Neural Dynamics For Embodied Cognition: Foundations

Gregor Schöner
Institute for Neural Computation (INI)
Ruhr-University Bochum

Survey

- Foundations 1: Neural dynamics [GS]
  - Introduction to Cedar/Instabilities in DFT [Stephan Sehring]
- Foundations 2: Dimensions/Binding [GS]
  - Cedar architecture: visual search [Raul Grieben]
- Foundations 3: Toward grounded cognition [GS]
  - Cedar architecture: relational grounding [Daniel Sabinasz]
- Foundations 4: Sequence generation [GS]
  - Cedar architecture sequence generation [Minseok Kang]

Survey

- Discussion [GS]
  - two forms of modularity
  - DFT vs. cognitive architectures
  - DFT vs. connectionism
  - DFT and neurosymbolics
  - DFT vs. VSA
  - DFT and embodiment, dynamical systems thinking
  - DFT vs mathematical psychology
  - Why model in DFT? How model in DFT?
  - Laws of the mind

Foundations 1: Neural dynamics

- Neuro-physics
- Neural dynamics
- Recurrent neural dynamics
- Neural fields: dynamics

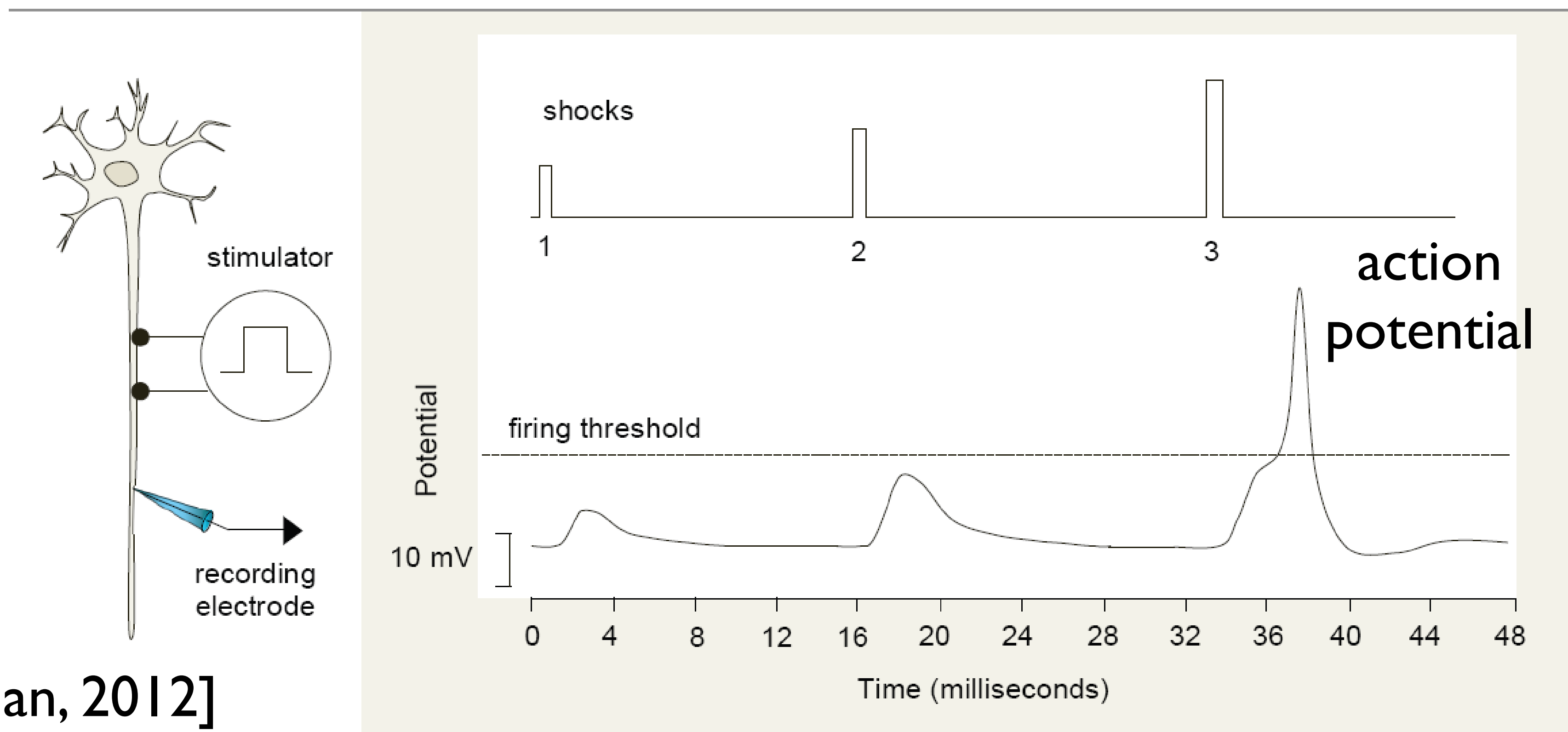
Neuro-physics

- membrane potential, $u(t)$, evolves as a dynamical system

$$\tau \dot{u}(t) = -u(t) + h + \text{input}(t)$$

$$\tau \approx 10 \text{ ms time scale}$$

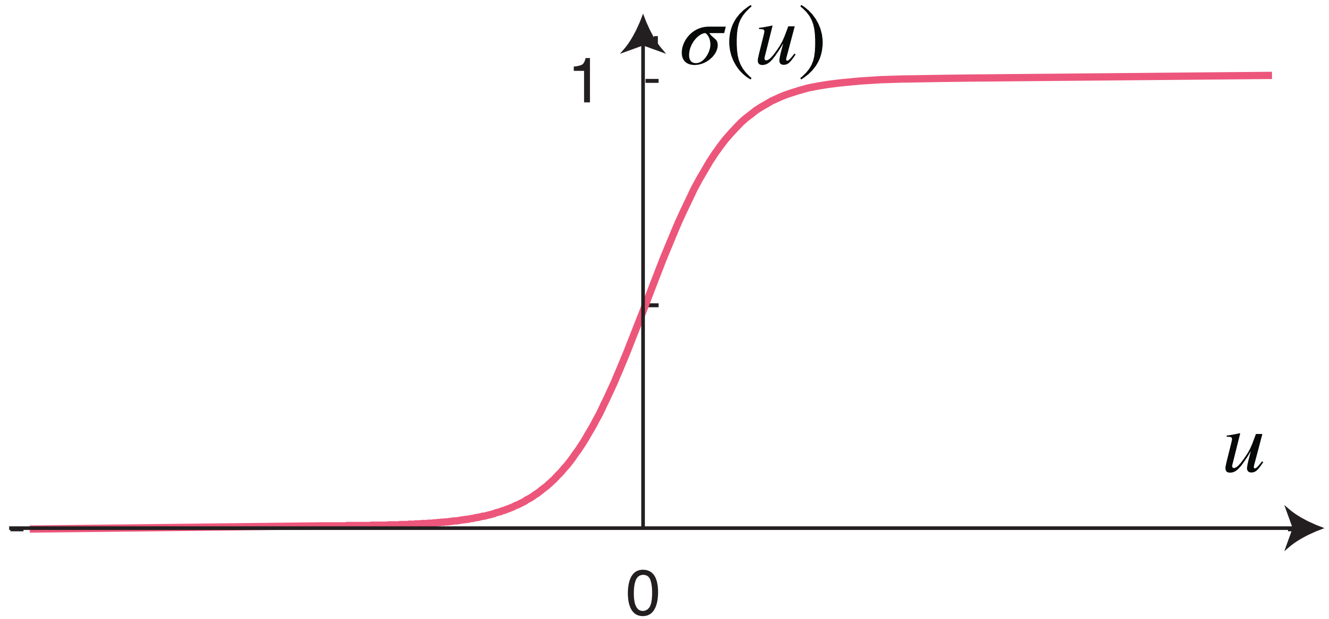
- only when membrane potential exceeds a threshold is activation transmitted to downstream neurons



[from: Tresilian, 2012]

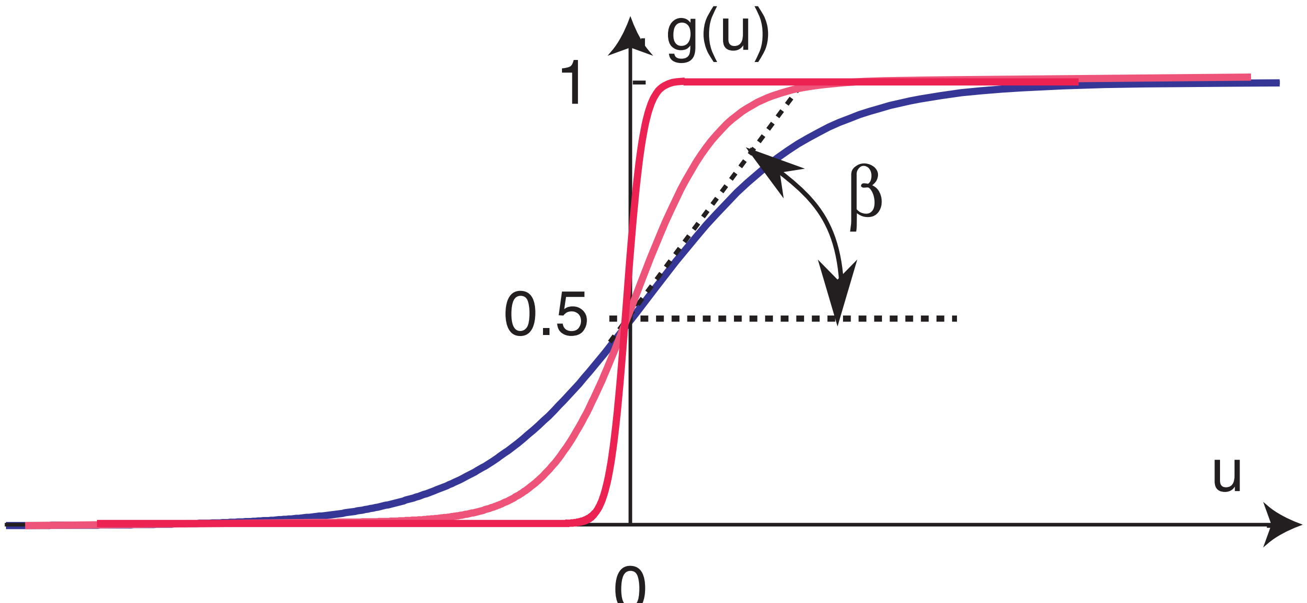
Neural dynamics

- spiking mechanism replaced by a threshold function
- that captures the effective transmission of spikes in populations



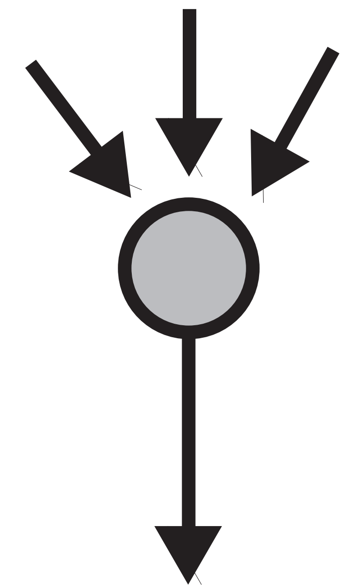
Neural dynamics

- replace spiking mechanism by sigmoid:
  - low levels of activation: not transmitted to downstream systems
  - high levels of activation: transmitted to downstream systems
- abstracting from biophysical details ~ population level membrane potential

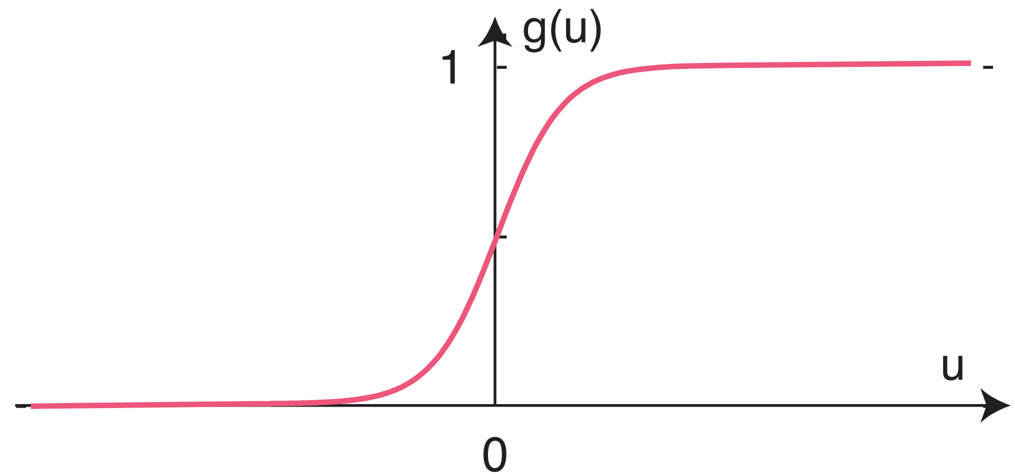


Connectionism

- employs the same abstraction: "neurons" sum input activations and pass them through a sigmoidal threshold function



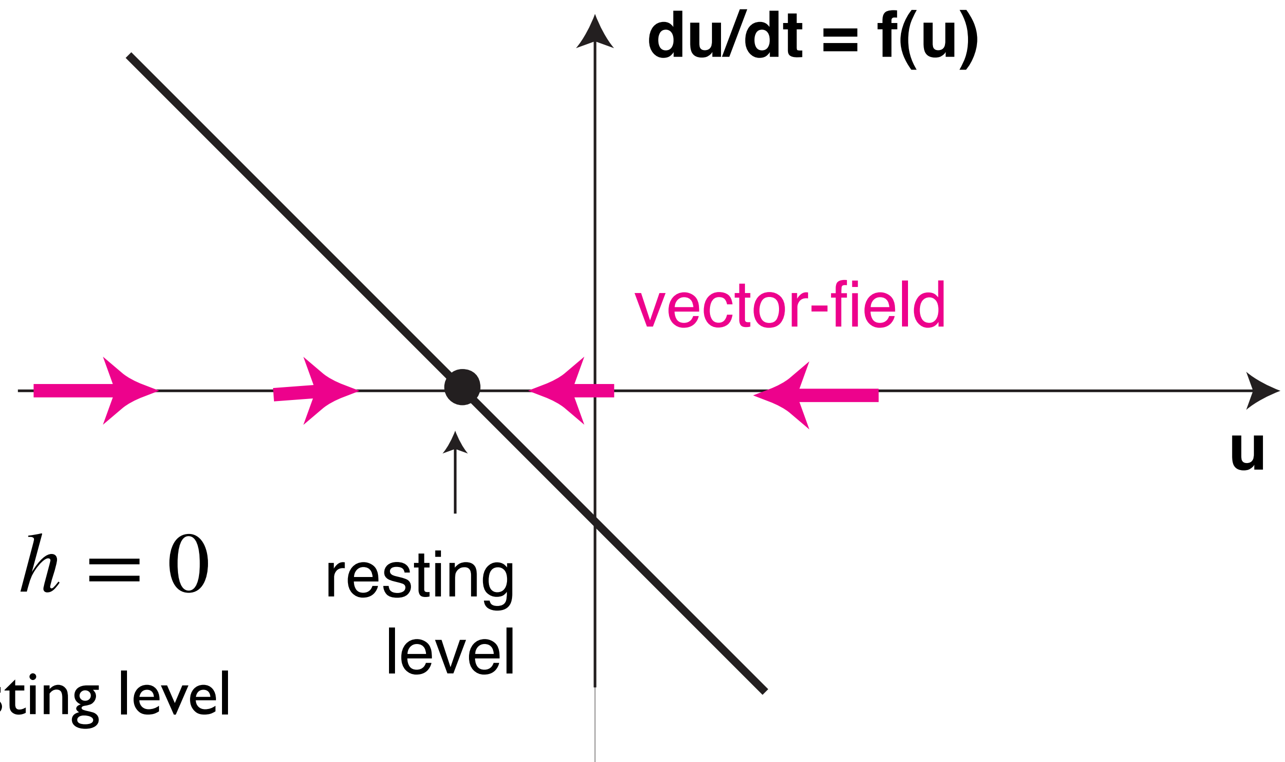
$$\text{output} = g\left(\sum(\text{inputs})\right)$$



Neural dynamics

- dynamical system: the present determines the future
- fixed point = constant solution = stationary state
- stable fixed point = attractor: nearby solutions converge to the fixed point

$$\tau\dot{u}(t) = -u(t) + h$$

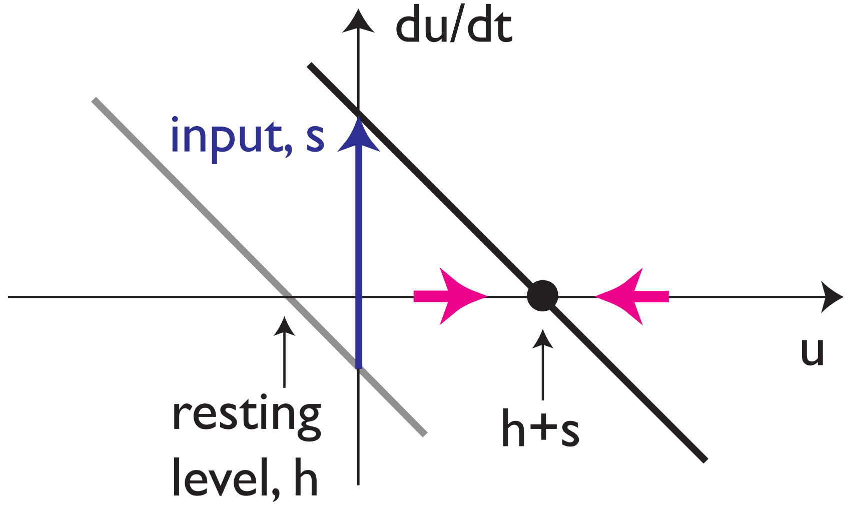


$$\tau\dot{u}_{\text{fp}} = -u_{\text{fp}} + h = 0$$

$$\Rightarrow u_{\text{fp}} = h$$ resting level

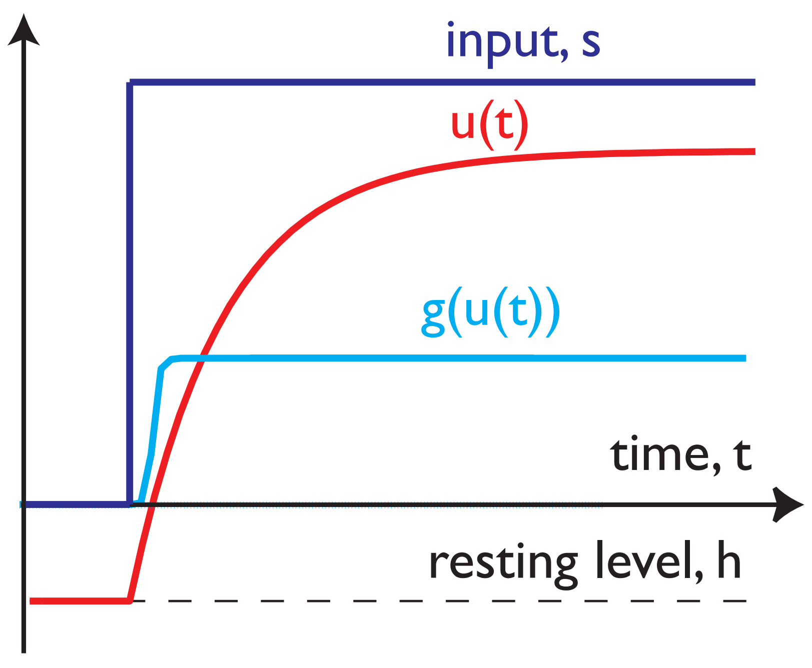
Neural dynamics

- inputs add to the rate of change of activation
  - positive: excitatory
  - negative: inhibitory



$$\tau \dot{u}(t) = -u(t) + h + s(t)$$

- input shifts the attractor
- activation tracks this shift
- $\sigma(u(t))$ transmitted to down-stream neurons

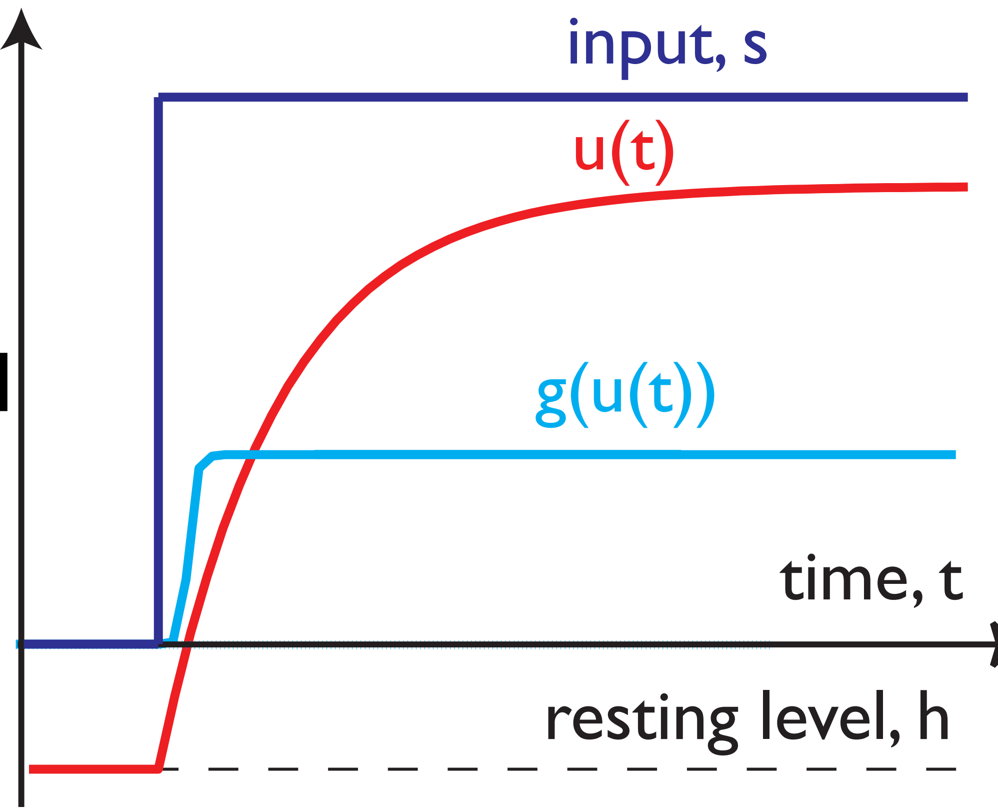


($\sigma(u)$ and $g(u)$ used interchangeably)

Neural dynamics

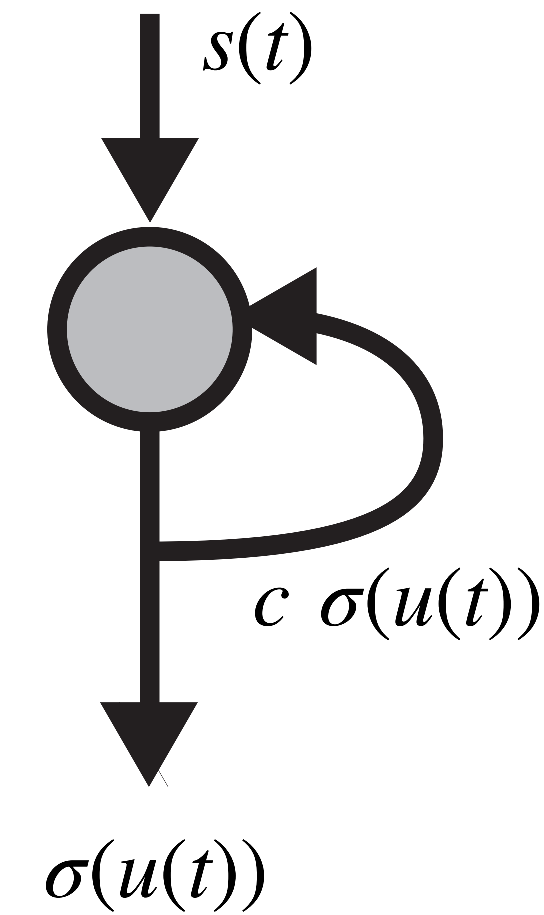
- so far, the dynamics just does low-pass filtering... (smoothing the time course)
- that would change as a step-function in a forward neural network
- when does neural dynamics make a real difference?

$$\text{output} = g\left(\sum(\text{inputs})\right)$$



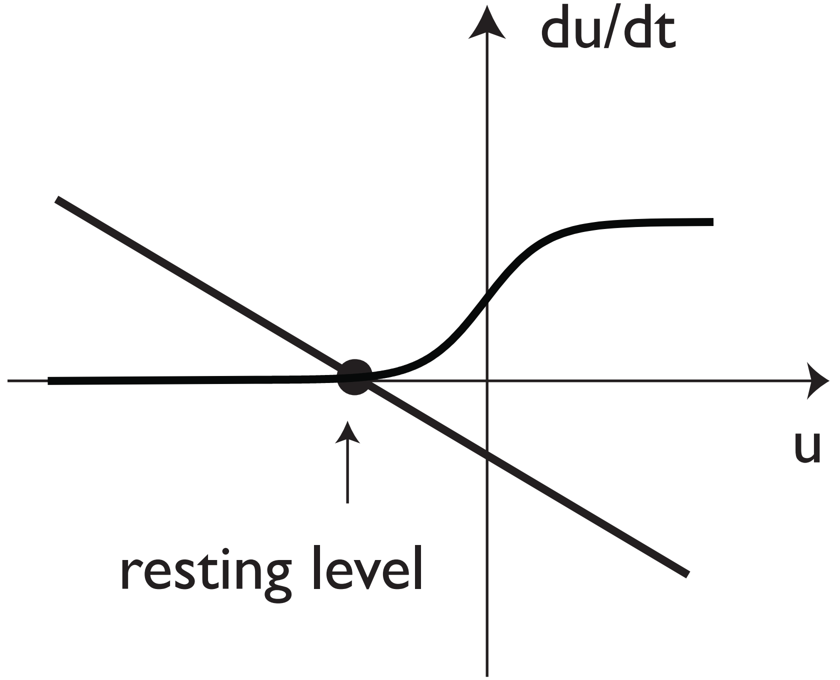
Neuronal dynamics with excitatory recurrent connection = interaction

- in recurrent networks, time is conceptually necessary as some inputs are outputs from the same neuron/population …
- “past outputs are new input”
- => dynamics

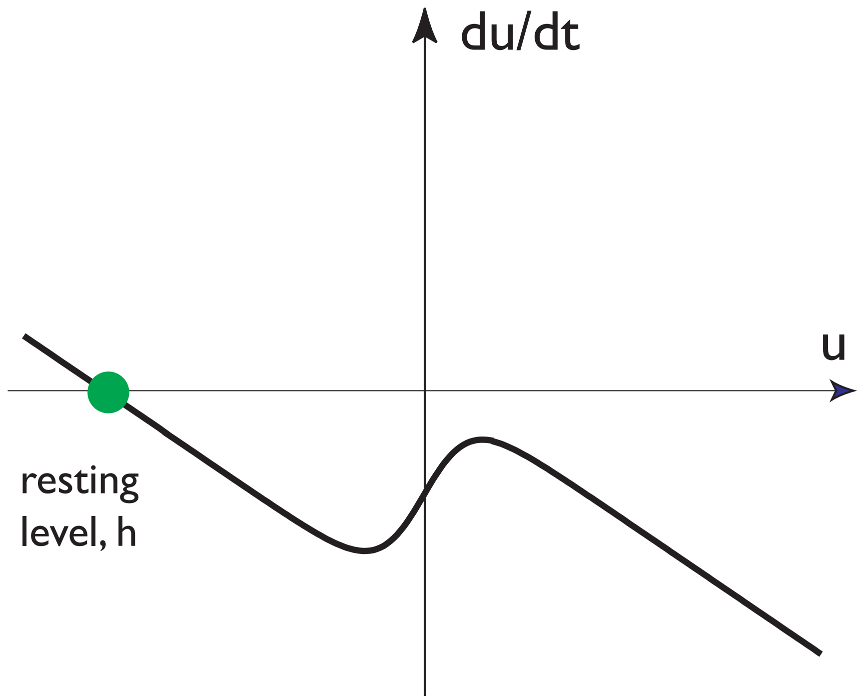


$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Neuronal dynamics with self-excitation



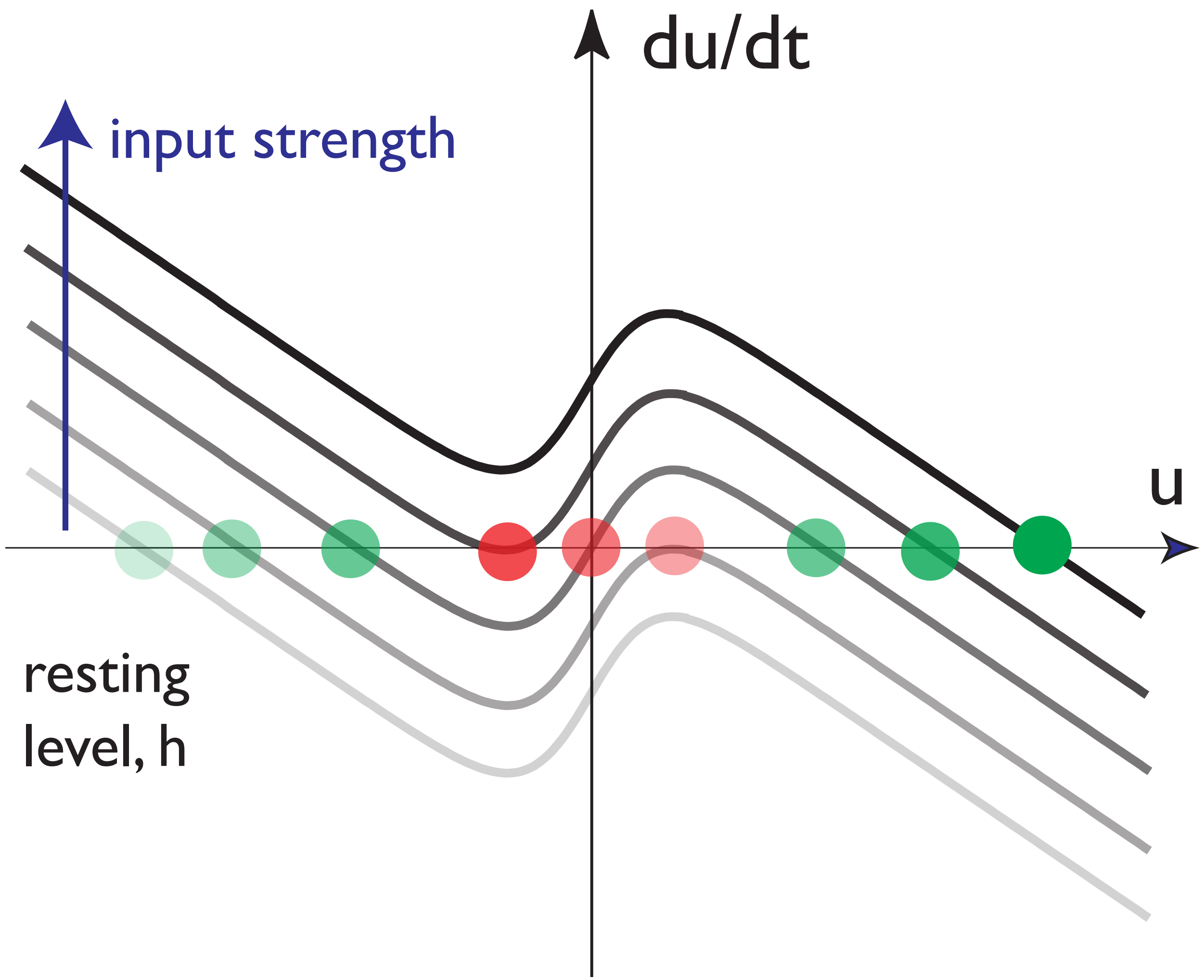
- nonlinear dynamics!



$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

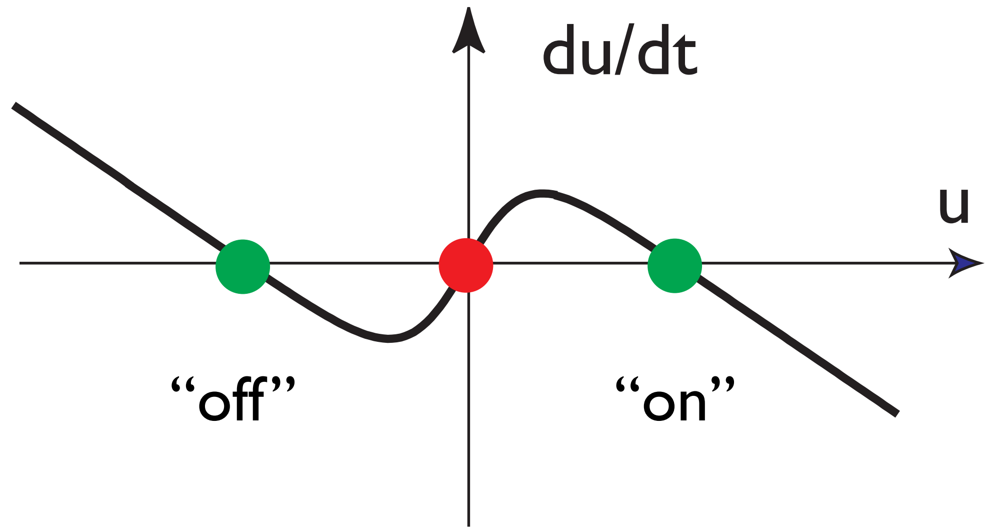
Neuronal dynamics with self-excitation

- varying input
- => number of attractors changes

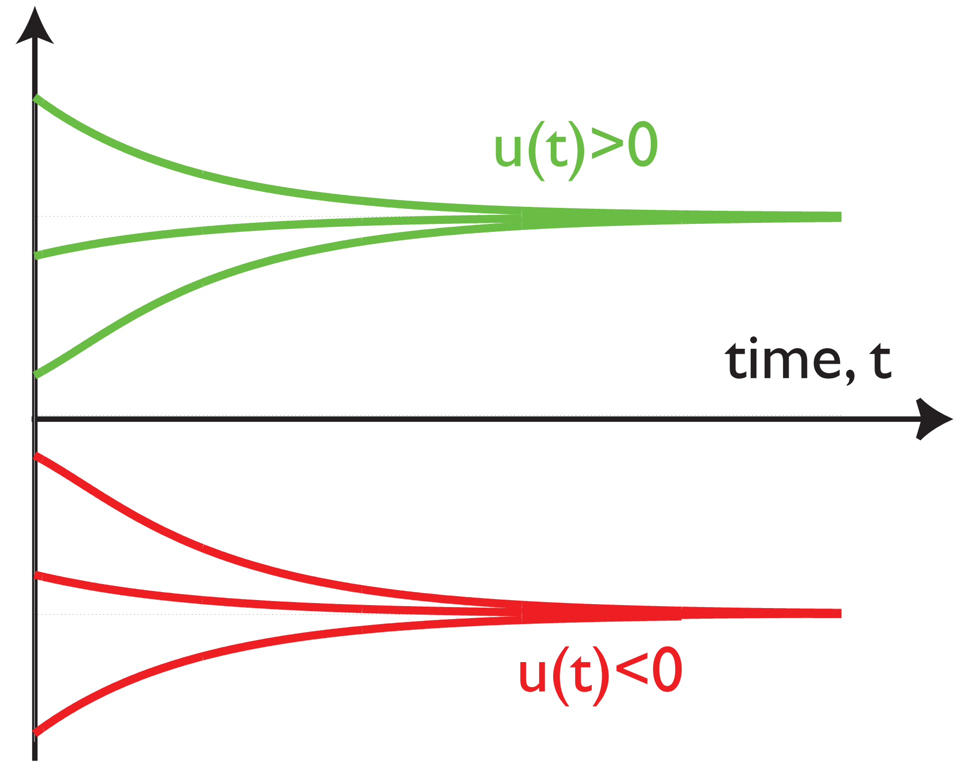


$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Neuronal dynamics with self-excitation



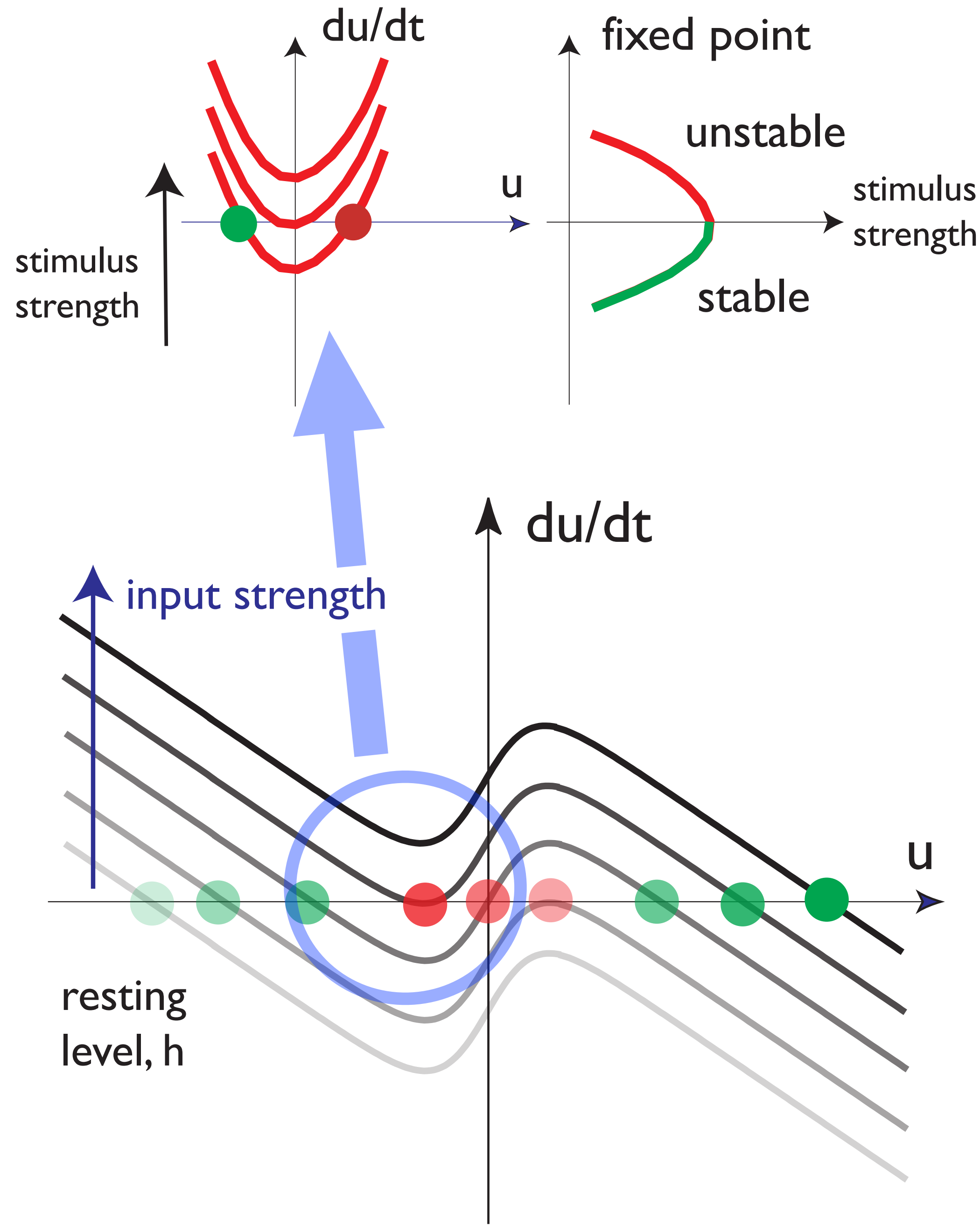
- at intermediate input levels: bistable dynamics
- "on" vs "off" state



$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Neuronal dynamics with self-excitation

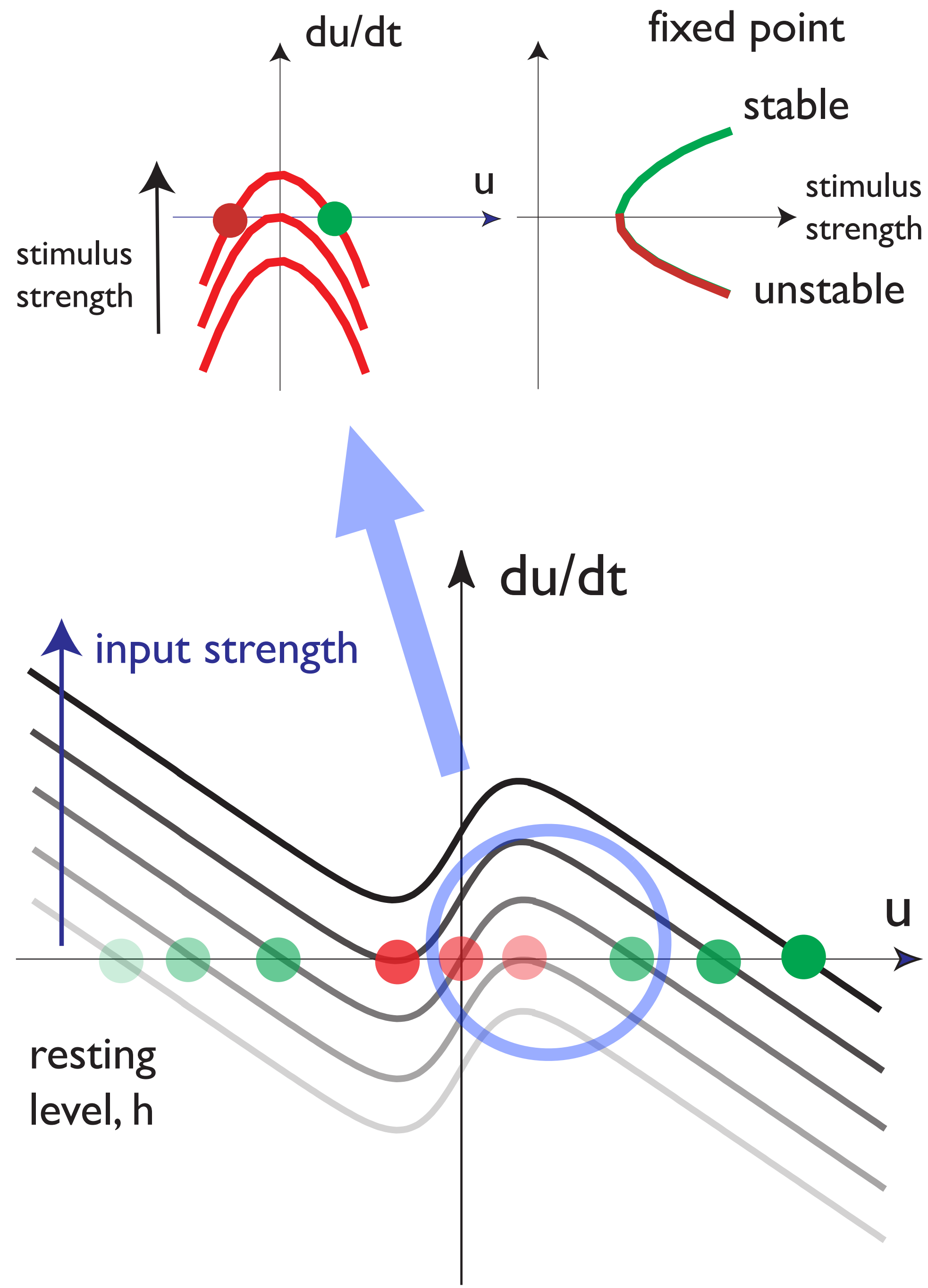
- increasing input strength => detection instability



$$\tau\dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Neuronal dynamics with self-excitation

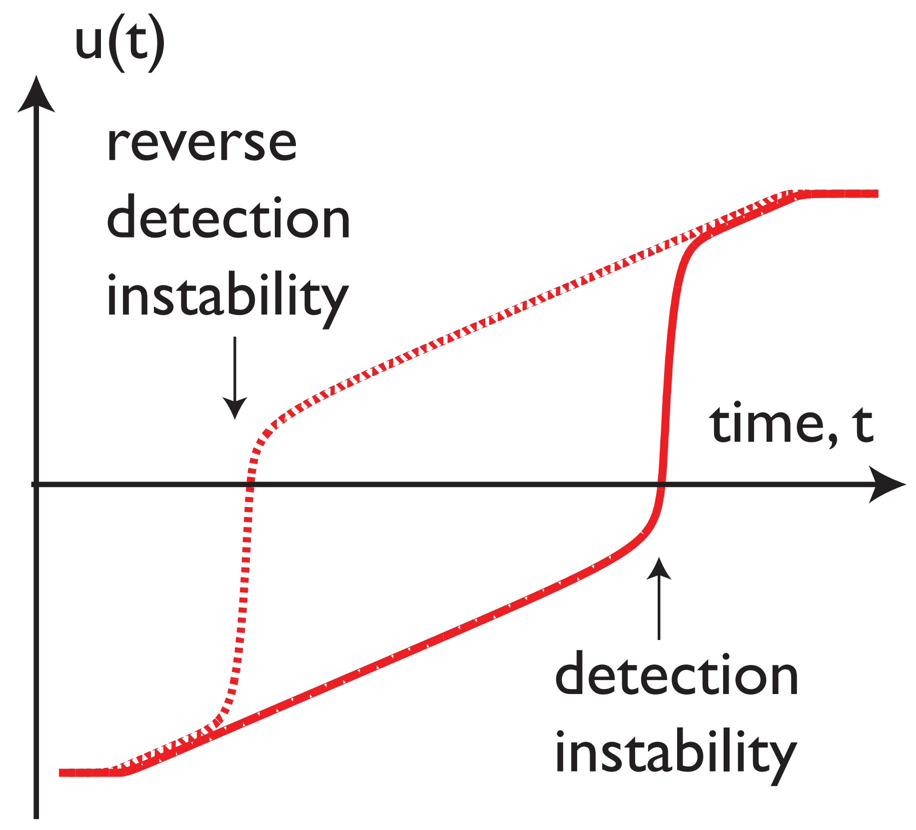
- decreasing input strength => reverse detection instability



$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

Neuronal dynamics with self-excitation

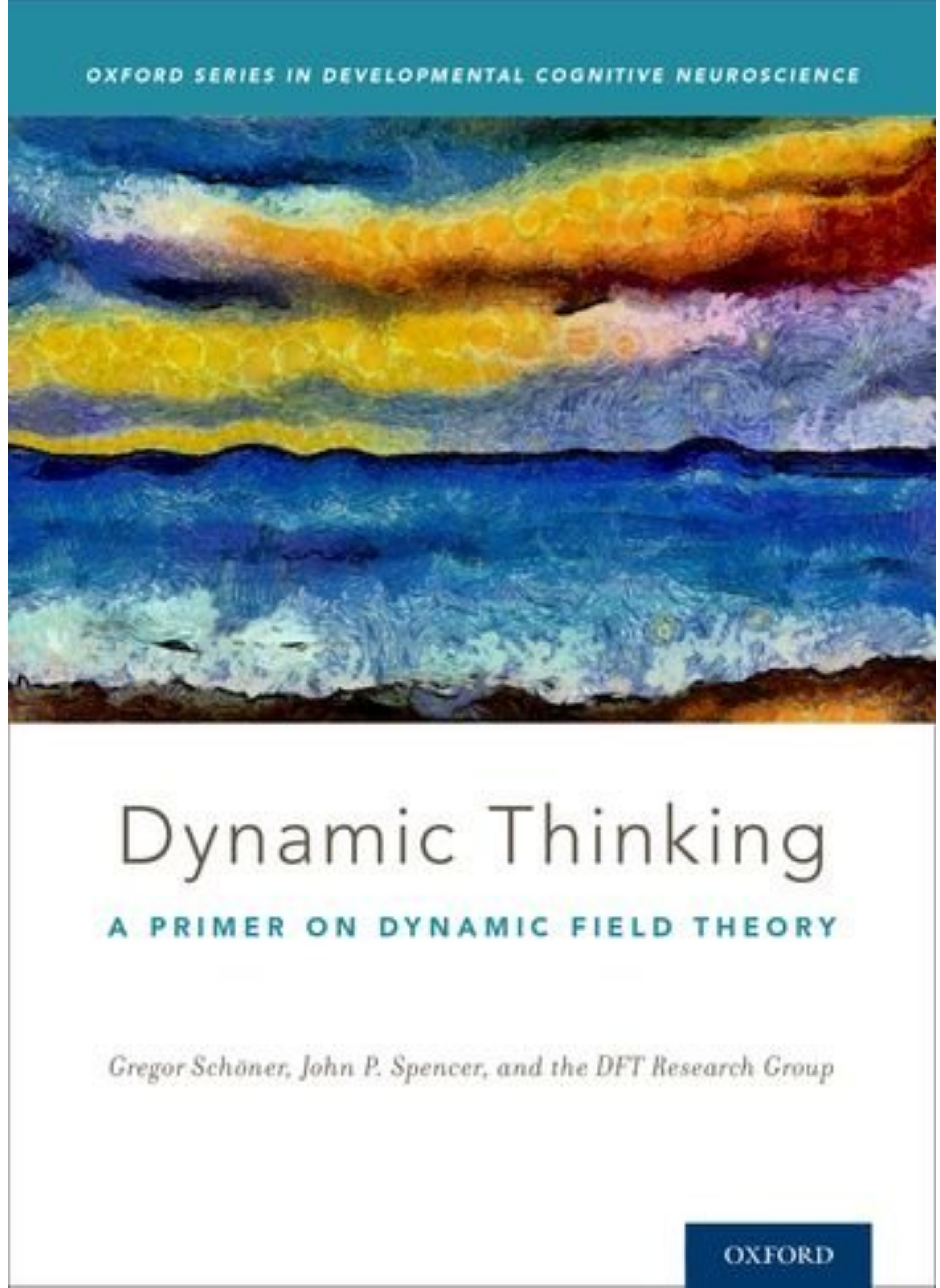
- the detection and its reverse create events at discrete times from time-continuous changes



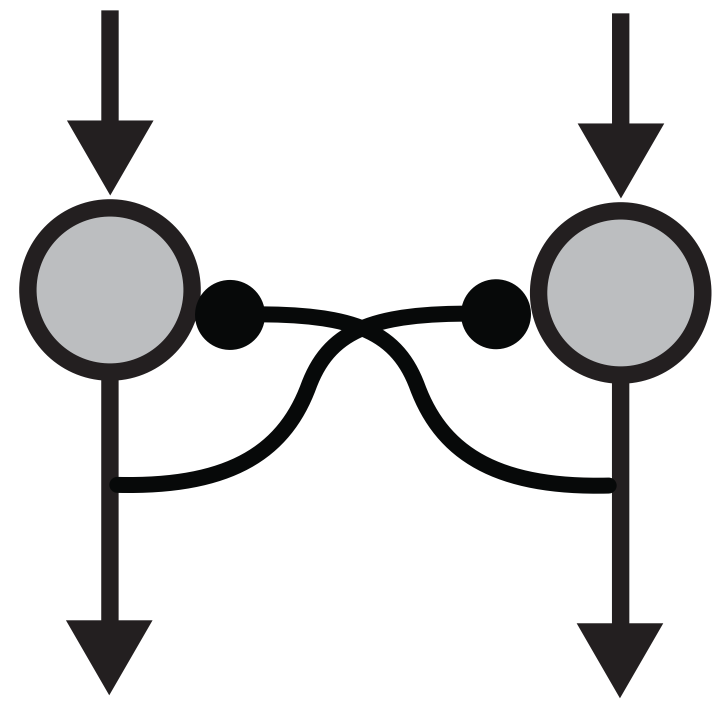
$$\tau \dot{u}(t) = -u(t) + h + s(t) + c\ \sigma(u(t))$$

=> simulation

- dynamicfieldtheory.org



Neuronal dynamics with inhibitory recurrent connectivity



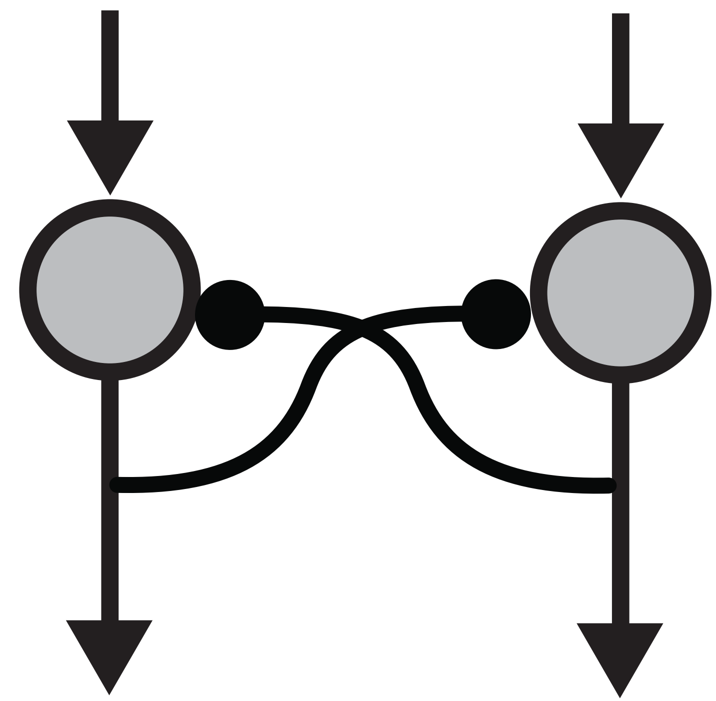
coupling/interaction

$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

Neuronal dynamics with inhibitory recurrent connectivity

- => competition/selection
- two possible attractor stats
  - $u_2 > 0$ and $u_1 < 0$
  - $u_2 < 0$ and $u_1 > 0$

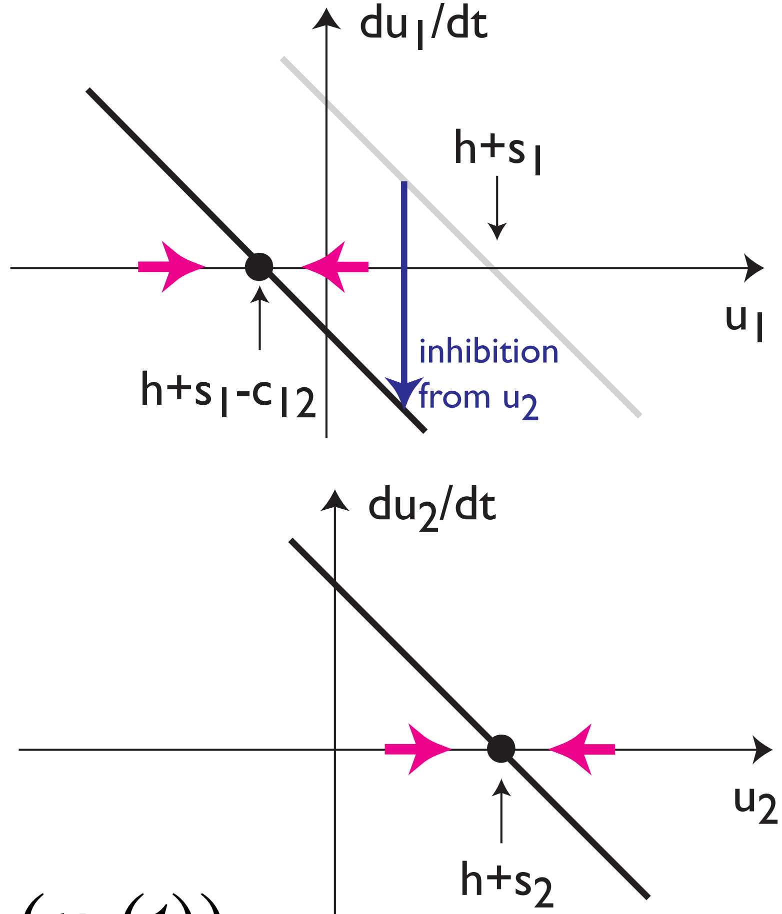


$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

Neuronal dynamics with inhibitory recurrent connectivity

- to visualize, assume that u_2 has been activated by input to a positive level
- => it inhibits u_1



$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

Neuronal dynamics with inhibitory recurrent connectivity

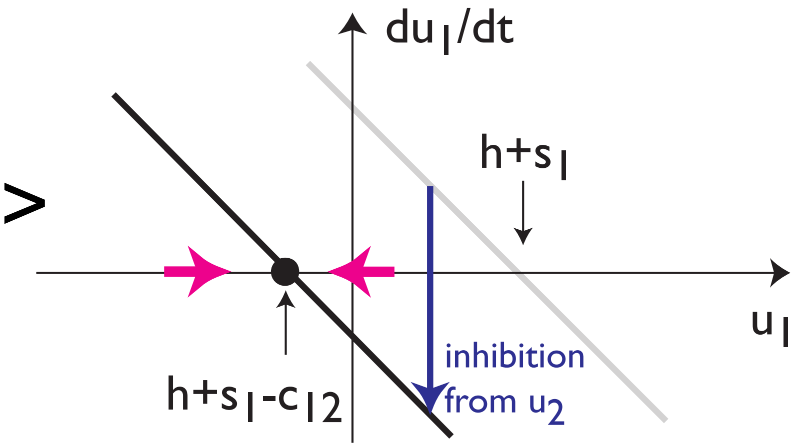
- symmetry: same logic if u_1 was initially activated it would prevent u_2 from activating
- => bistable selection of either u_1 or u_2

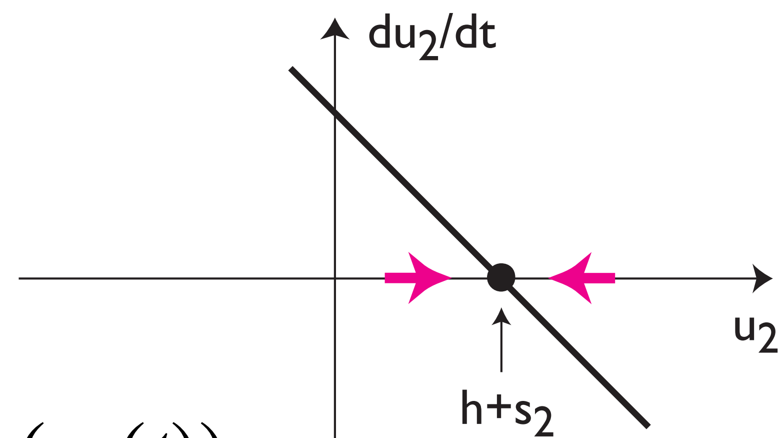
$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

Neuronal dynamics with inhibitory recurrent connectivity

- asymmetric case: e.g. more input to u_2 (better "match") => faster increase => u_2 selected
- => input advantage => time advantage => competitive advantage



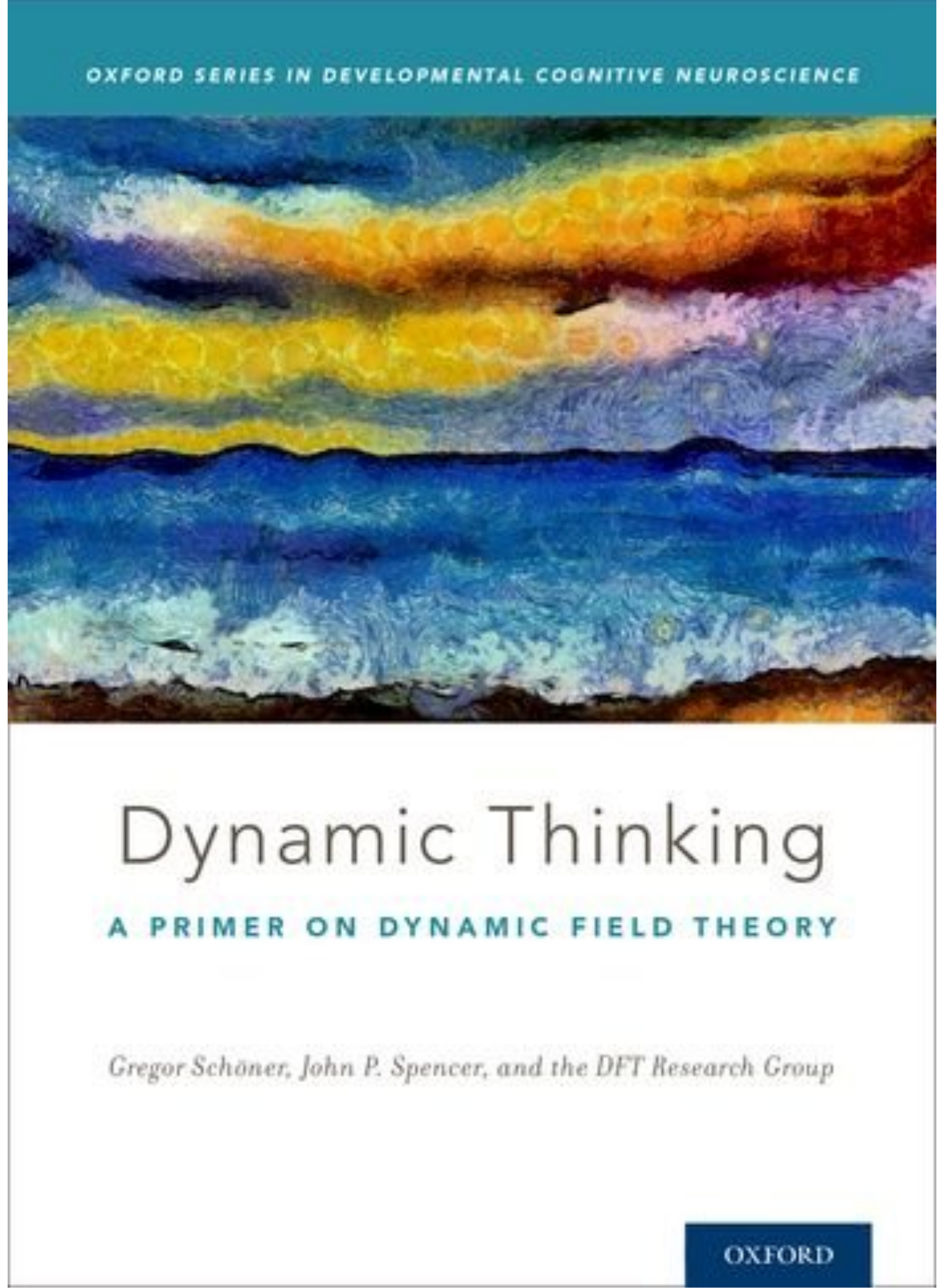


$$\tau \dot{u}_1(t) = -u_1(t) + h + s_1(t) - c_{12}\sigma(u_2(t))$$

$$\tau \dot{u}_2(t) = -u_2(t) + h + s_2(t) - c_{21}\sigma(u_1(t))$$

=> simulation

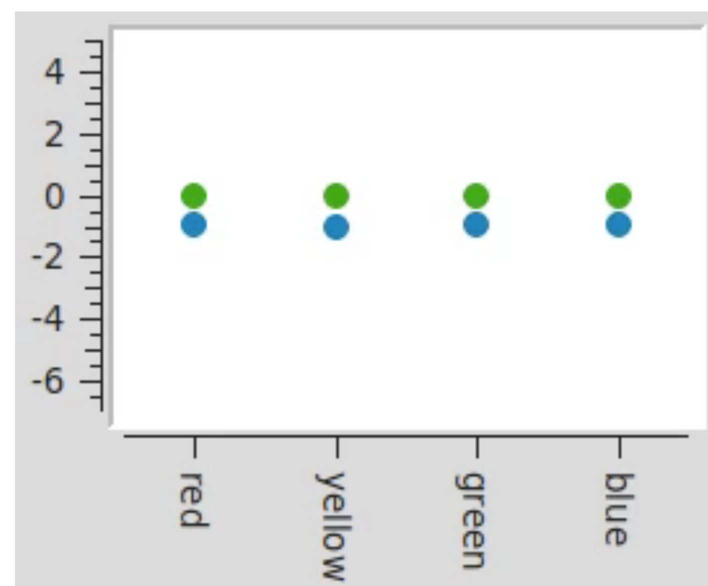
- dynamicfieldtheory.org



Neural dynamic nodes

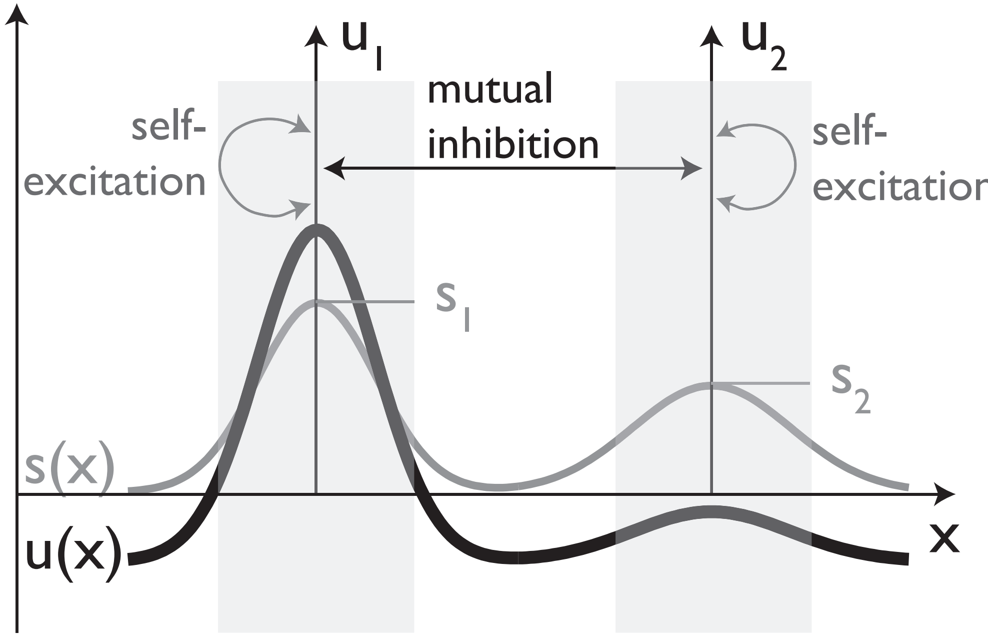
- discrete activation variables: nodes
  - that are self-excitatory: “on” vs “off” states, detection instability regulates switch between these
  - that are coupled inhibitorily: “on” states compete… selection

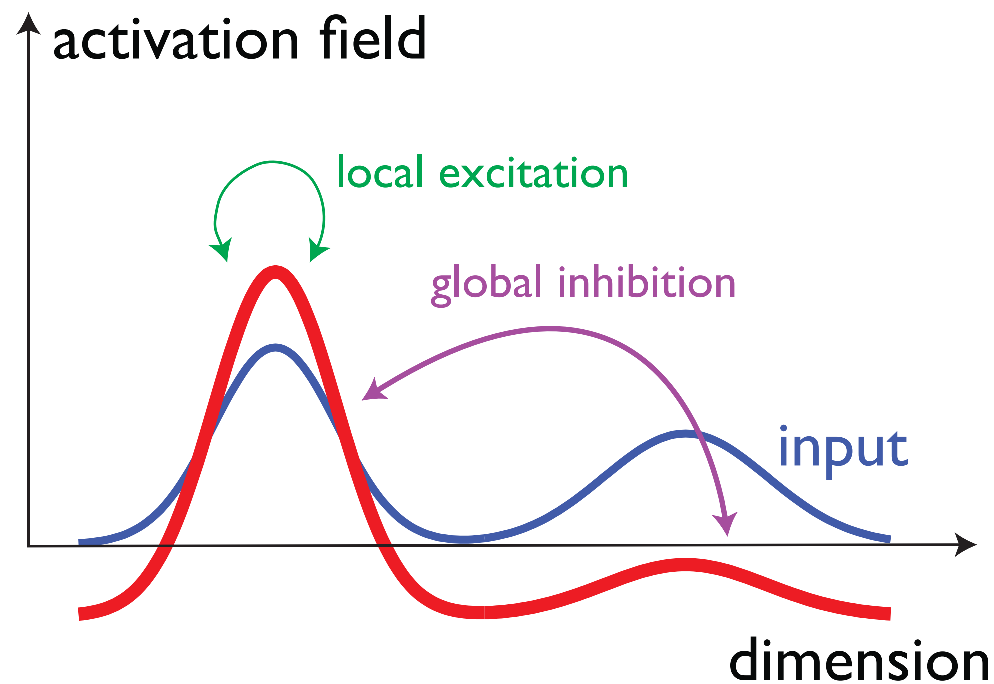
0-dimensional



Neural dynamics of fields

- embed activation variables in continuous dimensions, x
- detection: self-excitation => location excitation
- selection: => global inhibition
- interaction organized along a dimension, x…
- (meaning of dimension: next lecture)

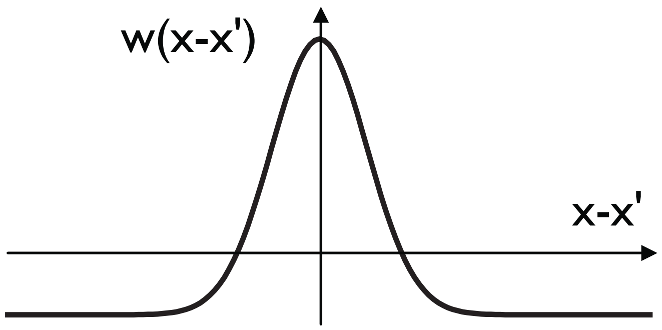


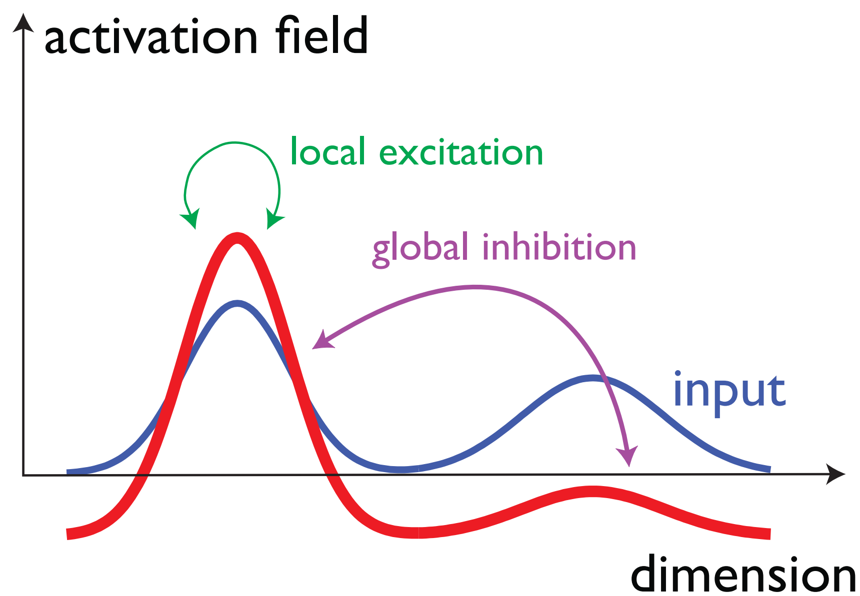


Neural dynamics of fields

- kernel: local excitatory interaction/ global inhibitory interaction

$$w(x - x') = w_{\text{exc}} e^{-\frac{(x - x')^2}{2\sigma^2}} - w_{\text{inh}}$$

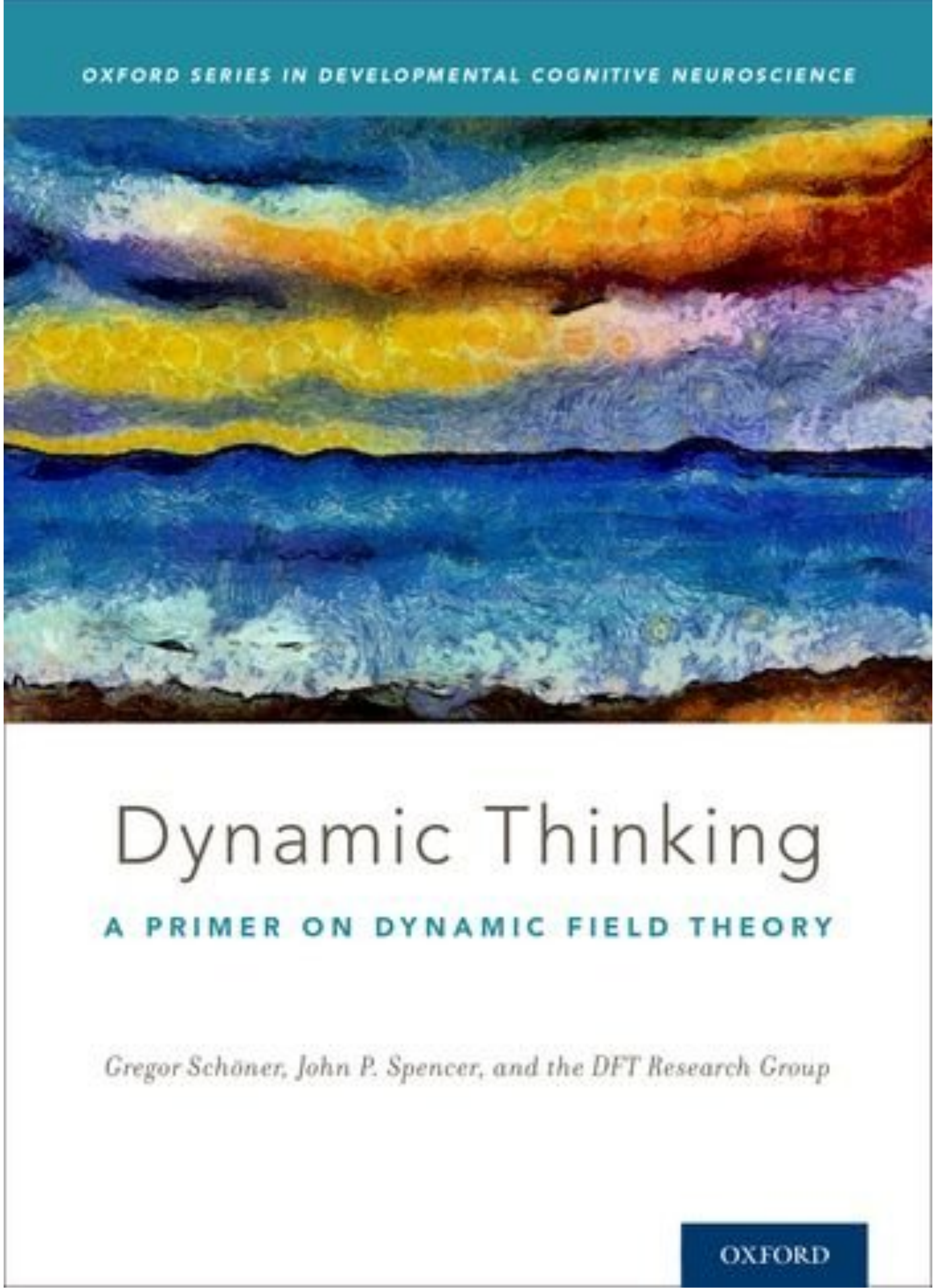




$$\tau \dot{u}(x,t) = -u(x,t) + h + s(x,t) + \int dx' \; w(x - x') \; \sigma(u(x'))$$

=> simulation

- dynamicfieldtheory.org



Attractors and their instabilities

- input driven solution (sub-threshold)
- self-stabilized solution (peak, supra-threshold)
- selection / selection instability
- working memory / memory instability
- boost-driven detection instability

↓ detection instability

↑ reverse detection instability

Noise is critical near instabilities

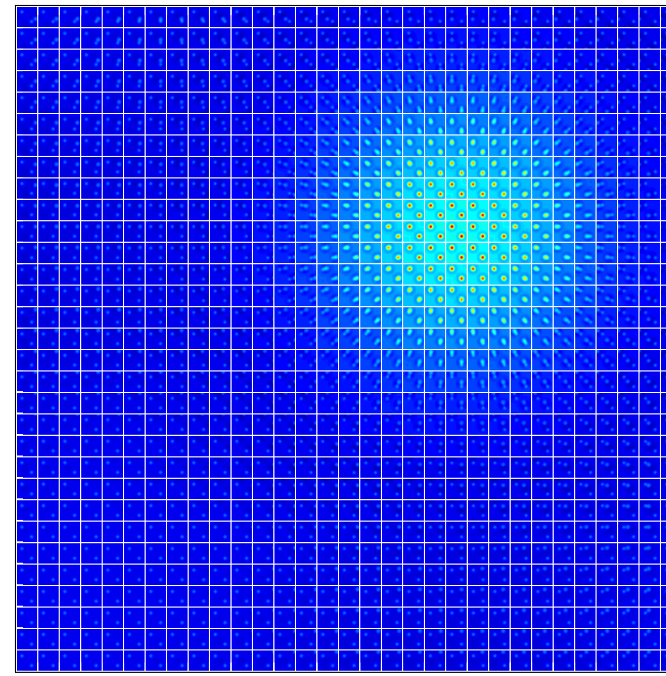
Dynamic regimes

- which attractors and instabilities arise as input patterns are varied
- examples
  - “perceptual regime”: mono-stable sub-threshold => bistable sub-threshold/peak => mono-table peak..
  - “working memory regime” bistable sub-threshold/peak => mono-table peak.. without mono-stable sub-threshold
  - single (“selective”) vs. multi-peak regime

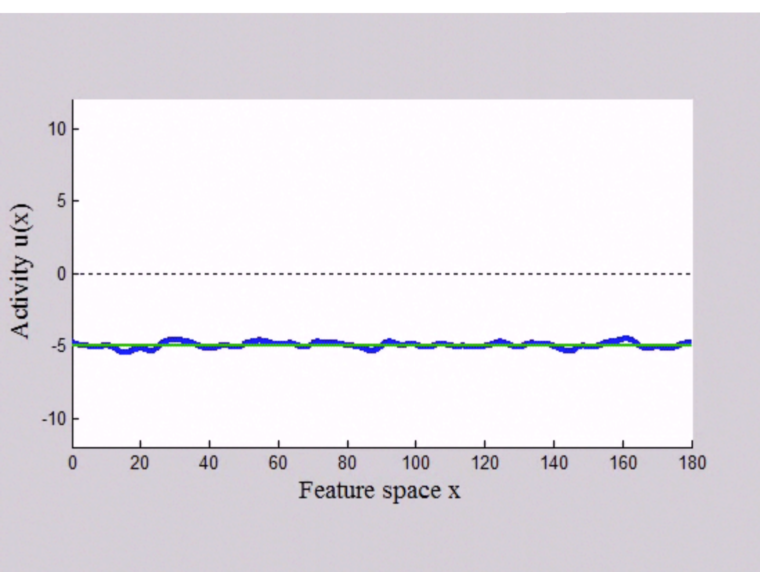
Embedding space may vary in dimensionality

- 1, 2, 3, 4… dimensions: peaks/ blobs as attractors

4-dimensional

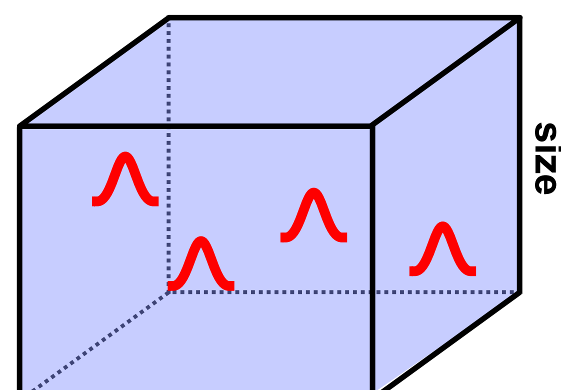


1-dimensional



2-dimensional

3-dimensional



Foundations 1: Neural dynamics

- Neuro-physics
- Neural dynamics
- Recurrent neural dynamics
- Neural fields: dynamics