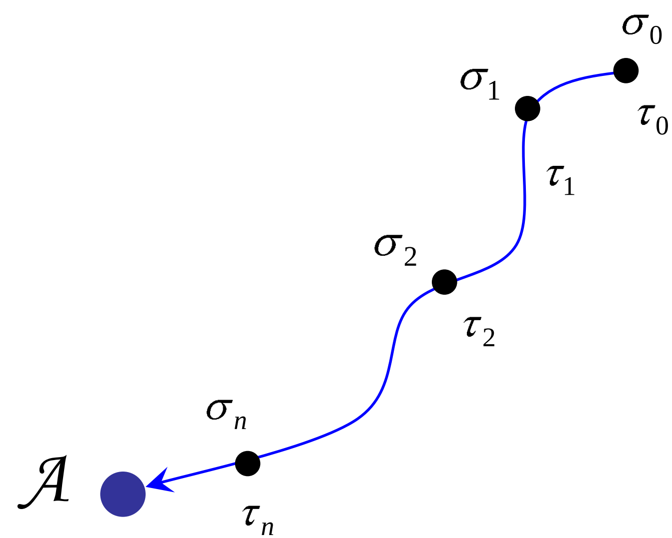
Prove: sequence of distances σ_i does not increase (i.e. converges)

Mathematical framework

- contraction mapping theorems
- method of Lyapunov functions
- small-gain theorems

Possible unstable convergence

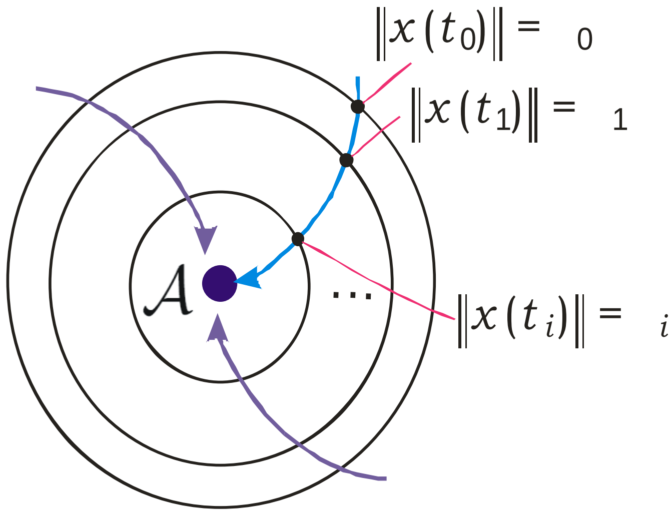
An asymptotically convergent trajectory that does not reach the target set in finite time …

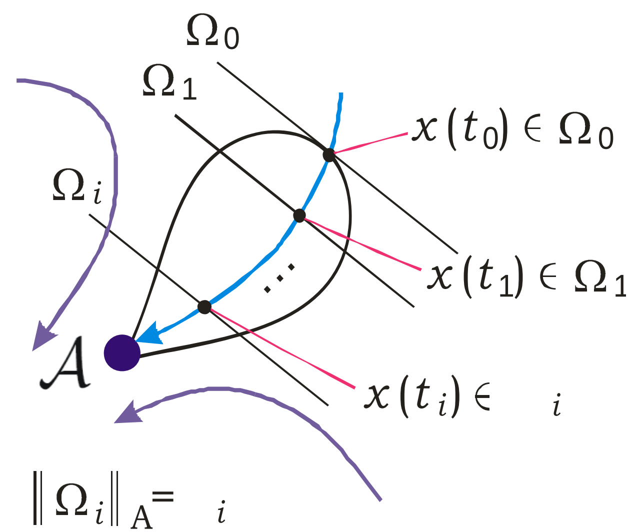


1) $\lim_{n\to\infty} \sigma_n = 0$

2) $\sum_n \tau_n = \infty$

Standard approaches	Proposed
	Weak attracting sets, concept of Milnor attracting sets
1) domains of attraction – neighborhoods 2) for autonomous systems implies stability	1) domains of attraction – sets of positive measure 2) possible to analyze unstable systems





Given: sequence of time instances t_i	**Prove:** sequence of partial sums t_i diverges
Prove: sequence of distances $_i$ does not increase (i.e. converges)	**Given:** sequence of distances $_i$

Mathematical framework	**Mathematical framework**
• contraction mapping theorems • method of Lyapunov functions • small-gain theorems	• Non-uniform small-gain theorem

4. Model. Design principles

1. There is a transitive low-dimensional invariant set (maximal attractor)

2. This attractor can broken into the smaller ones by “external perturbations”

3. Slight perturbation leads to that no other attractors emerge, but there are ghost attracting sets

4. Basins of attraction of these ghost attracting sets do not have common points with that of the resting state

4. Model. Diagrams, Equations and Parameters

Higher-dimensional, contracting

$$\dot{p}_i = -\tau_p (p_i - q_i)$$

$$\dot{q}_i = -\tau_q \left[q_i - g(p_i, w_i, k_i) \left(\sum_{j=1}^{N} c_j \varphi(y - \theta_j) + \varphi(u_i - \theta_{u_i}) \right) \right],$$

$$\tau_p, \tau_q \in R_{>0}, \quad \theta_i, \theta_{u_i} \in [0,1], \quad w_i \in R, \quad k_i \in R_{\geq 0},$$

$$g(p_i, w_i) = 1 + w_i \tanh(k_i p_i); \quad \varphi(z) = e^{-z^2}$$

"Echo"

Low-dimensional, exploring

$$\dot{x} = z(x + y - x(x^2 + y^2))$$

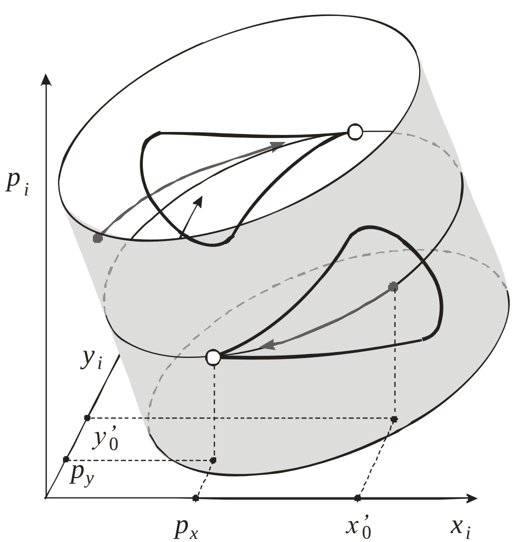
$$\dot{y} = z(x - y - y(x^2 + y^2)),$$

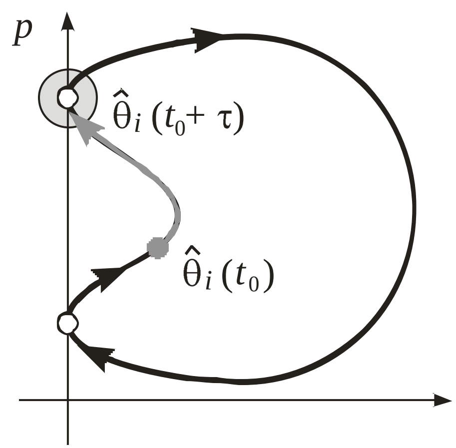
$$z = \gamma \left[\left(\sum_i 1 - f(q_i)\, \varphi(y - \theta_i) \right) + \delta \right], \quad \delta \in R_{\geq 0}$$

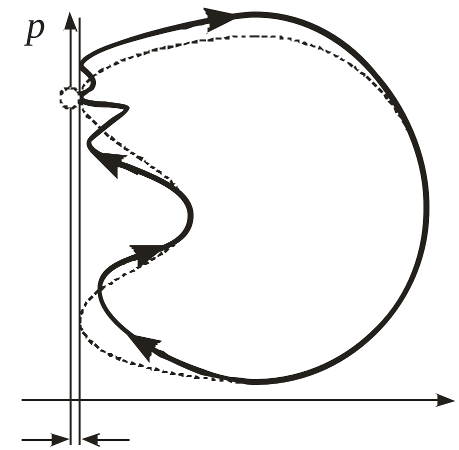
Important Parameters of the model

"Relaxation" constants

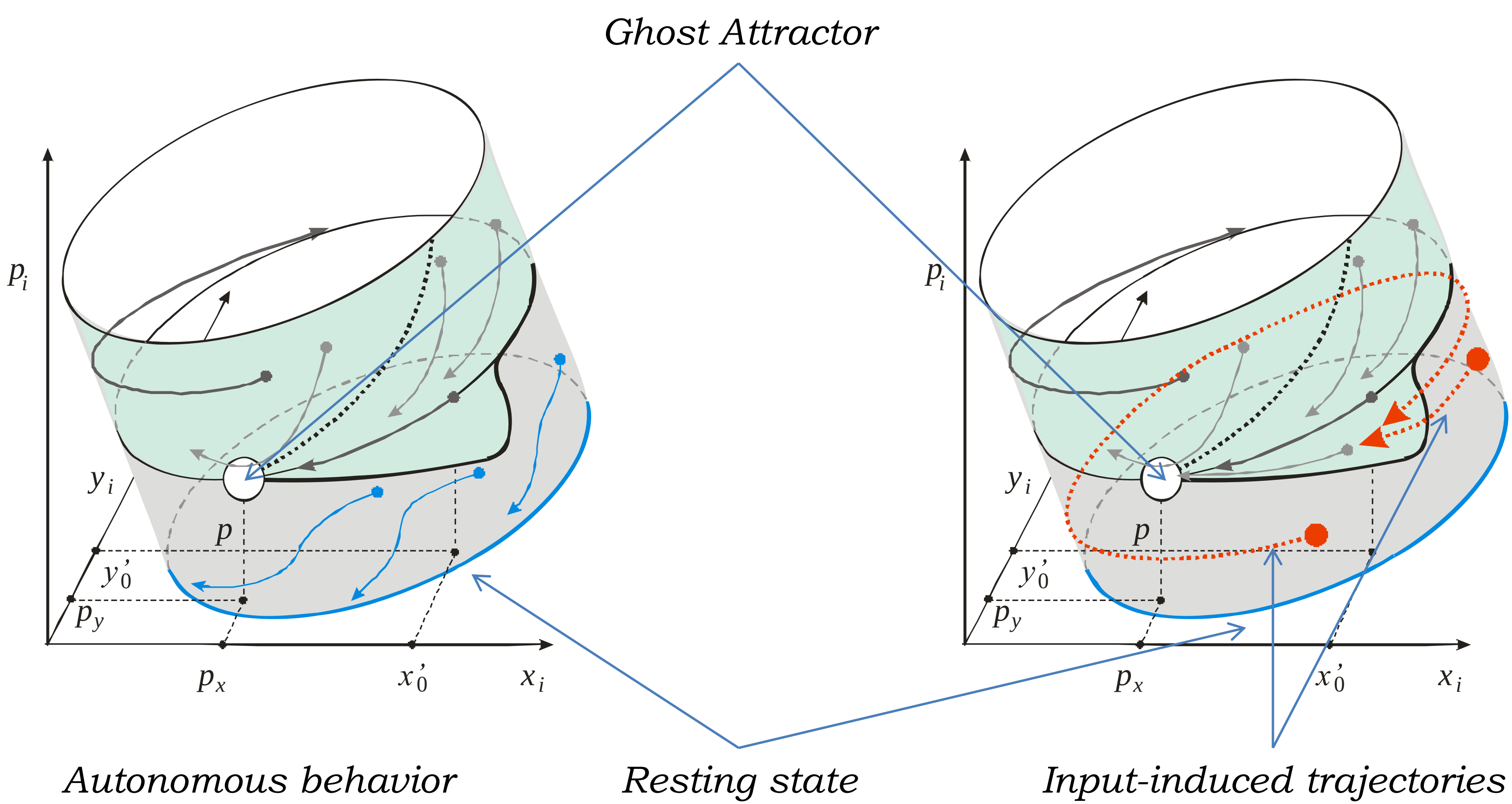
Maximal time of full "scan": T_s







4. Model. Input-induced memory diagrams



4. Model. Input-induced memory diagrams

4. Model

$$\dot{p}_i = -\tau_p(p_i - q_i)$$

$$\dot{q}_i = -\tau_q\left[q_i - g(p_i, w_i, k_i)\left(\sum_{j=1}^{N} c_j \varphi(y-\theta_j) + \varphi(u_i - \theta_{u_i})\right)\right],$$

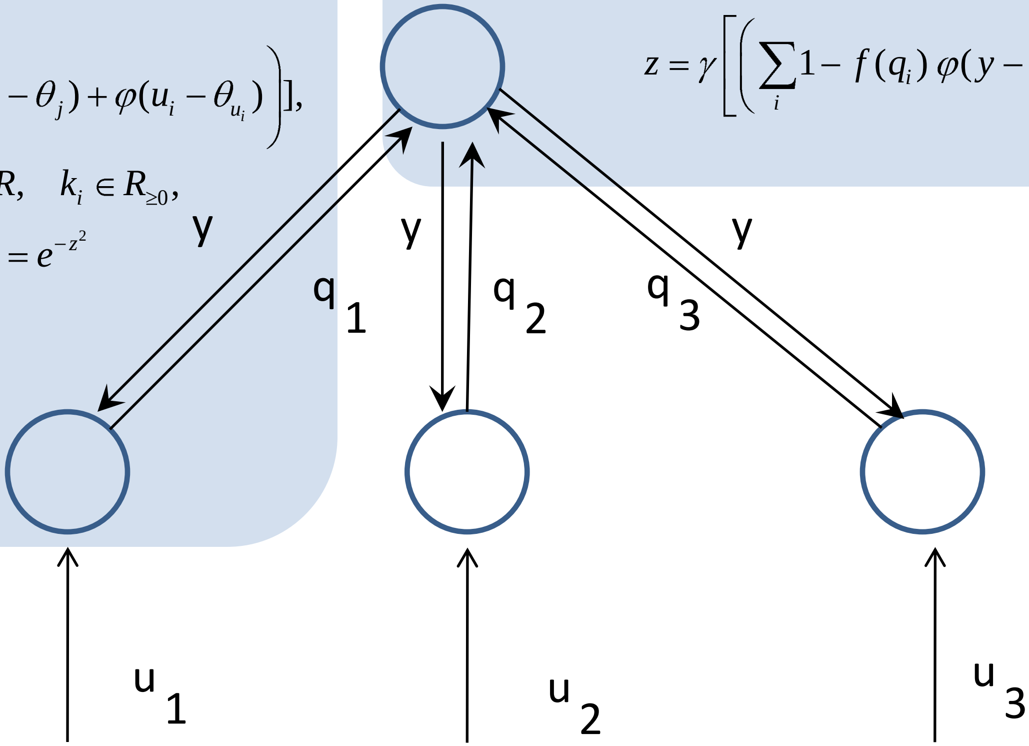
$$\tau_p, \tau_q \in R_{>0}, \quad \theta_i, \theta_{u_i} \in [0,1], \quad w_i \in R, \quad k_i \in R_{\geq 0},$$

$$g(p_i, w_i) = 1 + w_i \tanh(k_i p_i); \quad \varphi(z) = e^{-z^2}$$

$$\dot{x} = z(x + y - x(x^2 + y^2))$$

$$\dot{y} = z(x - y - y(x^2 + y^2)),$$

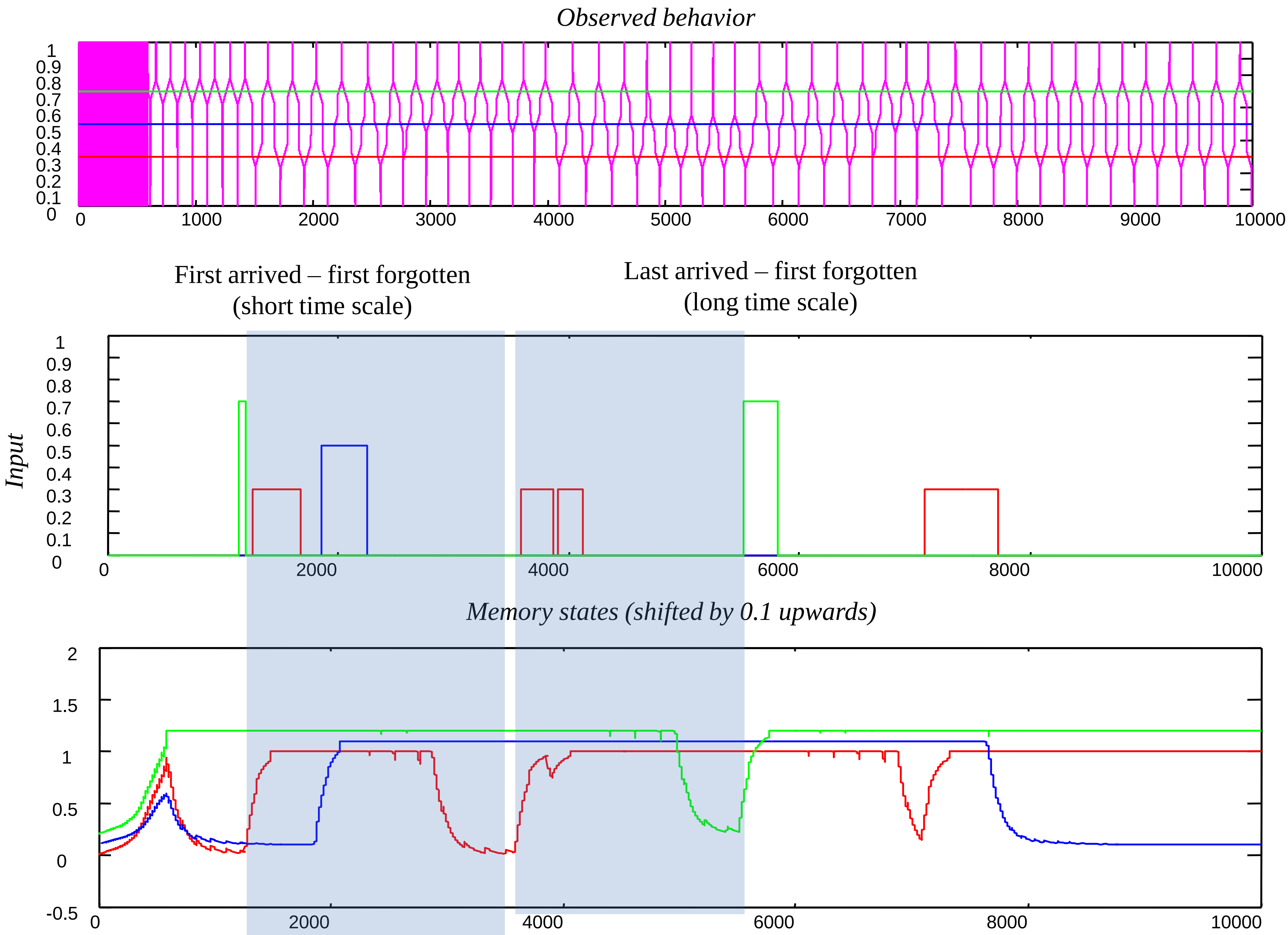
$$z = \gamma\left[\left(\sum_i 1 - f(q_i)\,\varphi(y - \theta_i)\right) + \delta\right], \quad \delta \in R_{\geq 0}$$



<u>What predictions can we make ? (depending on the parameters)</u>

- existence of memory
- illusions
- limitations of active (supra-threshold) memory

4. Model. Properties (3-node system)



4. Model. Properties

1. There are “active” and “non-active” ghost memory states
2. The number of “active” memory states is dependent on the strength of coupling
3. Recently induced memory state may “erase” ghosts that were induced earlier

Conjecture . *The model is computationally universal, i.e. it reproduces programs which a Turing machine with finite number of states can produce over a finite number of steps.*

<u>Idea of the proof:</u>

Model → Asynchronous Hopfield Networks → Universal computations

5. Discussion

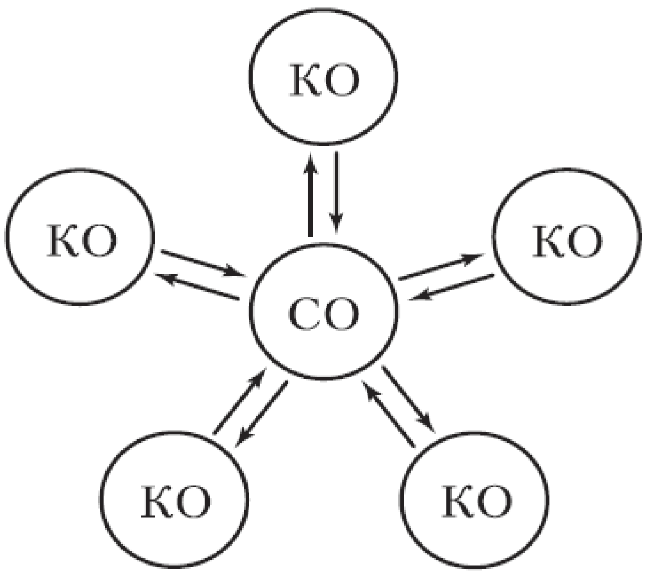
Extension to tori (point-ghosts → orbit-ghosts) leads to

"Neurolocator" (Kryukov, 2006)

Oscillatory memory (R. Borisyuk and Y. Kazanovich)

1999, 2003, 2006, and Biological Cybernetics, 2009

Models of memory and attention



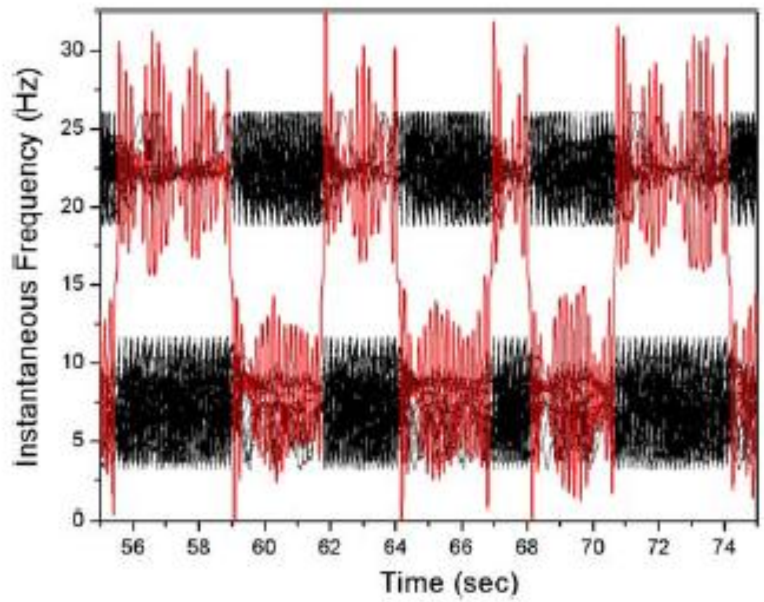


Fig. 2 Instantaneous frequency of the CO and POs as a function of time. There are two groups of POs and the natural frequencies of POS of one group are distributed in the interval [5, 10] and of another group in the interval [20, 25]. The frequency of CO "jumps" from the frequency range of one group of POs to the range of another, temporally synchronizing most oscillators in the "selected" group

6. Conclusion

1. Mathematical modelling of the brain is considered from the view point of dynamical (and controlled) systems. High-dimensional, evolving on (to) a transitive set, and “reducing” its complexity in response to stimulation

2. A novel concept of computation with ghost attractors has been presented. Our formal definition of ghost attractors is constructive. The concept unifies earlier frameworks (computing with attractors, Hirsh, or trajectories, Tsuda) and offers a resolution to the debate about which framework is better suited as a model of brain computations

3. A mathematical formalism is developed to study emergence of weak attractors in a class of systems described as an interplay between contacting higher-dimensional an exploring low-dimensional components

4. We presented a simple model realizing these features. Surprisingly, the model has certain computational universality (as the Hopfield nets do) and is operationally similar to more biologically plausible models such as neurolocator