Dynamics of Neural Systems
Local analysis of nonlinear systems II and Lyapunov functions

Martin A. Giese

Martin.giese@uni-tuebingen.de

Jan 13, 2025

Hertie-Institut
für klinische Hirnforschung

CIN
Werner Reichardt
Centre for Integrative Neuroscience

Overview

- Topological equivalence
- Lyapunov functions

Overview

- Topological equivalence
- Lyapunov functions

Topological equivalence

- Two subsets A and B of metric spaces are called **topologically equivalent** when there exists a continuous and continuously invertible mapping (homeomorphism) $h\colon A \to B$ between them.
- A **differentiable manifold** is a manifold that is locally similar to the Euclidean space and differentiable (detailed definition see Perko).

Center manifold theorem I

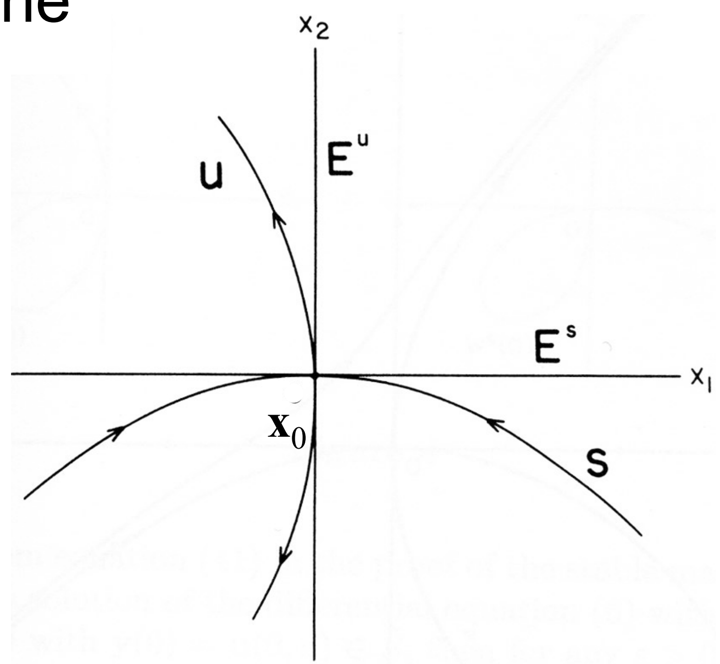
- Assume: $\mathbf{f} \in C^1$ has a fixed point $\mathbf{x}_0$, where the matrix $\mathbf{A} = \dfrac{\partial \mathbf{f}(\mathbf{x}_0)}{\partial \mathbf{x}}$ has k eigenvalues with negative and m with positive real parts, and $n - m - k$ with zero real part; then:
- there exists a k-dimensional **stable manifold** S that is tangent to the stable subspace E^s of the linearized system, which is invariant under the flow of the DEQ, i.e. $\boldsymbol{\varphi}_t(\mathbf{c}) \in S$ for $\mathbf{c} \in S$ and $t \geq 0$, and with: $\boxed{\lim_{t\to\infty} \boldsymbol{\varphi}_t(\mathbf{c}) = \mathbf{x}_0}$
- there exists an m-dimensional **unstable manifold** U that is tangent to the unstable subspace E^u, which is invariant under the flow ($\boldsymbol{\varphi}_t(\mathbf{c}) \in U$ for $\mathbf{c} \in U$ and $t \leq 0$), and with: $\boxed{\lim_{t\to-\infty} \boldsymbol{\varphi}_t(\mathbf{c}) = \mathbf{x}_0}$

Center manifold theorem II

- there exists an $(n\text{-}m\text{-}k)$-dimensional **center manifold** C that is tangent to the center subspace E^c, which is invariant under the flow of the DEQ.

Center manifold theorem II

- there exists an $(n\text{-}m\text{-}k)$-dimensional **center manifold** C that is tangent to the center subspace E^c, which is invariant under the flow of the DEQ.
- Remark: these manifolds are valid locally; it can be shown that they can be extended globally by propagating their points along the flow (global stable and unstable manifold); see Perko book for details.
- Illustration (2D): stable and unstable manifold



Example I

- Assume a system with three coupled nonlinear DEQ:

$$\boxed{\begin{aligned} \frac{\mathrm{d}x_1}{\mathrm{d}t} &= -x_1(t) \\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= -x_2(t) + x_1^2(t) \\ \frac{\mathrm{d}x_3}{\mathrm{d}t} &= x_3(t) + x_1^2(t) \end{aligned}}$$

- The origin $\mathbf{x} = \mathbf{0}$ is the only fixed point.
- From the linearized system we compute: $\mathbf{A} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\mathbf{0}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- The stable space E^{s} is this the $x_1 - x_2$ plane, and the unstable space E^{u} the x_3 axis.
- Solving the first equation one obtains a second non-autonomous linear equations with $x_1^2(t)$ as input.

Example II

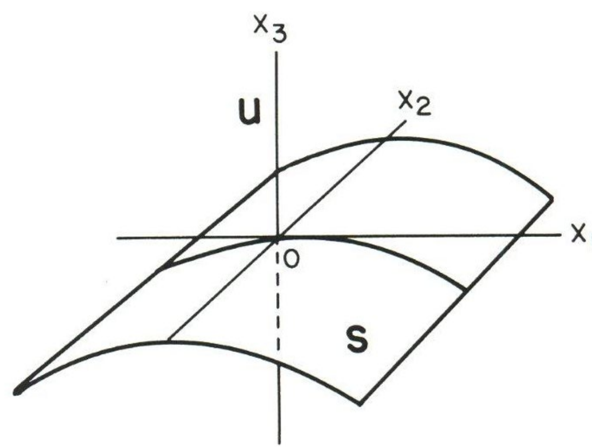
- The solution is: $\mathbf{x}(t) = \boldsymbol{\varphi}_t(\mathbf{x}_0) = \begin{pmatrix} x_{0,1}e^{-t} \\ x_{0,2}e^{-t} + x_{0,1}^2\left(e^{-t} - e^{-2t}\right) \\ x_{0,3}e^{t} + \dfrac{(x_{0,1})^2}{3}\left(e^{t} - e^{-2t}\right) \end{pmatrix}$
- This implies that $\lim_{t\to\infty} \boldsymbol{\varphi}_t(\mathbf{x}_0) = \mathbf{0}$ iff $x_{0,\mathbf{3}} + \dfrac{(x_{0,1})^2}{3} = 0$, implying:

$$S = \left\{\mathbf{x}_0 \in \mathrm{IR}^3 | x_{0,\mathbf{3}} = -\frac{(x_{0,1})^2}{3}\right\}$$ (stable manifold)

- Likewise, it follows that $\lim_{t\to-\infty} \boldsymbol{\varphi}_t(\mathbf{x}_0) = \mathbf{0}$ only iff $x_{0,1} = x_{0,2} = 0$ implying:

$$U = \left\{\mathbf{x}_0 \in \mathrm{IR}^3 | \, x_{0,1} = x_{0,2} = 0\right\}$$

(unstable manifold)



Hartman-Grobman theorem

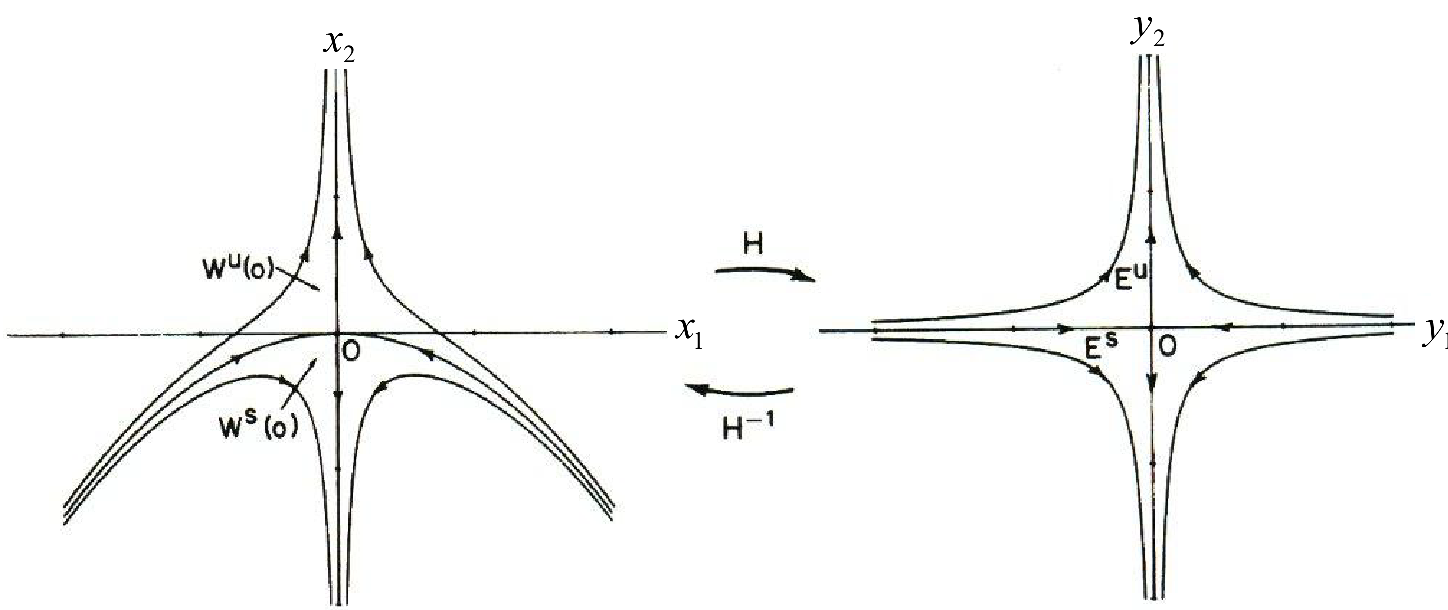
- Differential equations are called **topologically equivalent** if there exists a homeomorphism H (invertible map) that maps the trajectories onto each other.
- Assume: $\mathbf{f} \in C^1$ has a **hyperbolic** fixed point in zero (i.e. $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, and $\mathbf{A} = \dfrac{\partial \mathbf{f}(\mathbf{0})}{\partial \mathbf{x}}$ has no eigenvalues with zero real part) then the original system $\boxed{\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})}$ and the linearized system $\boxed{\dot{\mathbf{y}} = \mathbf{A}\mathbf{y}}$ are **topologically equivalent**; this means, for points $\mathbf{c}$ $(= \mathbf{x}_0)$ sufficiently near the origin, there exists a homeomorphism H so that:

$$\underbrace{H \circ \boldsymbol{\varphi}_t(\mathbf{c})}_{\mathbf{y}(t)} = e^{\mathbf{A}t} \underbrace{H(\mathbf{c})}_{\mathbf{y}_0} \qquad H\colon \mathbf{x} \to \mathbf{y}$$

Hartman-Grobman theorem

- Illustration:

Original system **Topologically equiv. linear system**

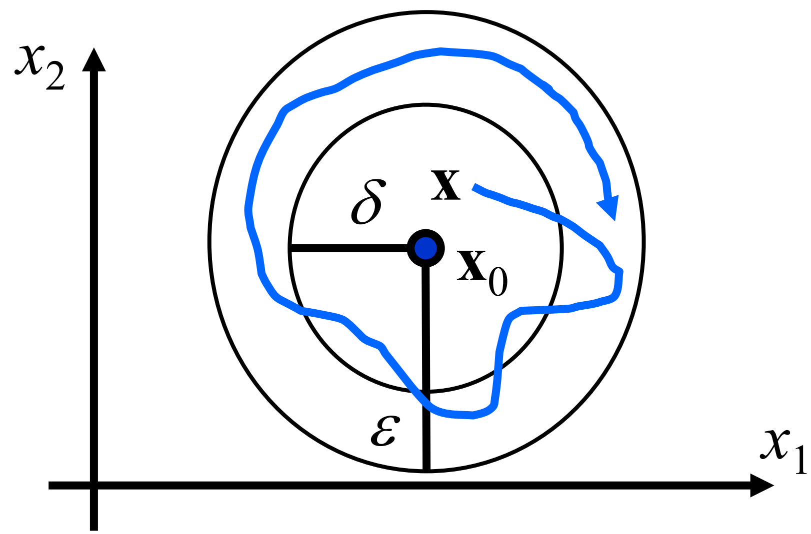


Overview

- Topological equivalence
- Lyapunov functions

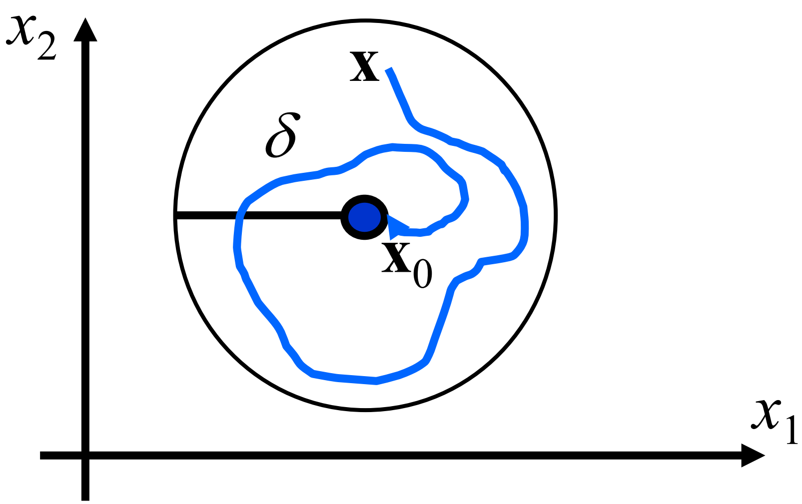
Stability of equilibrium points I

- Lyapunov theory helpful to characterize stability also non-hyperbolic fixed points.
- Assume $\boldsymbol{\varphi}_t$ is the flow of the differential equation: $\boxed{\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})}$
- An equilibrium point $\mathbf{x}_0$ is called **stable** if there exists for all $\varepsilon > 0$ a $\delta > 0$ such that for all $\mathbf{x} \in N_\delta(\mathbf{x}_0)$ and $t \geq 0$ is fulfilled: $\boldsymbol{\varphi}_t(\mathbf{x}) \in N_\varepsilon(\mathbf{x}_0)$ (where $N_\varepsilon(\mathbf{x})$ is a ball with radius ε about $\mathbf{x}$). $\rightarrow$ 'Solutions for $\mathbf{x}$ close $\mathbf{x}_0$ remain close.'
- An equilibrium is **unstable** if it is not stable.



Stability of equilibrium points I

- The equilibrium point $\mathbf{x}_0$ is **asymptotically stable** if there exists $\delta > 0$ such that for all $\mathbf{x} \in N_\delta(\mathbf{x}_0)$: $\lim_{t\to\infty} \boldsymbol{\varphi}_t(\mathbf{x}) = \mathbf{x}_0$.
- The set of all points for which this condition is fulfilled is called **domain / basin of attraction** of $\mathbf{x}_0$.
- Remark that these conditions require a certain behavior for trajectories passing the neighborhood of $\mathbf{x}_0$, not only for the ones passing through $\mathbf{x}_0$.



Stability of equilibrium points II

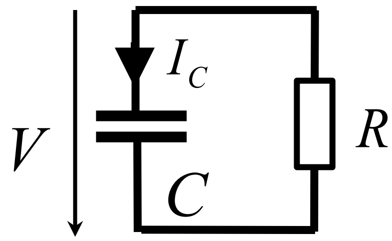
- Examples: Linear system:
  - stable node is asymptotically stable.
  - unstable node or a saddle point is unstable.
  - center is stable, but not asymptotically stable.
- A hyperbolic fixed point is obviously unstable if at least one eigenvalue has positive real part.

Lyapunov theory: motivation I

- Assume simple linear model of a membrane (RC circuit) with the DEQ:

$$I_C = C\dot{V}(t) = -V(t)/R \qquad \boxed{\dot{V}(t) = \frac{\mathrm{d}V}{\mathrm{d}t} = \frac{-V(t)}{RC}}$$

Passive membrane circuit



- Because this system does not contain energy sources and a resistor, eventually the voltage $V(t)$ will converge to zero.
- This implies $V = 0$ is an asymptotically stable fixed point of the dynamics.
- Remark: The electrical energy stored in the capacitor is

$$E(t) = \int_0^t V(t')I_C(t')\,\mathrm{d}t' = C\int_0^t V(t')\frac{\mathrm{d}V}{\mathrm{d}t'}\mathrm{d}t' = C\frac{V^2(t)}{2}$$

(where we assumed $V(0) = 0$).

Lyapunov theory: motivation II

- This energy decays along the trajectories of the system, unless V is already in the fixed point $V = 0$ because:

$$\dot{E}(t) = \frac{\mathrm{d}}{\mathrm{d}t} C \frac{V^2(t)}{2} = CV(t)\dot{V}(t) = -V^2(t)/R \le 0$$

- Geometrically, the changes of the system state define a descending motion on the ‘landscape’ that is defined by the function E.
- In general, a passive system that dissipates energy and does not contain energy sources will be stable because the state in which all energy has been dissipated will be stable.

State functions

- Electrical energy is an example of a **state function**, which depends on the dynamical state of the system.
- For a general DS $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, we require that state functions $E(\mathbf{x})$ are scalar and have continuous spatial derivatives (C^1 functions).
- A state function is called **positive definite** in a region R of the neighborhood of a singular point $\mathbf{x}_0$ if:
  $E(\mathbf{x}_0) = 0$ and
  $E(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{x}_0$
- In general, the time derivative of the state function is then given by: $\dot{E}(\mathbf{x}) = \left(\dfrac{\partial E}{\partial \mathbf{x}}\right)^T \mathbf{f}(\mathbf{x})$

Lyapunov theorem

A.M. Lyapunov

- Direct method of Lyapunov proven in his thesis (1892).
- Core idea: definition of a generalized energy function (Lyapunov function) $E(\mathbf{x})$ that decreases on the trajectories of the dynamics.
- **Theorem:** If in region R there exists a (C^1) positive definite state function (with $E(\mathbf{x}_0) = 0$ and $E(\mathbf{x}) > 0$ if $\mathbf{x} \neq \mathbf{x}_0$) then:
  a) if $\dot{E}(\mathbf{x}) \leq 0$ everywhere then $\mathbf{x}_0$ is stable.
  b) if $\dot{E}(\mathbf{x}) < 0$ everywhere (in region R) except for $\mathbf{x}_0$ then $\mathbf{x}_0$ is asymptotically stable.
  c) if $\dot{E}(\mathbf{x}) > 0$ everywhere (in region R) except for $\mathbf{x}_0$ then $\mathbf{x}_0$ is unstable.

Example

- Consider the system: $\dot{x}_1 = -x_1 - x_2 - 3x_1x_2^2$

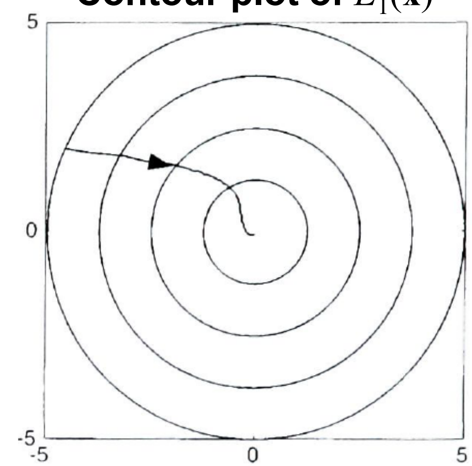
  $\dot{x}_2 = -x_2 + x_1$

- Lyapunov functions: $E_1(\mathbf{x}) = x_2^2 + x_1^2$

  (see Wilson book) $E_2(\mathbf{x}) = x_1^2 + x_2^2 - x_2^4$

- Around the singular point $\mathbf{x} = \mathbf{0}$ both functions are positive definite.
- E_1 proves stability in the whole phase space, E_2 only in the strip $|x_2| < 1$.
- Lyapunov functions are thus **not unique**; the proven stability region depends on the chosen function.

Contour plot of $E_1(\mathbf{x})$



Finding Lyapunov functions I

- No principled approach; often simple techniques work.
- Example: divisive feedback (see Lecture 7):

$$\boxed{\begin{aligned}\tau\frac{\mathrm{d}u_1}{\mathrm{d}t} &= -u_1(t) + \frac{s}{1+u_2(t)}\\ \tau\frac{\mathrm{d}u_2}{\mathrm{d}t} &= -u_2(t) + 2u_1(t)\end{aligned}}$$

- Lyapunov function can be found using following **theorem**: If the C^1 functions $F(x_1, x_2)$ and $G(x_1, x_2)$ have only a finite number of joint zeros (points with $F(x_1, x_2) = G(x_1, x_2) = 0$) then the following function with $|\varepsilon| < 1$

$$\boxed{E(\mathbf{x}) = (1/2)F^2(\mathbf{x}) + \varepsilon F(\mathbf{x})G(\mathbf{x}) + (1/2)G^2(\mathbf{x})} \qquad \boldsymbol{x} = [x_1, x_2]^{\mathrm{T}}$$

is **positive definite** in regions around each zero.

- **Corollary:** For $a,\ c < 0$ and $|b| < 2\sqrt{ac}$ the function $\boxed{aF^2(\mathbf{x}) + bF(\mathbf{x})G(\mathbf{x}) + cG^2(\mathbf{x})}$ is **negative definite** in these regions.

Finding Lyapunov functions II

- For the example choose $\varepsilon = 0$ and resulting Lyapunov function:

$$\begin{cases} F(u_1,u_2) = \frac{1}{\tau}\left(-u_1 + \dfrac{s}{1+u_2}\right) \\ G(u_1,u_2) = \frac{1}{\tau}(-u_2 + 2u_1) \end{cases}$$

$$E(u_1,u_2) = \frac{1}{2}(F^2 + G^2) = \frac{1}{2\tau}\left(\left(-u_1 + \frac{s}{1+u_2}\right)^2 + (-u_2 + 2u_1)^2\right)$$

Positive definite according to the theorem (with $\varepsilon = 0$)!

- From this follows:

$$\dot{E}(u_1,u_2) = F\left(\frac{\partial F}{\partial u_1}F + \frac{\partial F}{\partial u_2}G\right) + G\left(\frac{\partial G}{\partial u_1}F + \frac{\partial G}{\partial u_2}G\right)$$

Has the form assumed in the corollary!

$$= F^2\frac{\partial F}{\partial u_1} + FG\left(\frac{\partial F}{\partial u_2} + \frac{\partial G}{\partial u_1}\right) + G^2\frac{\partial G}{\partial u_2}$$

Finding Lyapunov functions II

- Taking into account $a = c = \frac{\partial F}{\partial u_1} = \frac{\partial G}{\partial u_2} = -\frac{1}{\tau}$ and using the corollary follows the negative definiteness of the last function if $|b| = \left|\frac{\partial F}{\partial u_2} + \frac{\partial G}{\partial u_1}\right| = \frac{1}{\tau}\left|\frac{-s}{(1+u_2)^2} + 2\right| < \frac{2}{\tau}$.

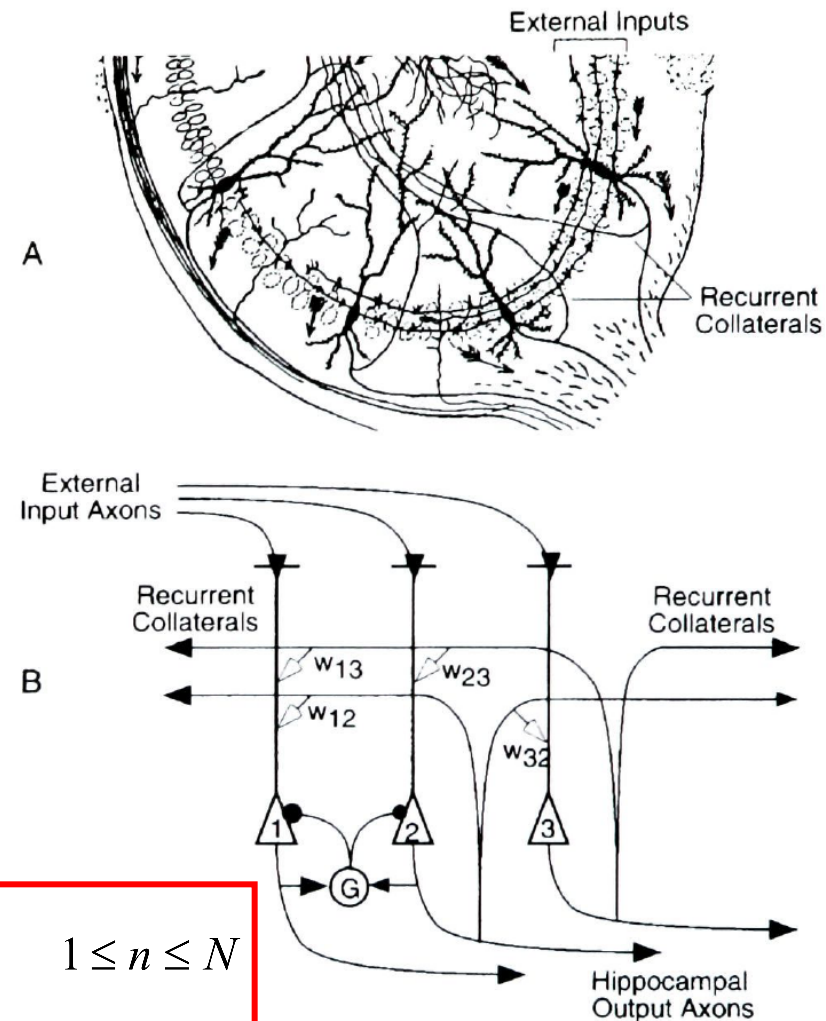
  This condition is fulfilled for $u_2 > \frac{\sqrt{s}}{2} - 1.$

- This implies that the fixed point $u_{0,2} = 2u_{0,1} = \frac{-1+\sqrt{1+8s}}{2}$ is within that region because $u_{0,2} > \frac{\sqrt{s}}{2} - 1$ for any $s \geq 0$. Therefore, E is really a Lyapunov function that is valid in the relevant region $u_{0,2} \geq 0$.

Memory networks I

- Important application of Lyapunov functions in neuroscience (Hopfield; Grossberg, …)
- Modeling long-term memory, e.g. in hippocampus.
- Model: Recurrent neural network with hebbian learning; example:

Hippocampal CA3 network



$$\tau \frac{\mathrm{d}u_n}{\mathrm{d}t} = -u_n(t) + \Theta\left(-v(t) + \sum_{m=1}^{N} w_{mn} u_m(t) \right) \qquad 1 \le n \le N$$

Inhibition dynamics:

$$\tau \frac{\mathrm{d}v}{\mathrm{d}t} = -v(t) + c \sum_{m=1}^{N} u_m(t) \quad \text{with} \quad \Theta(u) = K_1 \frac{u^2 1(u)}{K_2^2 + u^2}$$

Sigmoidal threshold

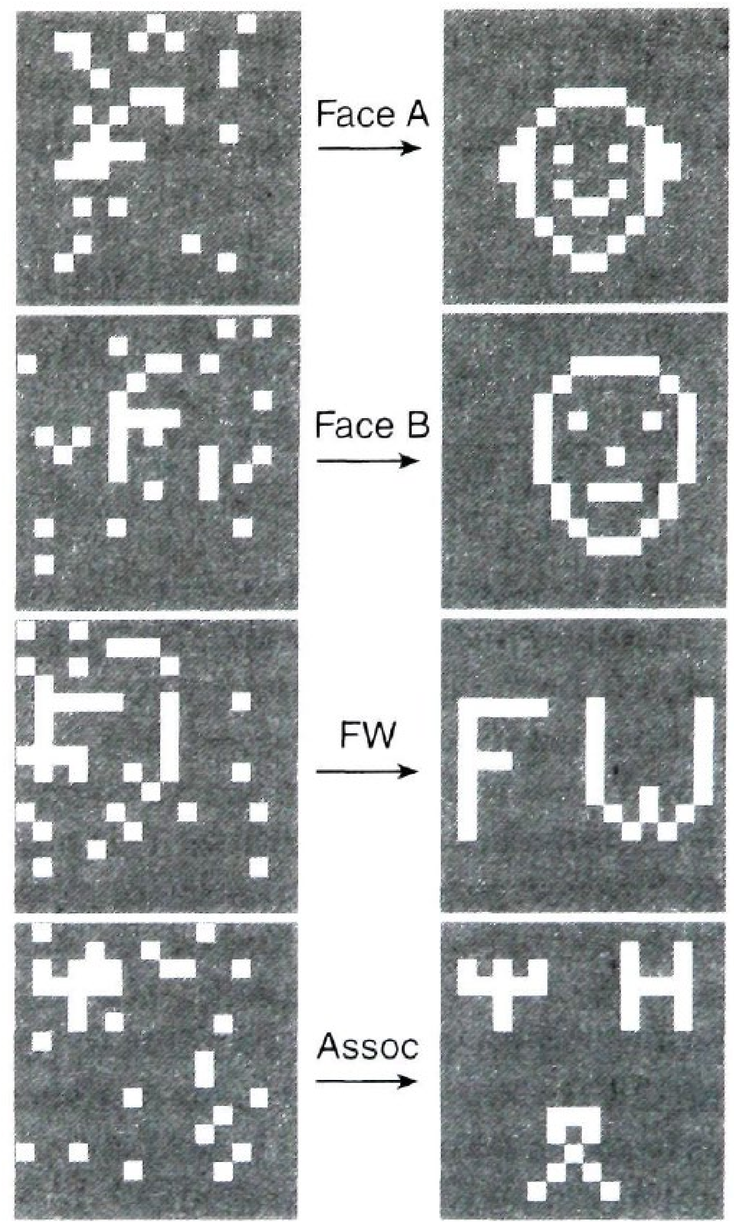
Memory networks II

- **Hebbian learning rule**:

$$\boxed{\tau_w \frac{\mathrm{d}w_{mn}}{\mathrm{d}t} = -w_{mn} + 1\left(u_m(t) - \frac{K_1}{2}\right)\left(u_{\boldsymbol{n}}(t) - \frac{K_1}{2}\right)}$$

- We assume that the learning dynamics is much slower than the activation dynamics; this allows us to treat the weights as quasi-constant.
- The network is an **auto-associative memory**: stored patterns can be completed by the recurrent dynamics.
- Suboptimal patterns suppressed by inhibition.

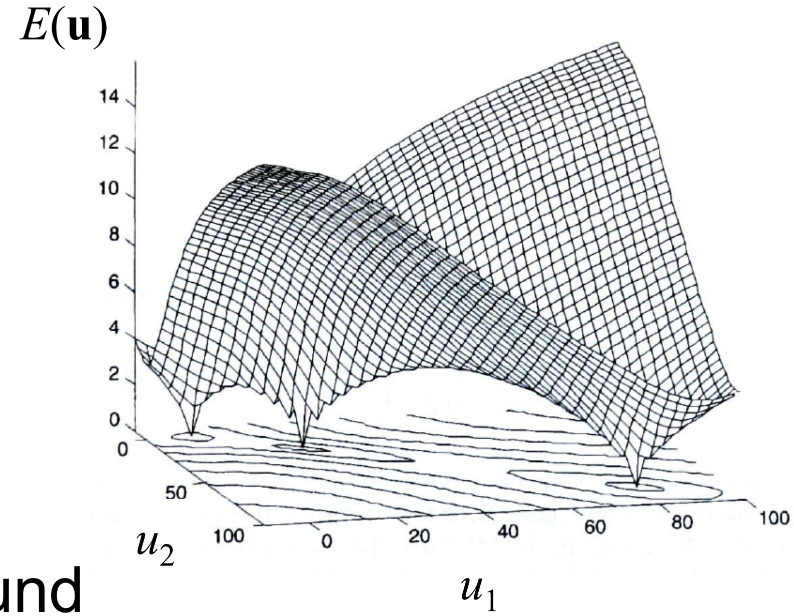
Pattern completion



Memory networks III

- Two neurons without inhibition:

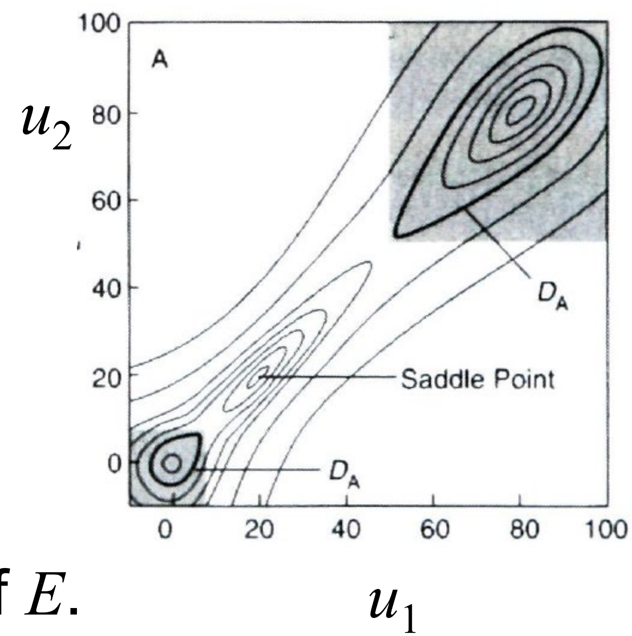
$$\boxed{\begin{aligned}\frac{\mathrm{d}u_1}{\mathrm{d}t} &= \frac{1}{\tau}\left(-u_1(t)+\Theta\left(u_2(t)\right)\right) = f_1(u_1,u_2)\\ \frac{\mathrm{d}u_2}{\mathrm{d}t} &= \frac{1}{\tau}\left(-u_2(t)+\Theta\left(u_1(t)\right)\right) = f_2(u_1,u_2)\end{aligned}}$$



- A Lyapunov function can be found using the last theorem:

$$\boxed{E(\mathbf{u}) = (1/2)(f_1^2(u_1,u_2)+f_2^2(u_1,u_2))}$$

- From $\dot{E}(\mathbf{u}) = -\frac{f_1^2}{\tau}-\frac{f_2^2}{\tau}+ f_1 f_2\left(\frac{\partial f_1}{\partial u_2}+\frac{\partial f_2}{\partial u_1}\right)$ follows $\dot{E}(\mathbf{u})<0$ for $\frac{\partial f_1}{\partial u_2}+\frac{\partial f_2}{\partial u_1}=\frac{\Theta'(u_2)+\Theta'(u_1)}{\tau}<\frac{2}{\tau}$. This is fulfilled in the gray regions.



- Stable states correspond to minima of E.

Things to remember

- Center manifold theorem $\rightarrow$ 2)
- Hartman-Grobman theorem $\rightarrow$ 2)
- Lyapunov theory $\rightarrow$ 2,3)

Literature (for this lecture)

1) Dayan, P. & Abbott, L.F. (2001 / 2005) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* MIT Press, Cambridge MA, USA. Chapter 7.

2) Perko, L. (1998) *Differential Equations and Dynamical Systems.* Springer-Verlag, Berlin. Chapter 2.

3) Wilson, H.R. (1999) *Spikes, Decisions, and Actions.* Oxford University Press, UK. Chapters 6 and 14.