… toward perception, and cognition… Dynamic Field Theory

Gregor Schöner

up to this point we have examined processes of action planning and control:

- that generate time courses of behavioral variables to steer a system toward desired states while satisfying constraints

autonomy

- for both planning and control, autonomy comes from the availability of sensory information about the world= the environment and the system itself
  - detecting targets, obstacles
  - estimating ego-position and -orientation
  - recognizing land-marks ect.
- => perception = extracting from sensory signals information about the world

autonomy

- autonomy also involves making major behavioral changes, activating or deactivating different behaviors. Ultimately, this always happens under the influence of perception, specifically in response to detection events.
  - e.g., detecting that a grasp as succeeded my activate lifting and transporting the object
- this will be the topic of the next lecture on "behavioral organization"

perception

- 1) detection
- 2) estimation
- 3) classification/recognition

detection

- detection=decide if a particular signal/object etc is present
- Examples:
  - target detection from radar signals
  - detection of communication signals from radio waves
- theoretical approaches:
  - signal detection theory, with varying amounts of prior information about signals and noise (models)
  - framework: statistical hypothesis testing

estimation

- estimation=determine the value of a continuously valued parameter from data, given the presence of a signal (which was detected)
- examples:
    - navigation: determine ego-position from distance sensors, maps, beacons
    - control: estimate parameters of plant
    - motion planning constraints: estimate pose and position of targets

estimation: tracking

- tracking in the special, but common case that the parameters that must be estimated vary continuously in time, so that previous estimated values can be used to to estimate updated values
- example: tracking a target
- theoretical approaches
  - (optimal) estimation theory based on various amounts of a priori knowledge about the system
  - Optimal filtering, Kalman filtering, particle filters

classification

- classification=given that a signal has been detected, assign that signal to one class within a set of discrete classes
- examples:
  - binary classification (target yes or no)
  - decoding in (digital) telecommunication
  - recognition: letters, speech, objects, ...

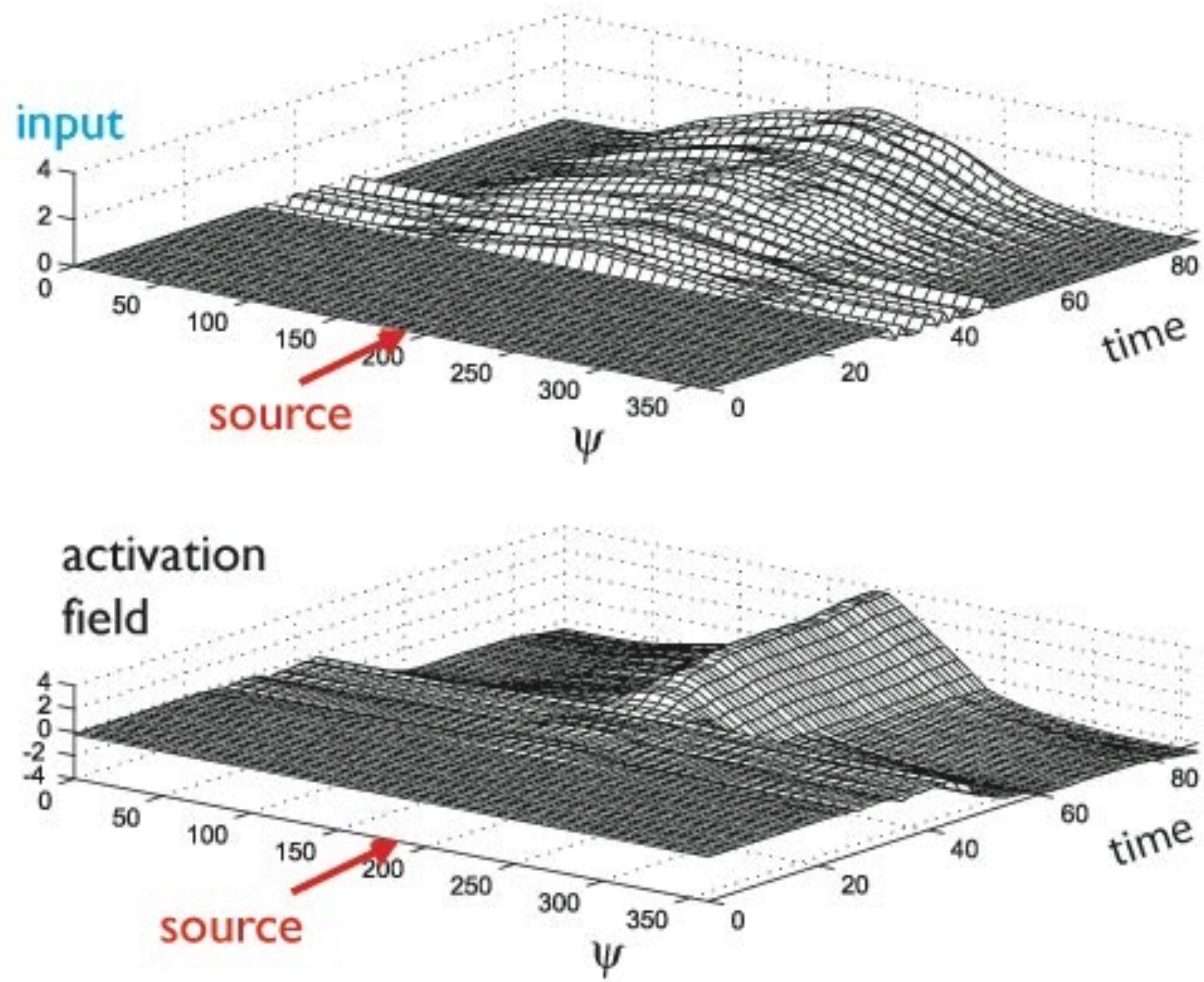
classification

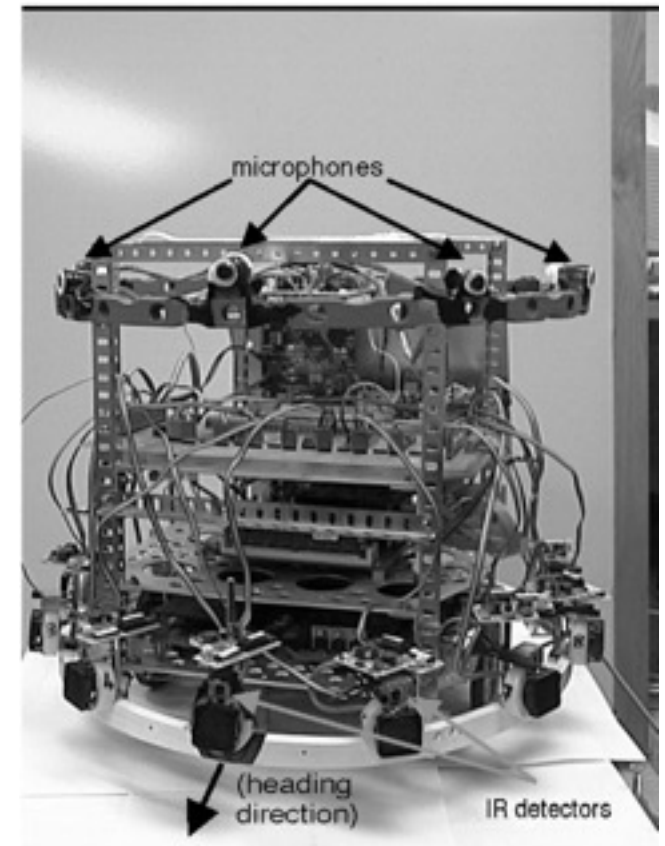
- theoretical approaches:
  - statistical hypothesis testing within metrics of feature/code space to separate distributions (discrimination)
  - (detection being a special case of classification)
  - neural networks, learning
  - statistical learning theory: support vector machines
  - link to coding: optimal code that maximize distances in code space between classes

Dynamic Field Theory

- supports these perceptual and cognitive processes through (neural) dynamics

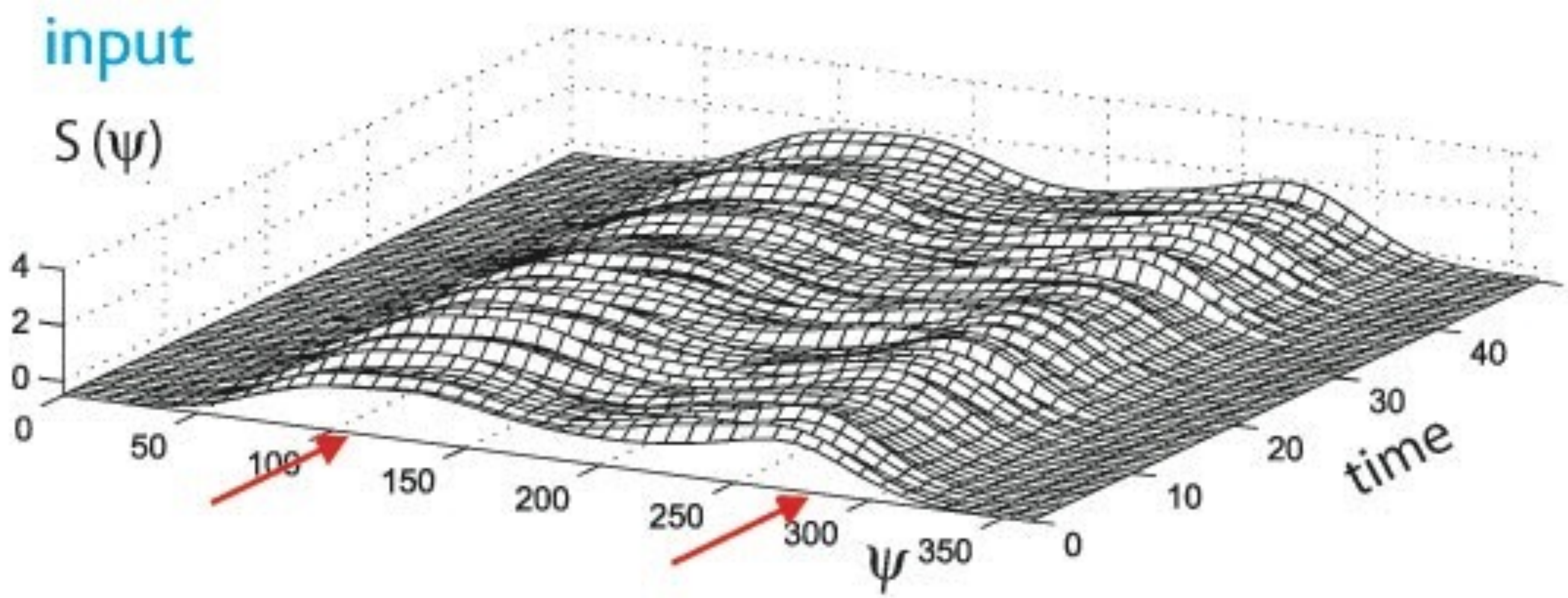
detection instability on a phonotaxis robot

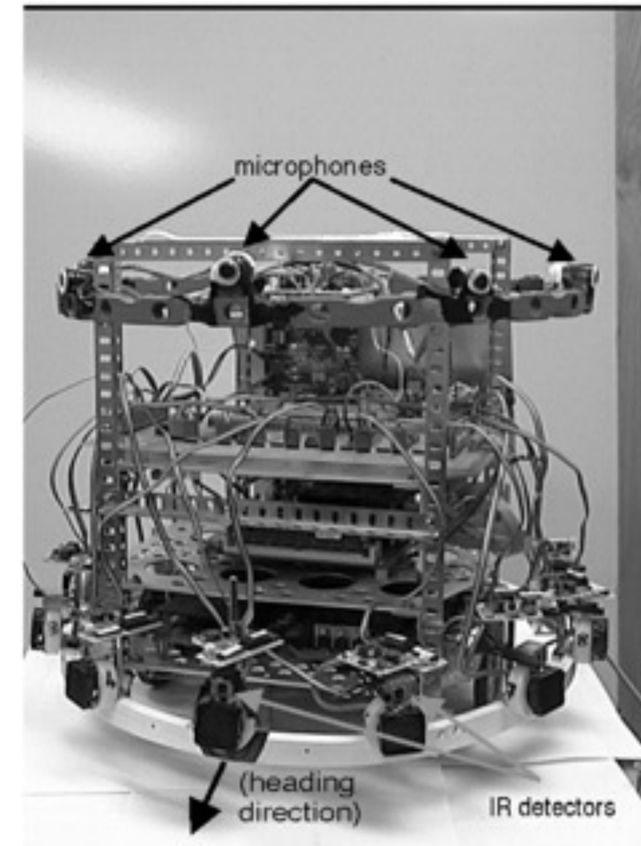


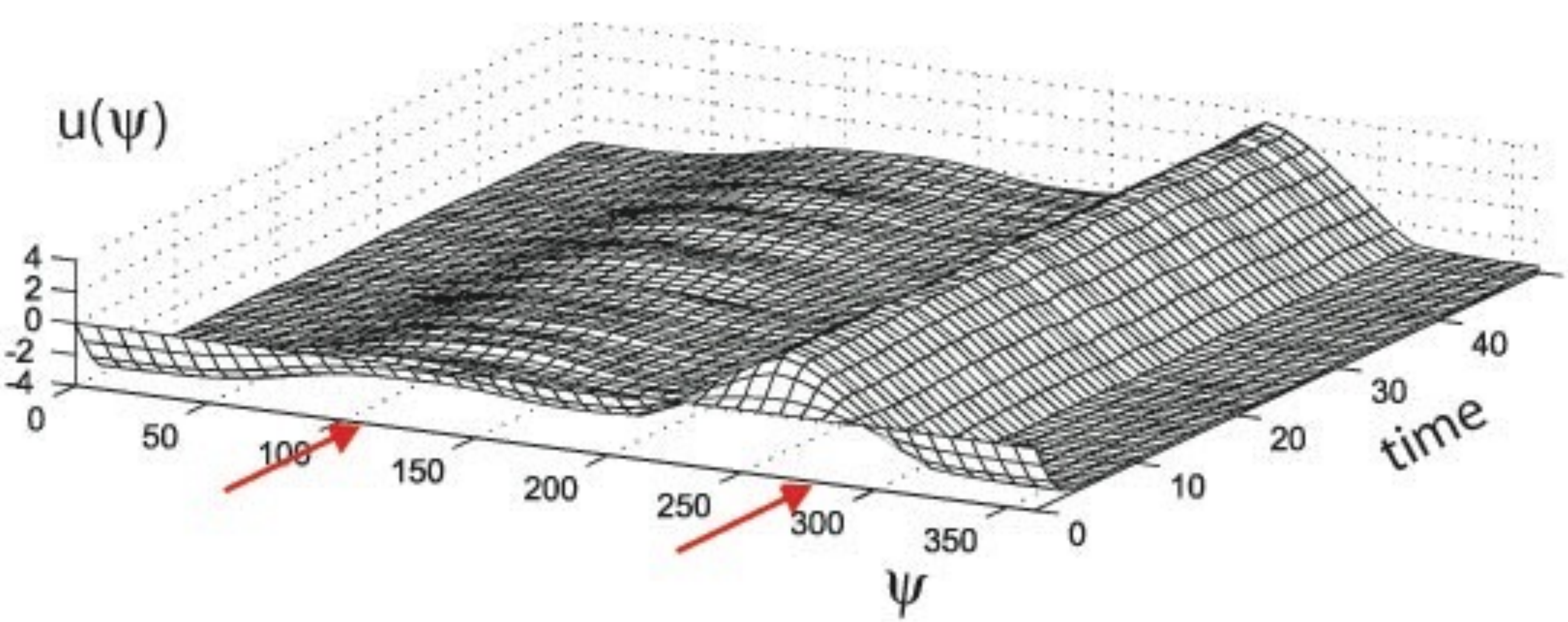


[from Bicho, Mallet, Schöner: Int. J. Rob. Res., 2000]

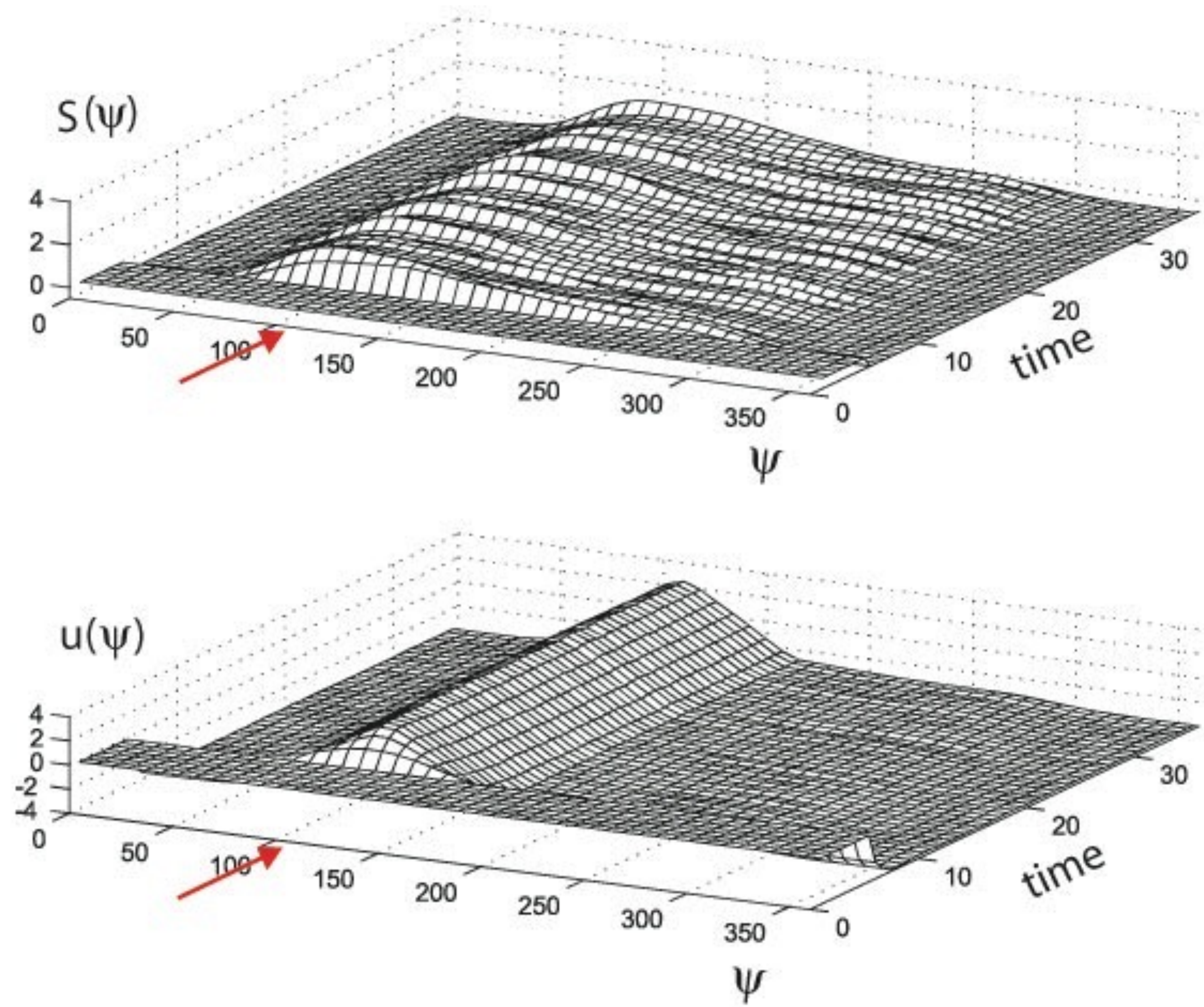
target selection on phonotaxis vehicle



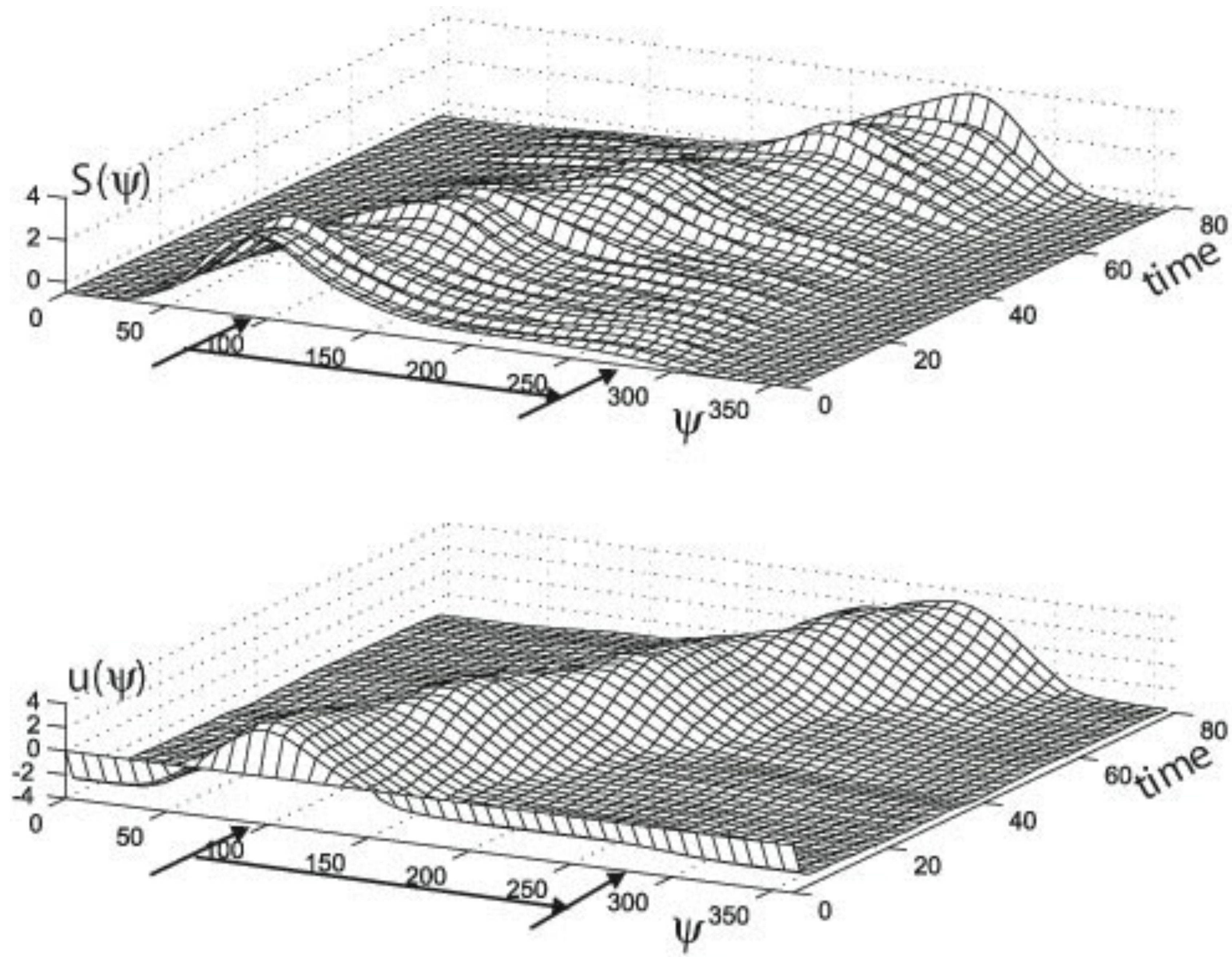




robust estimation



tracking


S(ψ)
4
2
0
0
50
100
150
200
250
300
350
ψ
0
20
40
60
80
time
u(ψ)
4
2
0
-2
-4
0
50
100
150
200
250
300
350
ψ
0
20
40
60
80
time

Dynamic Field Theory

- dimensions
- activation fields
- field dynamics: peaks, instabilities

Dimensions

- different categories of behavior and percepts each form continua, embedded in spaces
  - e.g., the space of possible reaching movements: spanned by the direction in space of the hands velocity
  - e.g., the spaces of possible shapes, colors, poses of a segmented visual object
- => inspired by the central nervous system. Analogy in AI:
  - discrete categories of behaviors/representations
  - continuous parameters of these behaviors/representations

Dimensions

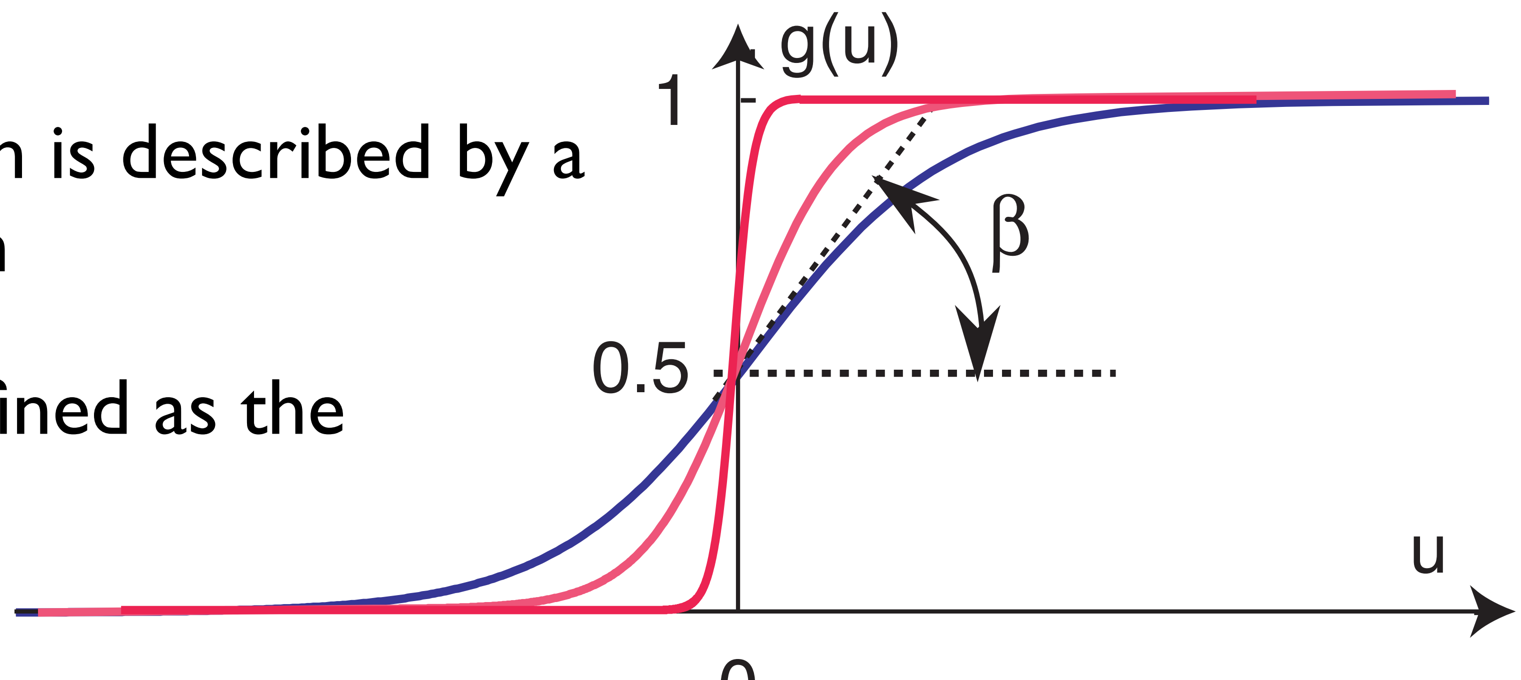
- the continuous spaces form the dimension over which activation fields will be defined
  - homologous to sensory surfaces, e.g., visual or auditory space (retinal, allocentric, ...)
  - homologous to motor surfaces, e.g., saccadic end-points or direction of movement of the end-effector in outer space
  - feature spaces, e.g., localized visual orientations, color, impedance, ...
  - abstract spaces, e.g., ordinal space, along which serial order is represented

Activation

- activation: the notion of an "inner" state of a neural network that is used to mark what is significant about neural activity (=has impact)
- variables that may represent the "inner" state of a neural network…
  - membrane potential of neurons?
  - spiking rate?
  - ... population activation… elaborated in lecture course of the WS on neural dynamics

Activation

- activation: a real number that characterizes the inner state of a "neuron", and abstracts from biophysical details
  - low levels of activation: state of the "neuron" is not transmitted to other systems (e.g., to motor systems)
  - high levels of activation: state is transmitted to other systems
  - that significance of activation is described by a sigmoidal threshold function
  - activation level "zero" is defined as the threshold of that function

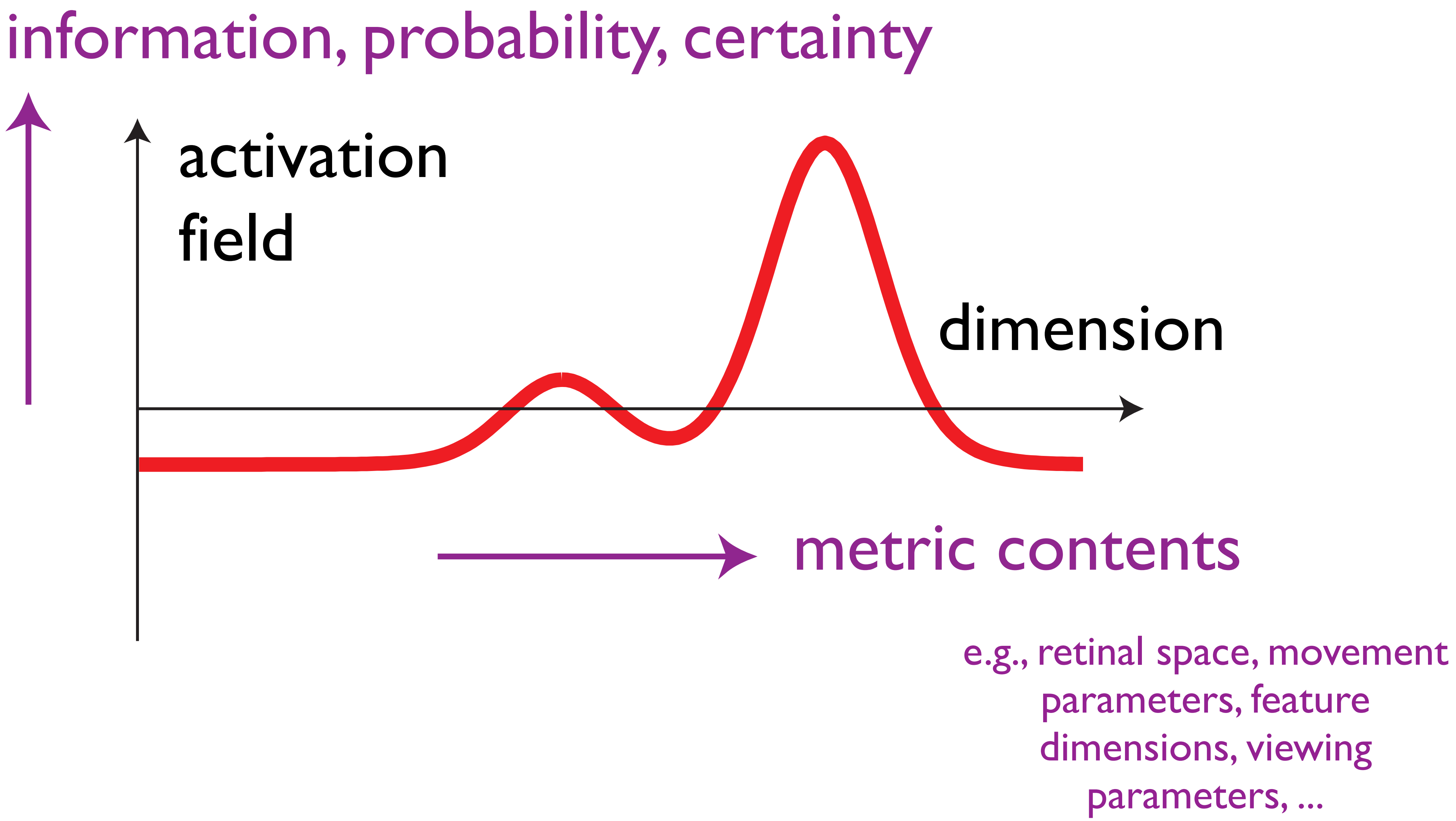


Activation

- connectionist notion of activation: same idea, tied to individual neurons, while DFT looks at more abstract fields of activation
- compare to abstract notion of activation in cognitive architectures defined as production systems (ACT-R, SOAR)
  - related in some way, but also different... activation in thos systems measures how far a module is from emitting its output... and is used to predict how long it will take to finish the computation

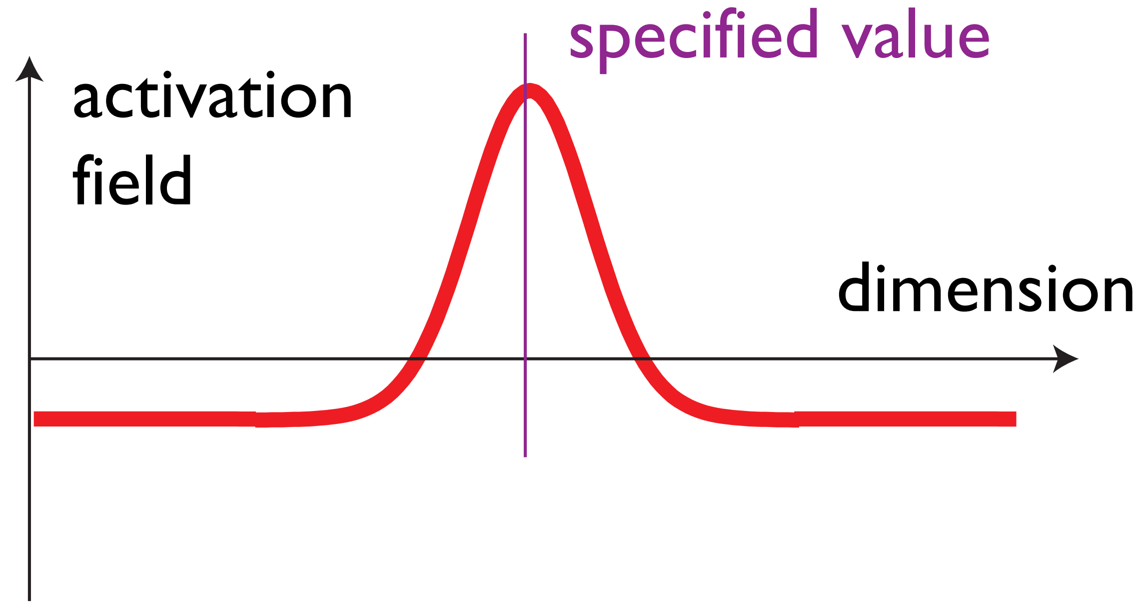
Activation fields

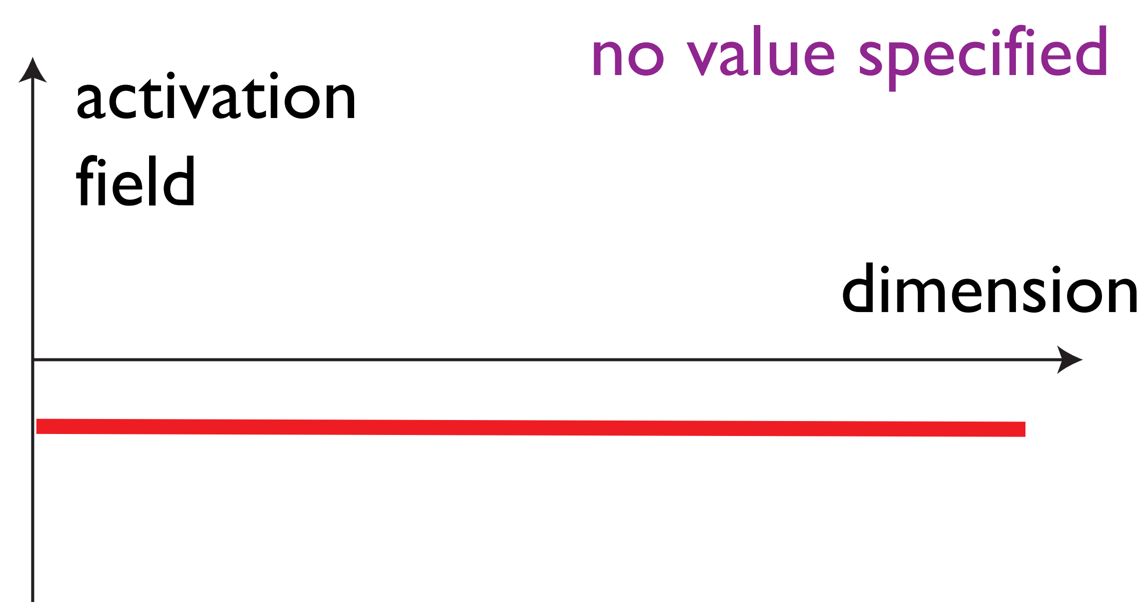
- combine activation and dimensions



Activation fields

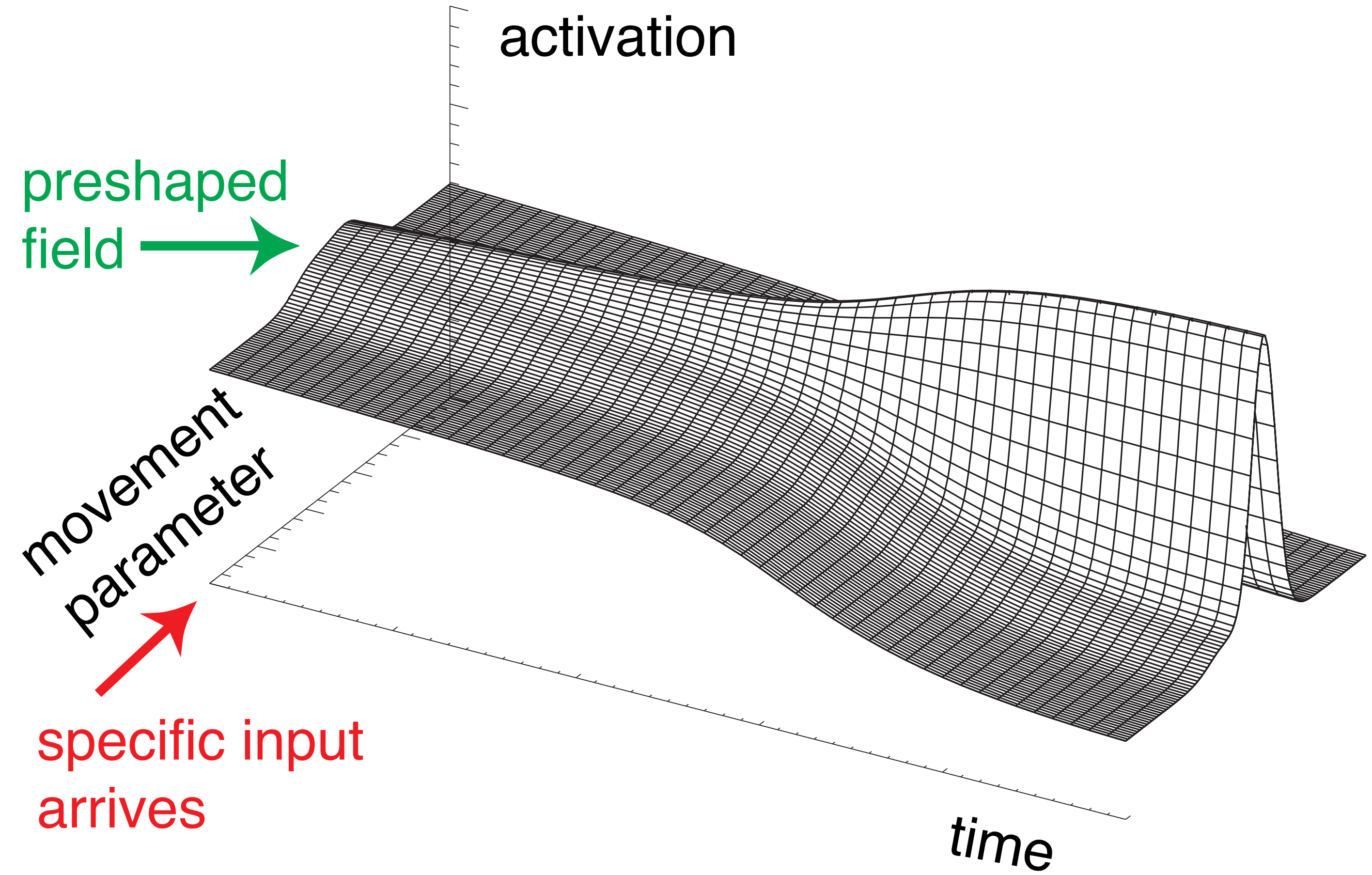
- may represent different states of affairs:
  - localized activation peak: a specific value along the dimension is specified and information about the dimension is thus available
    - had been detected/instantiated
    - and has been estimated/planned
  - flat, sub-threshold activation: no information is available, no value is specified

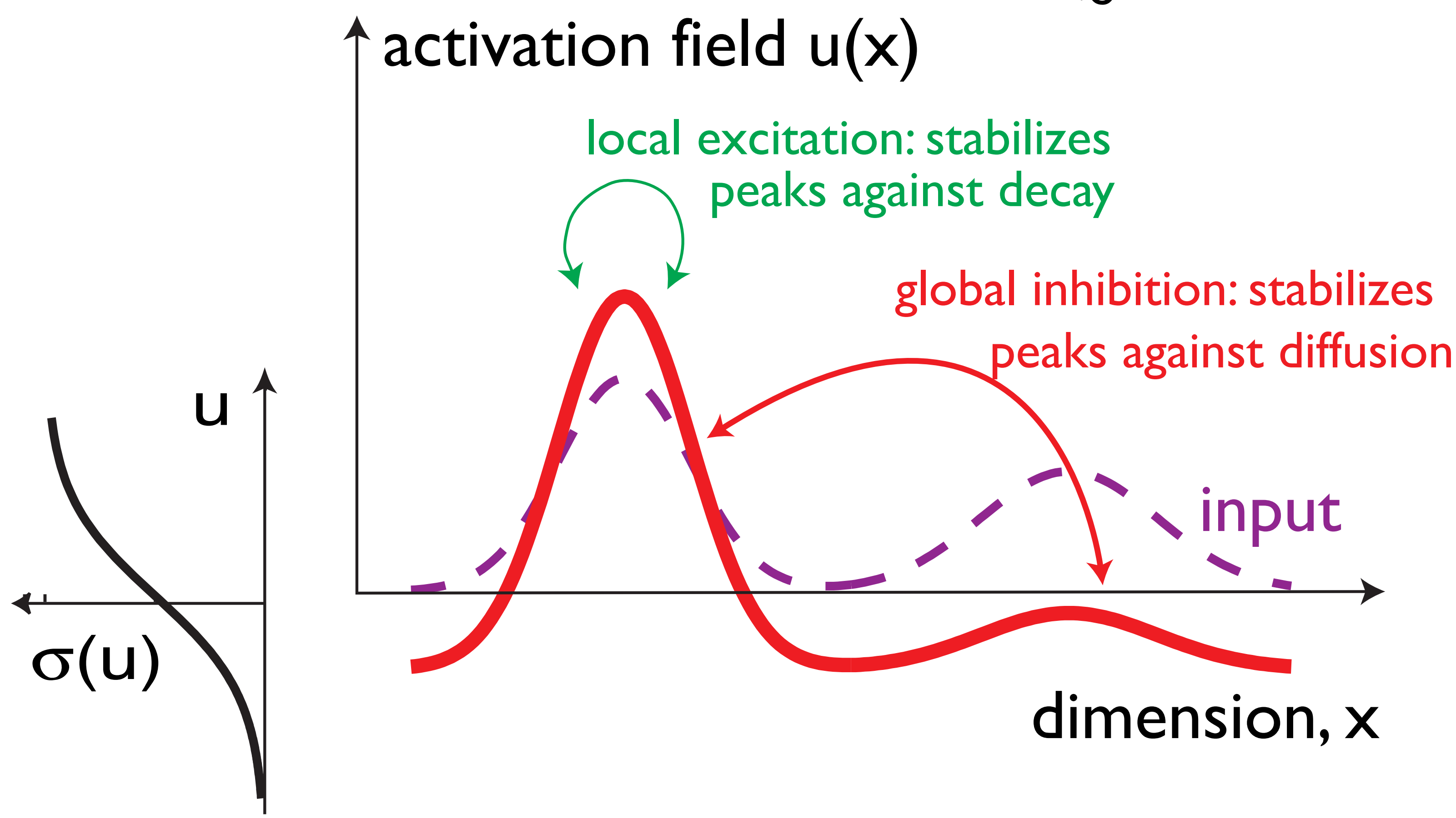




The dynamics activation fields

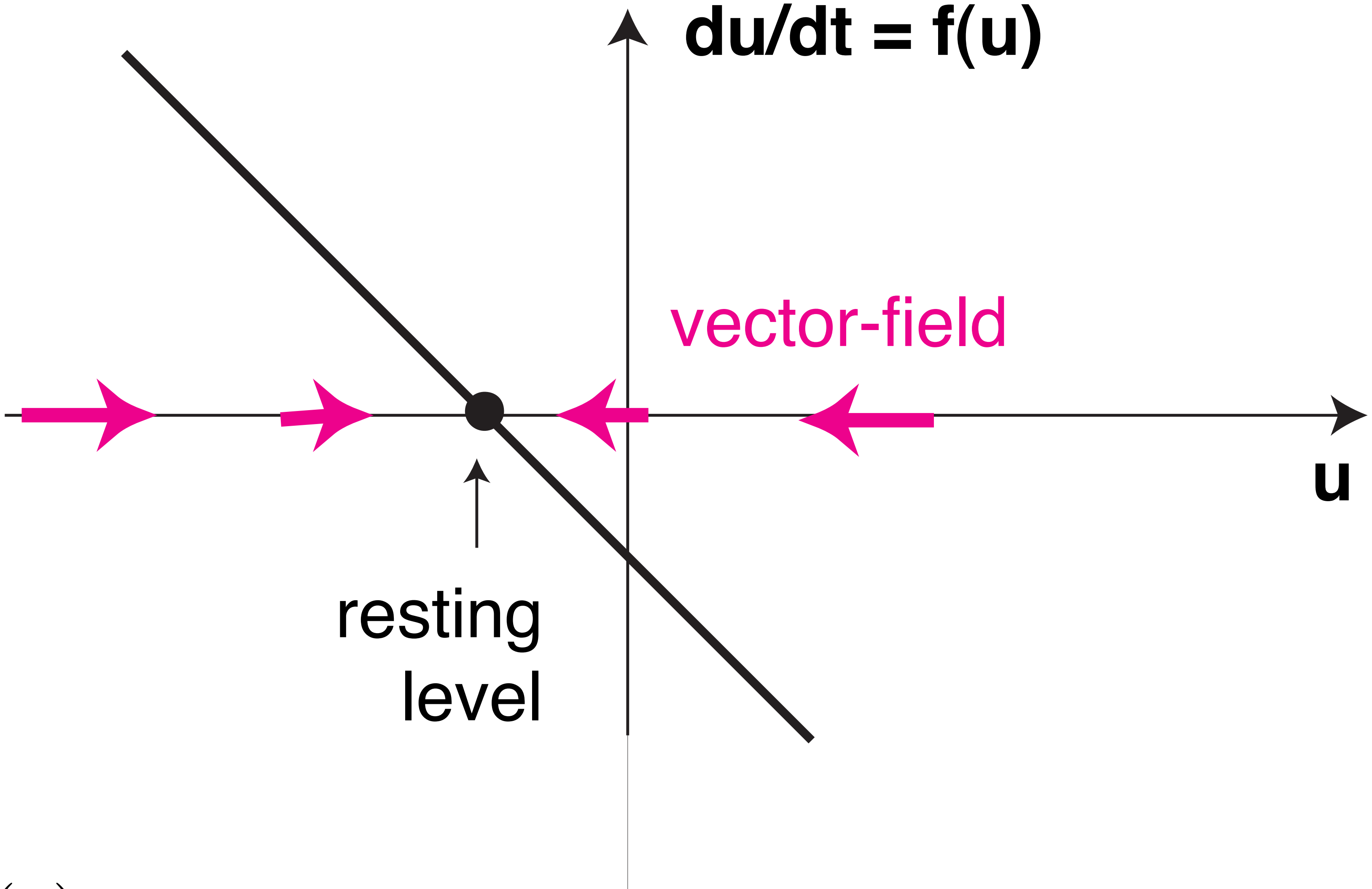
- drives the evolution in time of activation
- such that sub-threshold patterns of activation or peaks are attractors
- either may become unstable in instabilities that are critical to DFT





Neural dynamics

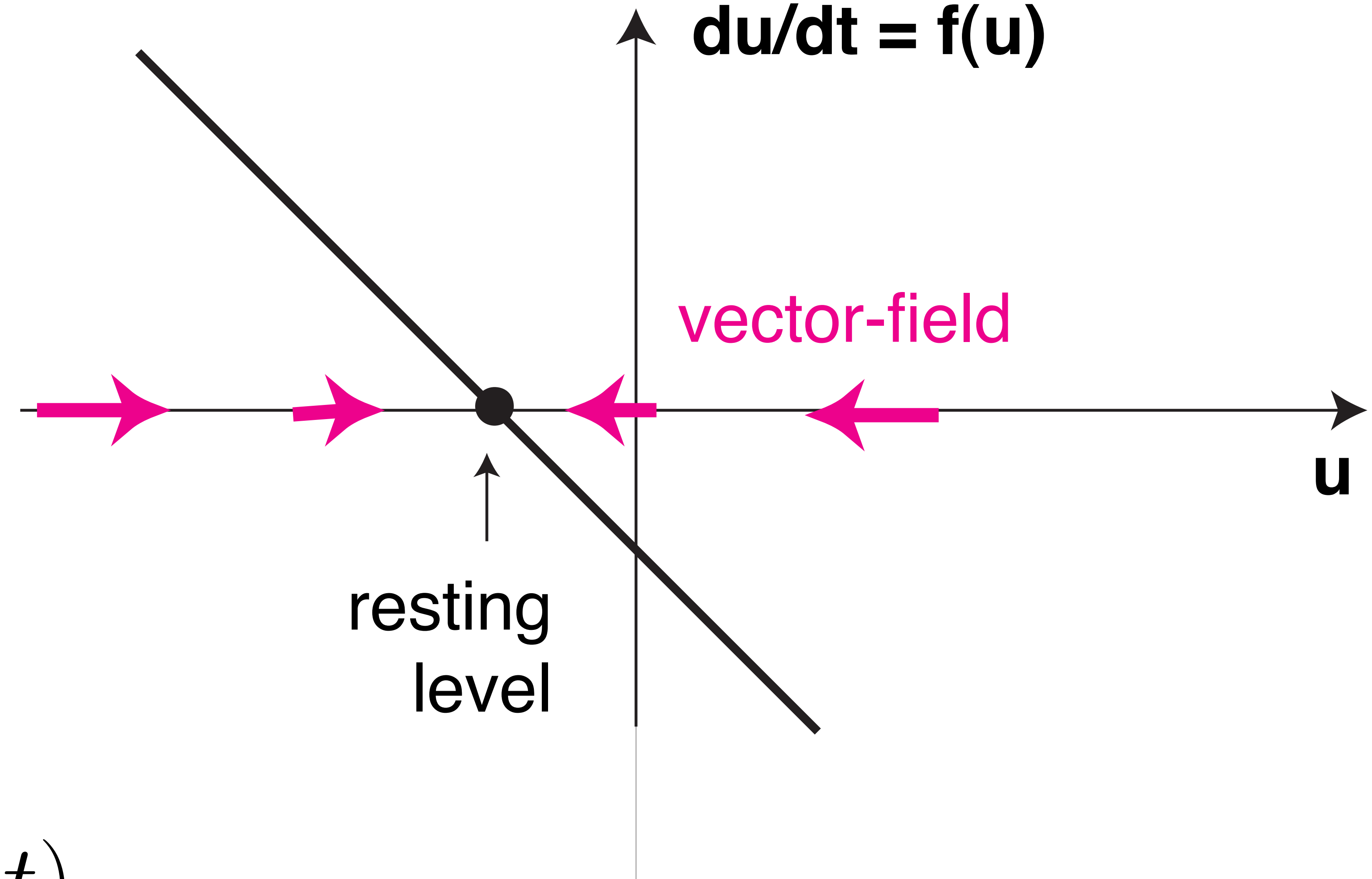
- first, consider a single activation variable u(t) and its dynamics



$$\frac{du(t)}{dt} = \dot{u}(t) = -u(t) + h \qquad (h < 0)$$

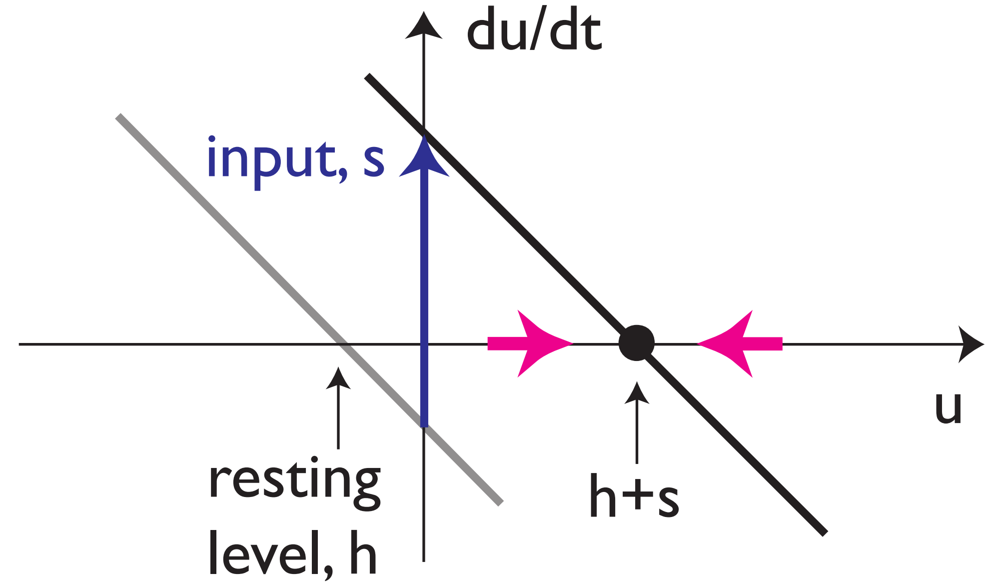
Neural dynamics

- stationary state=fixed point= constant solution
- stable fixed point: nearby solutions converge to the fixed point=attractor

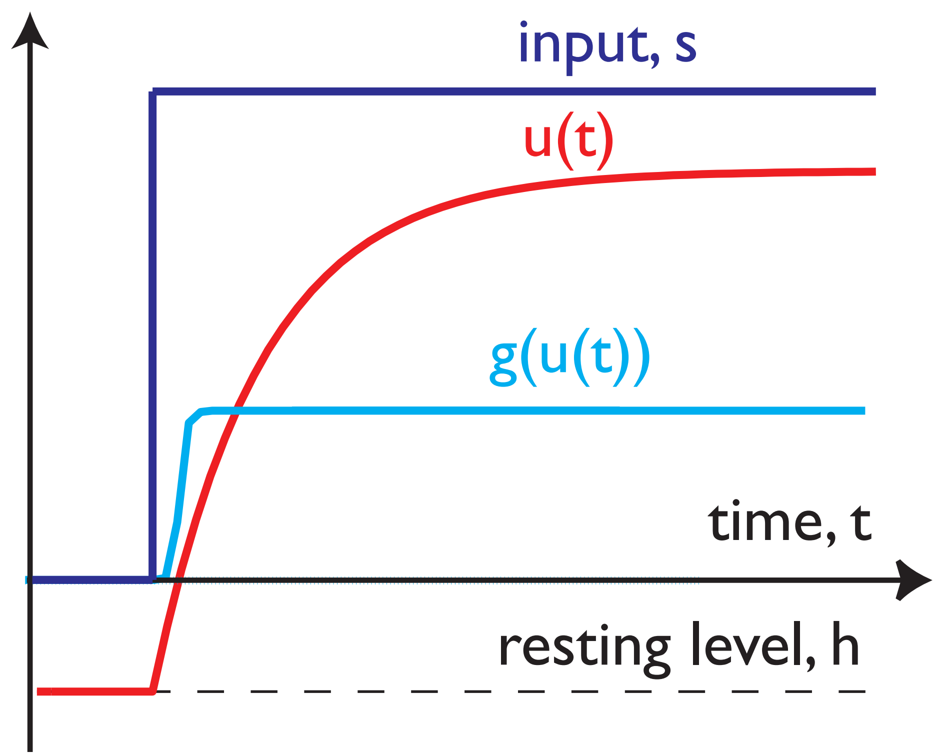


$$\frac{du(t)}{dt} = \dot{u}(t) = -u(t) + h \qquad (h < 0)$$

Neuronal dynamics

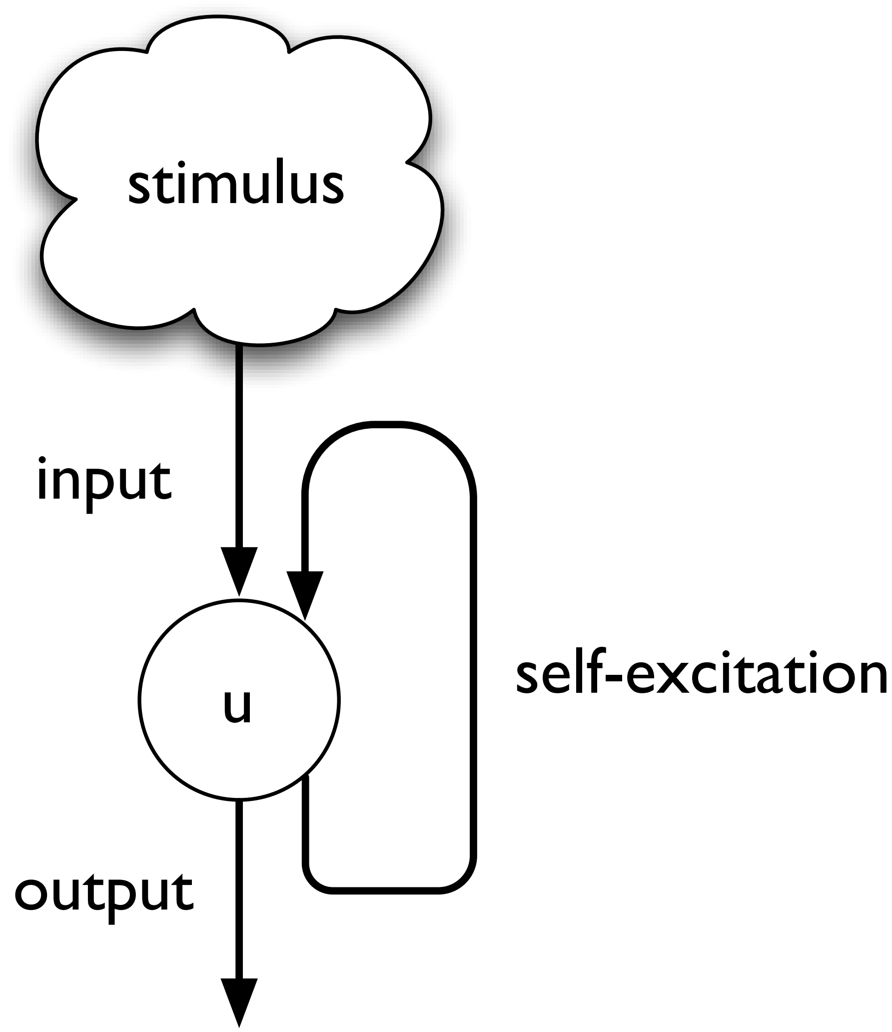


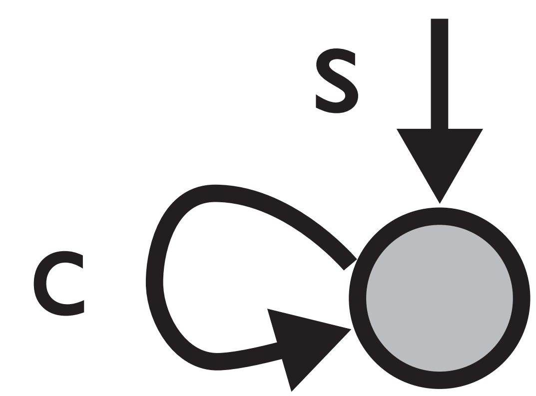
- inputs=contributions to the rate of change
  - positive: excitatory
  - negative: inhibitory
- => shifts the attractor
- activation tracks this shift (stability)



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

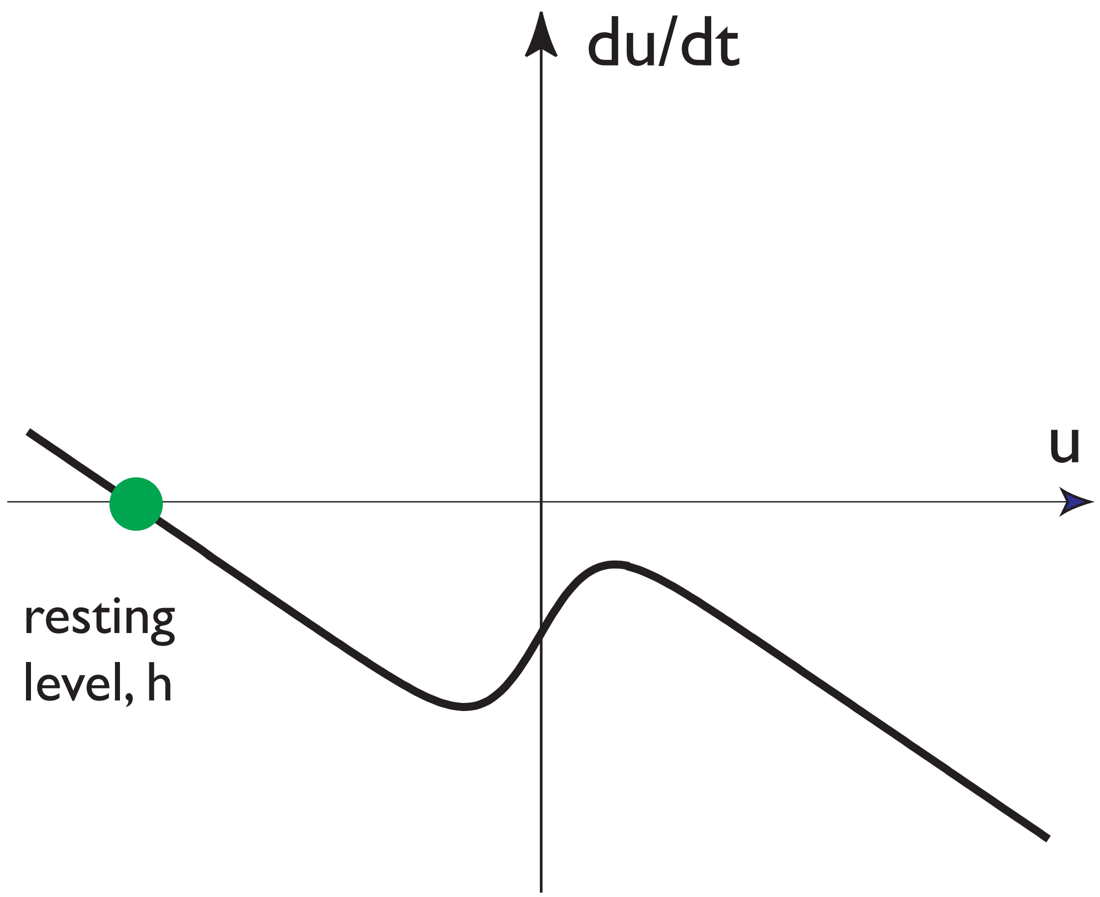
Neuronal dynamics with self-excitation

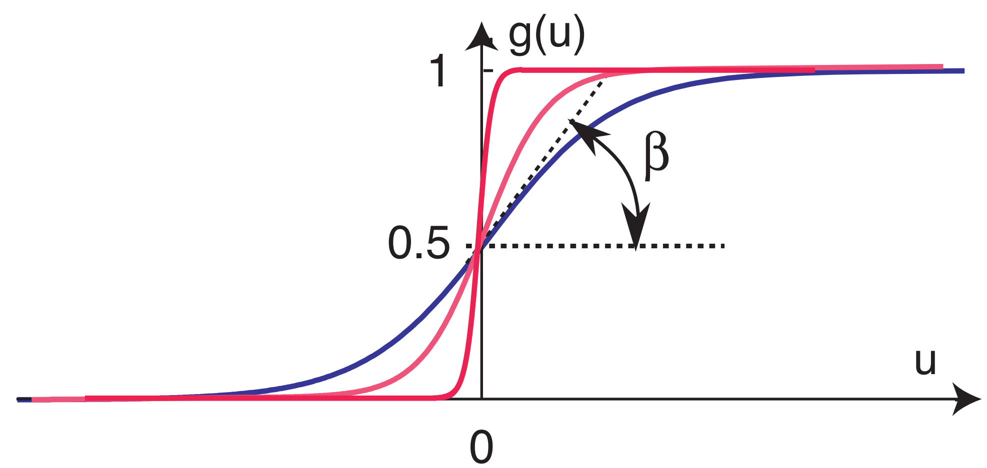




$$\tau \dot{u}(t) = -u(t) + h + S(t) + c\sigma(u(t))$$

Neuronal dynamics with self-excitation

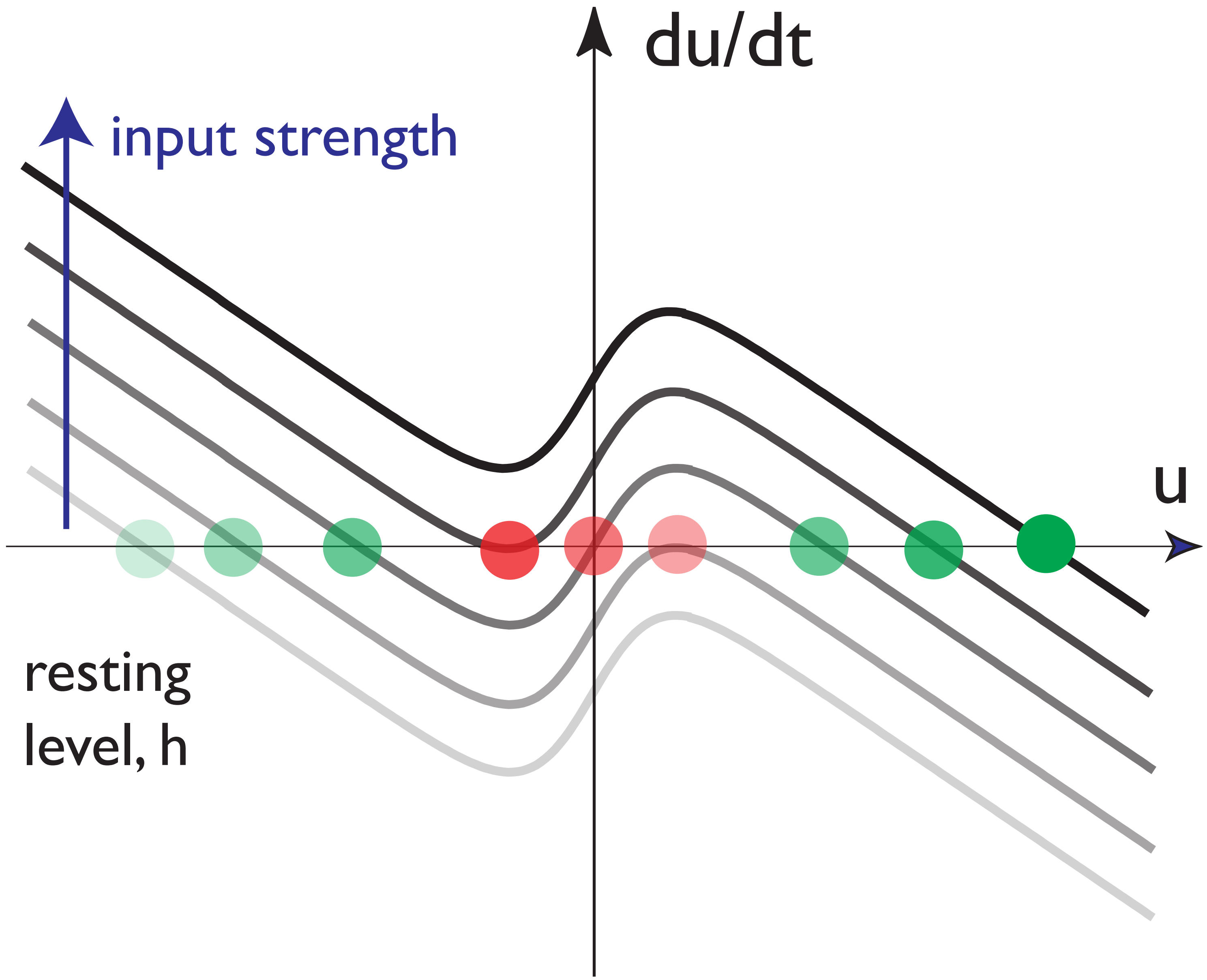




$$\tau \dot{u}(t) = -u(t) + h + S(t) + c\sigma(u(t))$$

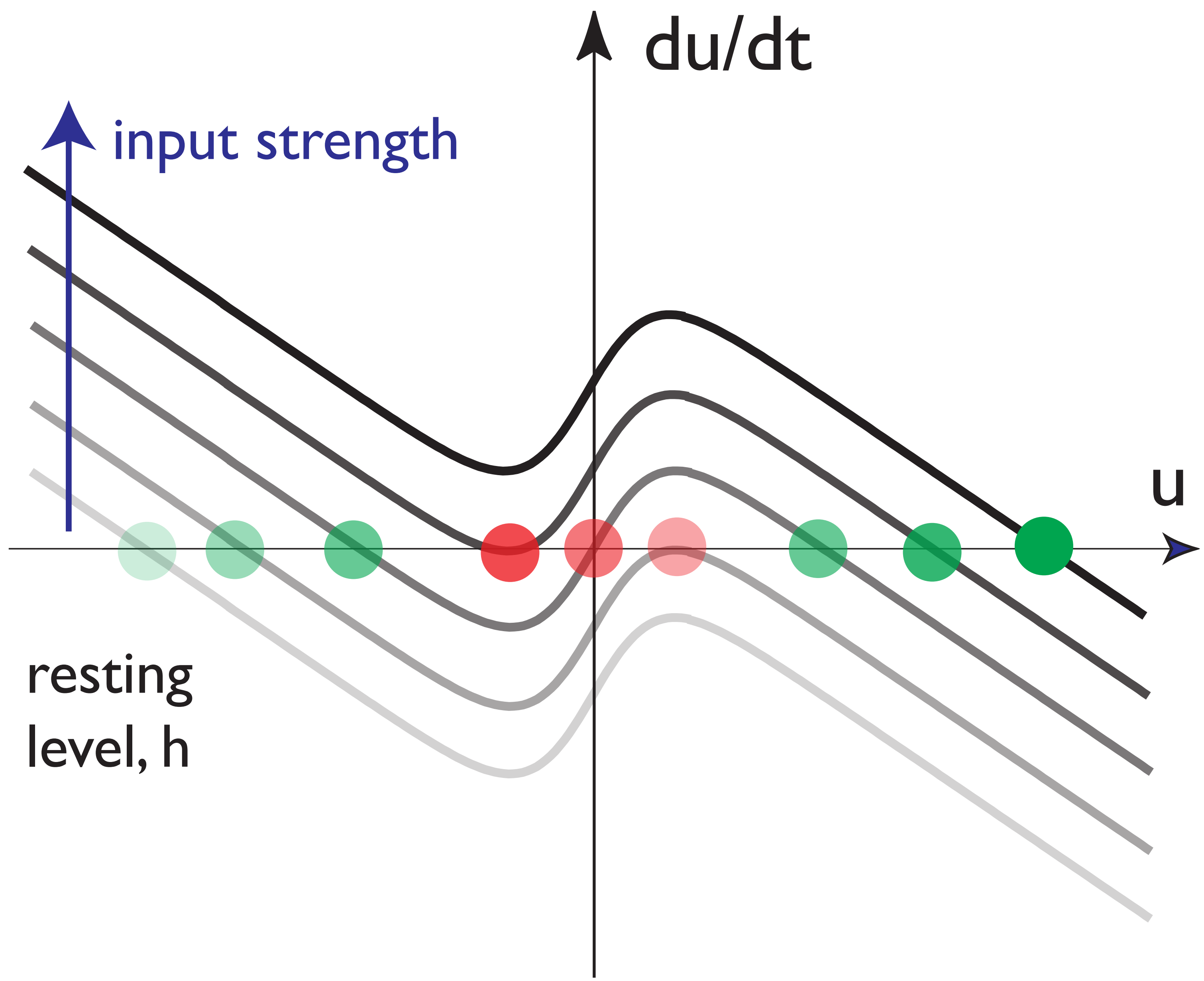
Neuronal dynamics with self-excitation

- stimulus input



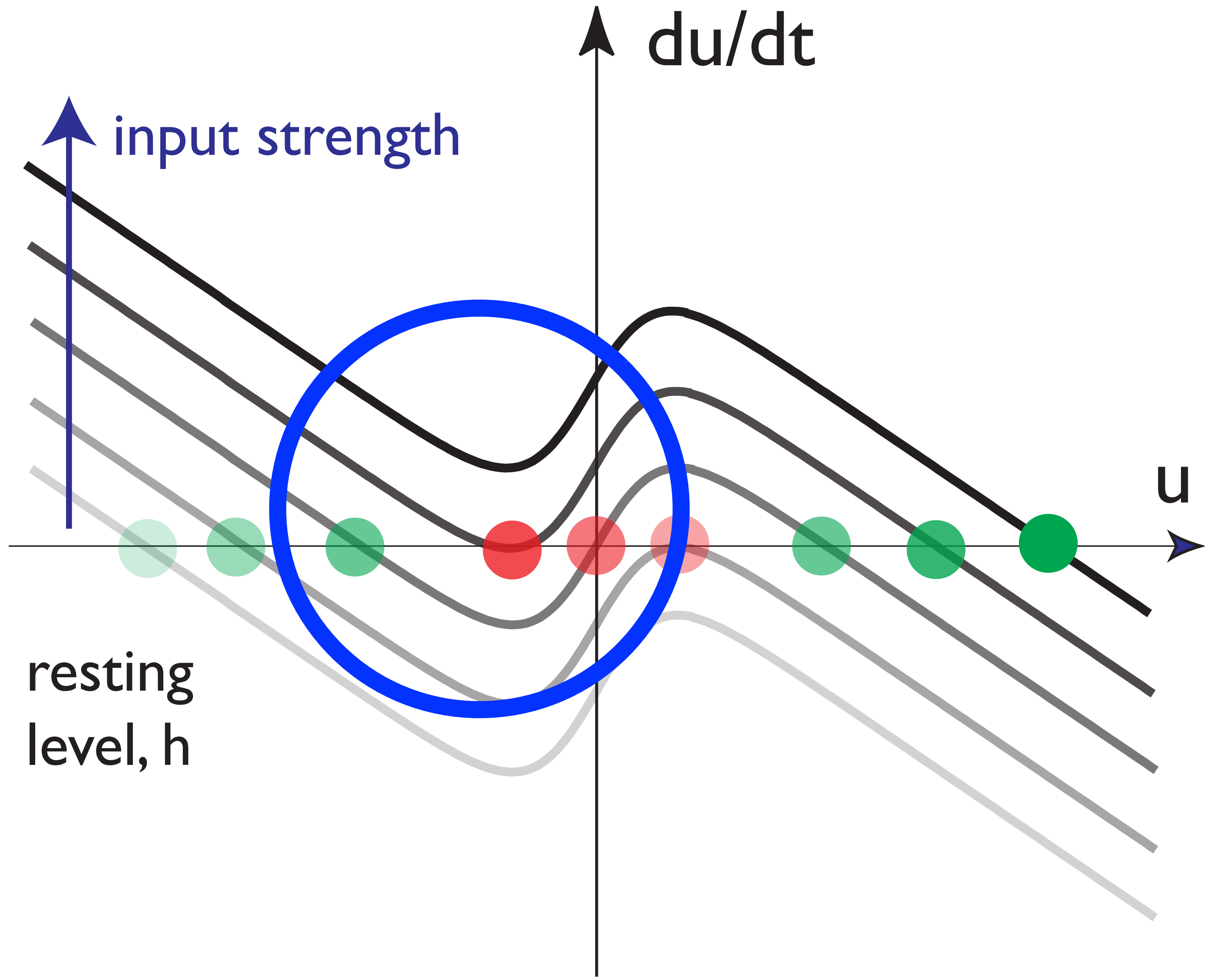
Neuronal dynamics with self-excitation

- with varying input strength system goes through two instabilities: the detection and the reverse detection instability



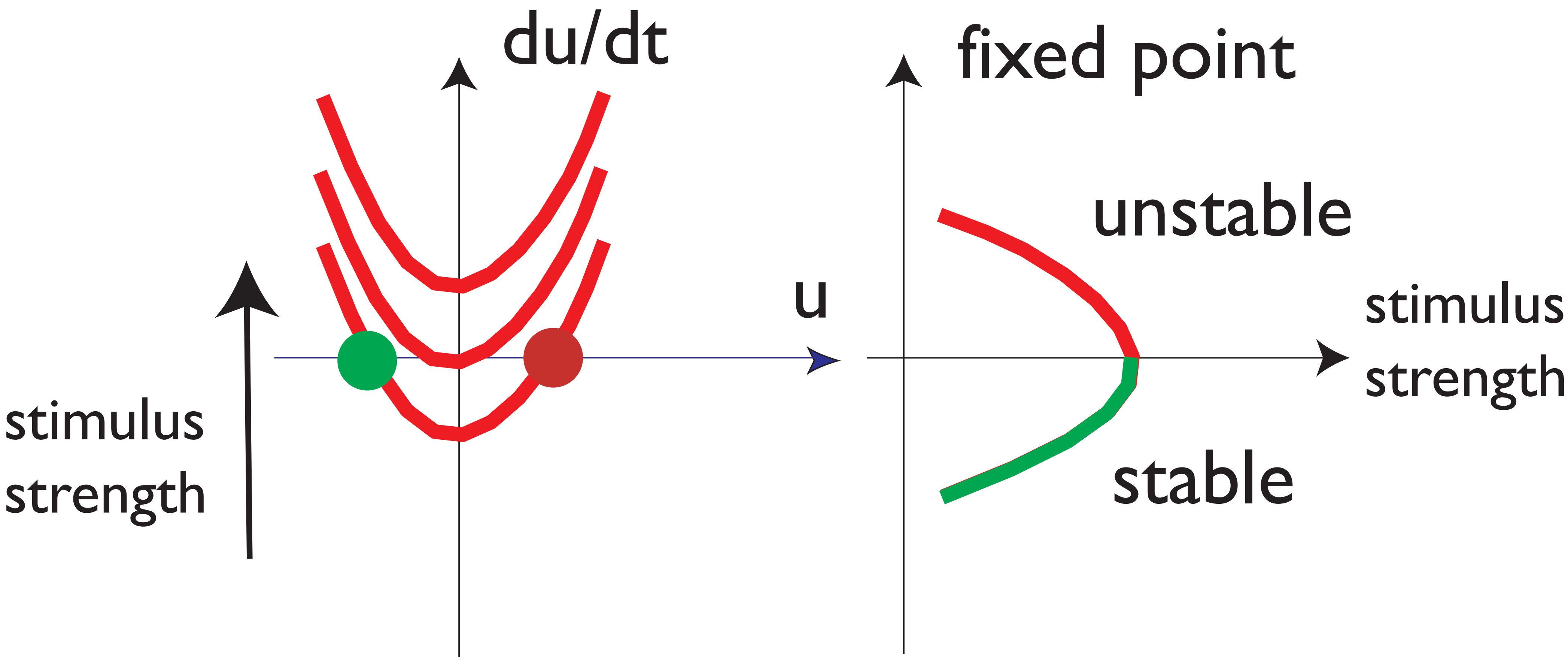
Neuronal dynamics with self-excitation

- with varying input strength system goes through two instabilities: the detection and the reverse detection instability



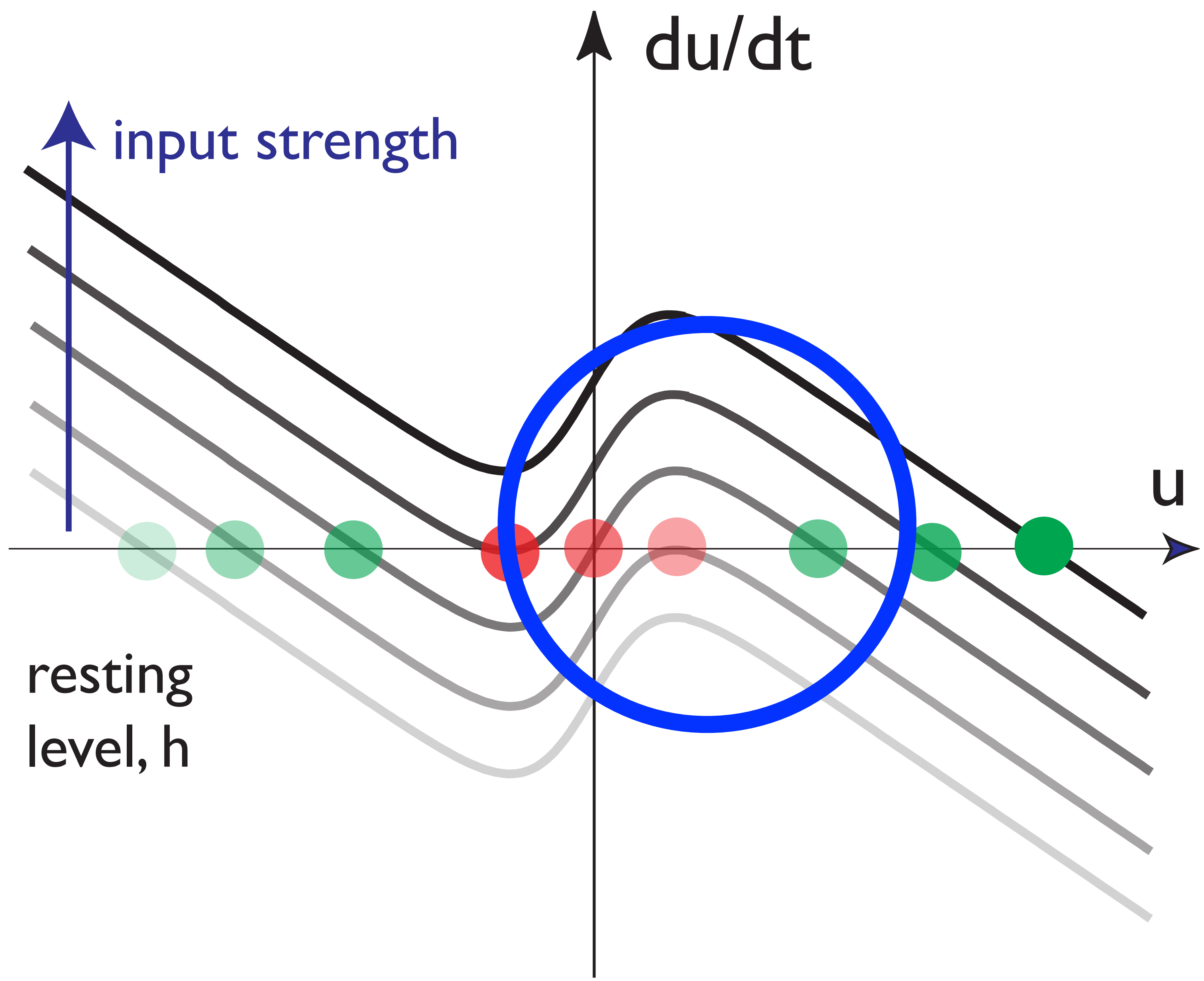
Neuronal dynamics with self-excitation

- detection instability



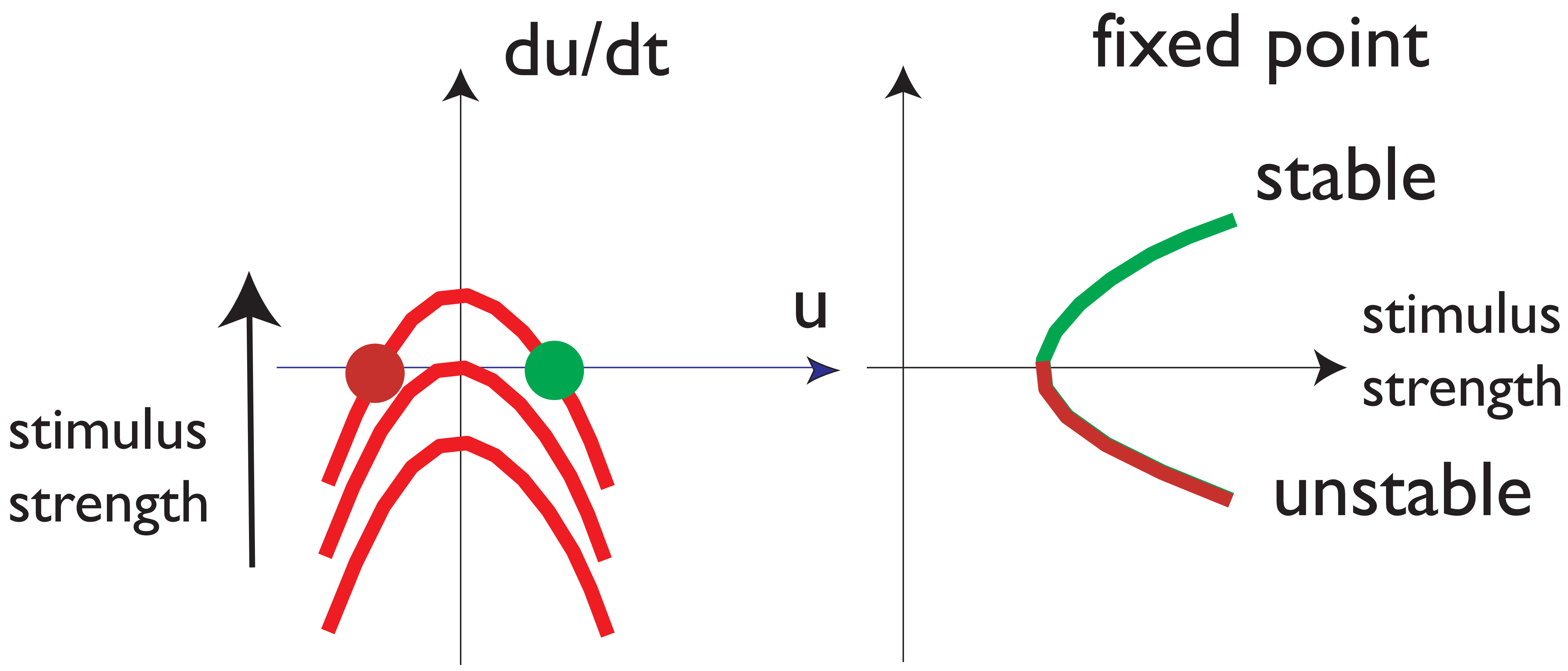
Neuronal dynamics with self-excitation

- with varying input strength system goes through two instabilities: the detection and the reverse detection instability

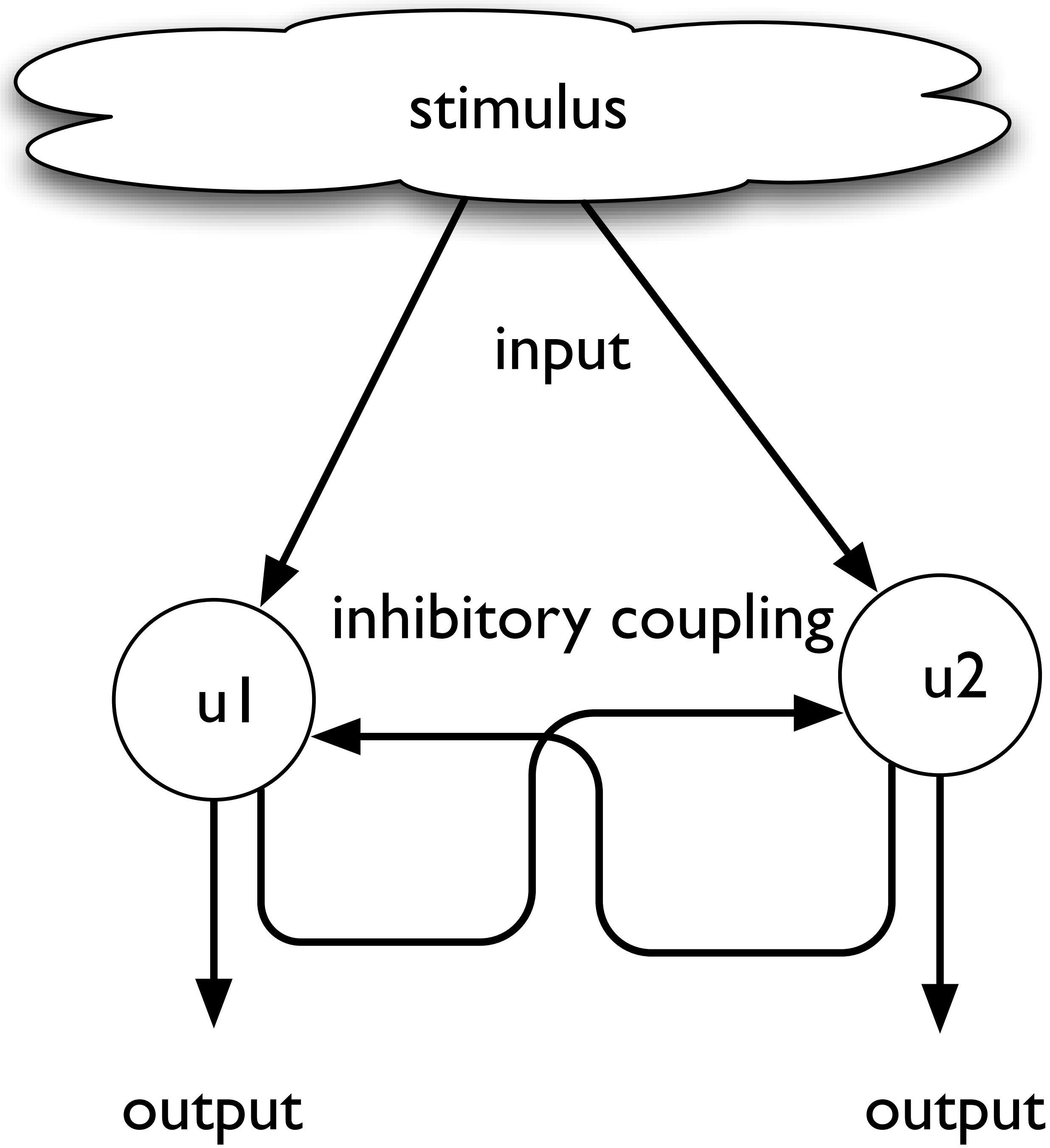


Neuronal dynamics with self-excitation

- reverse detection instability



Neuronal dynamics with competition



$$\tau \dot{u}_1(t) = -u_1(t) + h - \sigma(u_2(t)) + S_1$$
$$\tau \dot{u}_2(t) = -u_2(t) + h - \sigma(u_1(t)) + S_2$$

Neuronal dynamics with competition

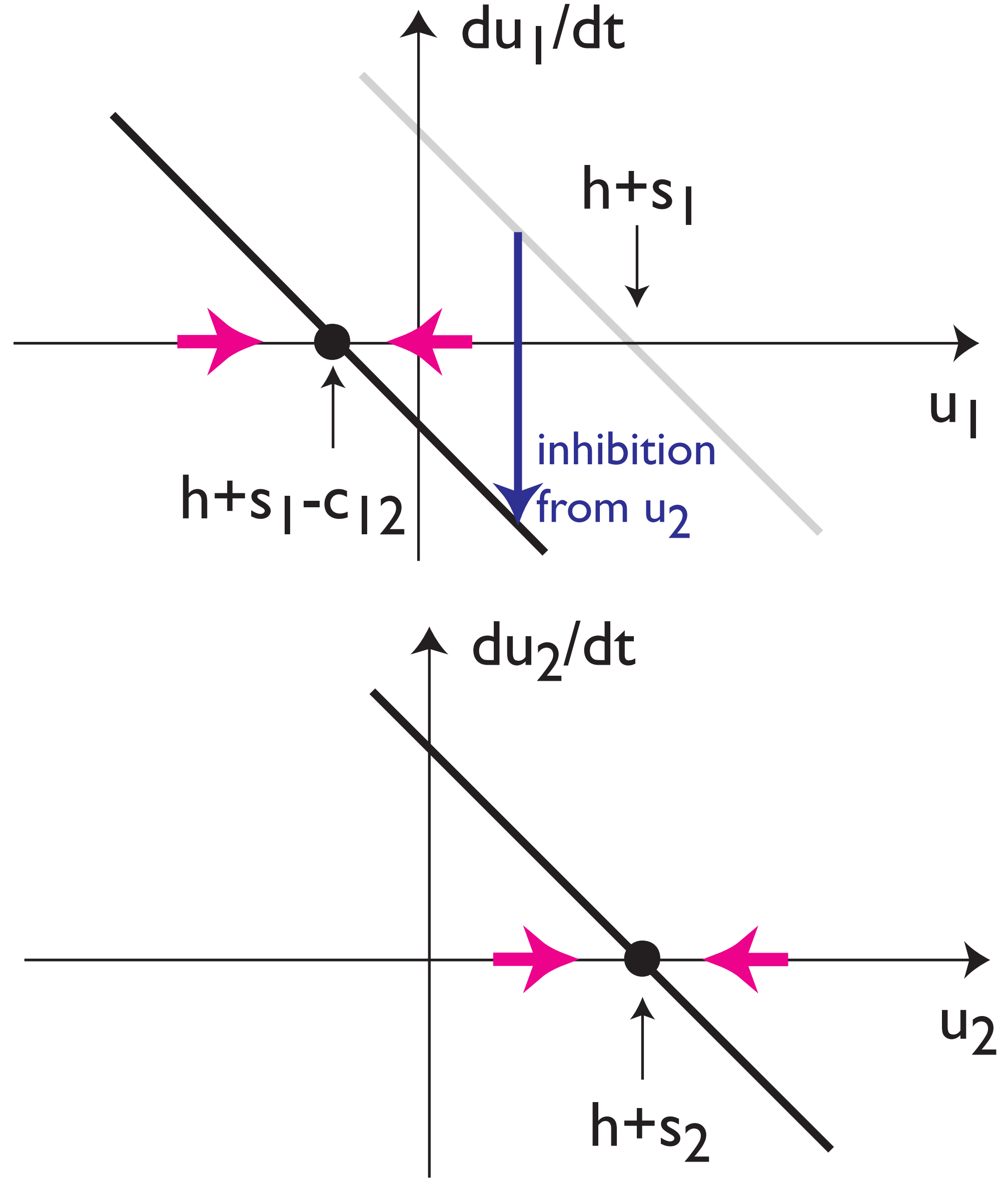
- the rate of change of activation at one site depends on the level of activation at the other site
- mutual inhibition

$$\begin{aligned} \tau \dot{u}_1(t) &= -u_1(t) + h - \sigma(u_2(t)) + S_1 \\ \tau \dot{u}_2(t) &= -u_2(t) + h - \sigma(u_1(t)) + S_2 \end{aligned}$$

sigmoidal nonlinearity

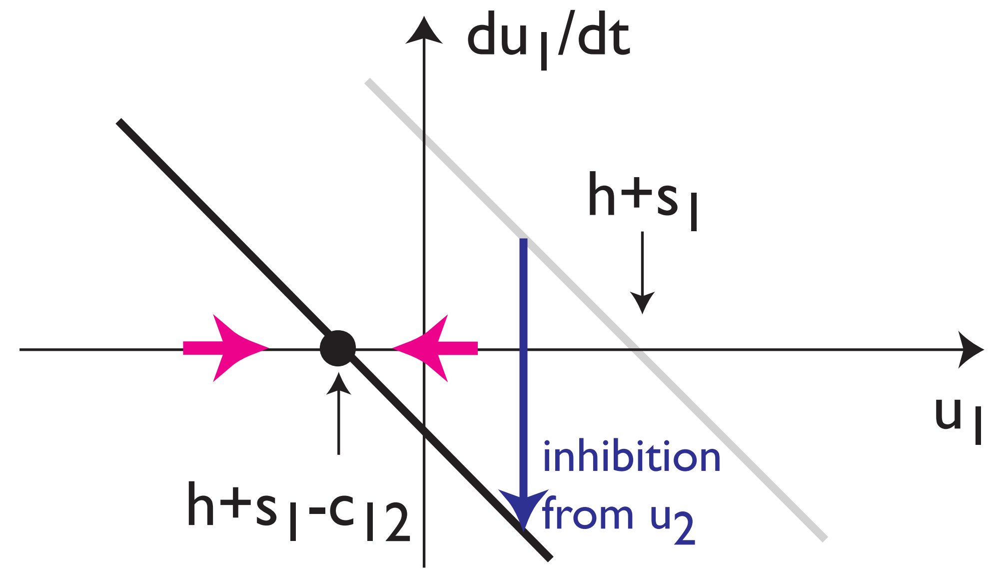
Neuronal dynamics with competition

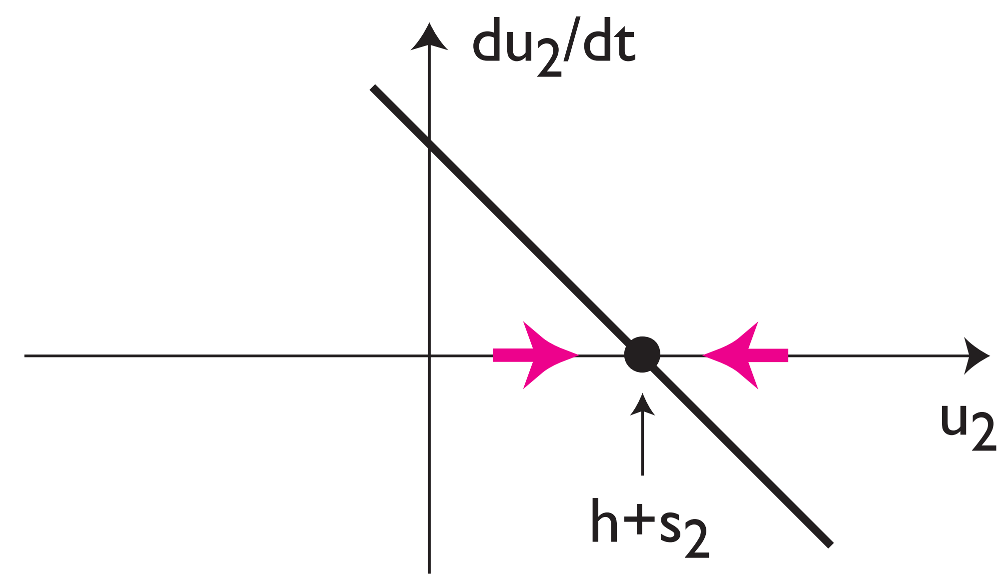
- to visualize, assume that u_2 has been activated by input to positive level
- => then u_1 is suppressed



Neuronal dynamics with competition

- why would u_2 be positive before u_1 is? E.g., it grew faster than u_1 because its inputs are stronger/inputs match better
- => input advantage translates into time advantage which translates into competitive advantage

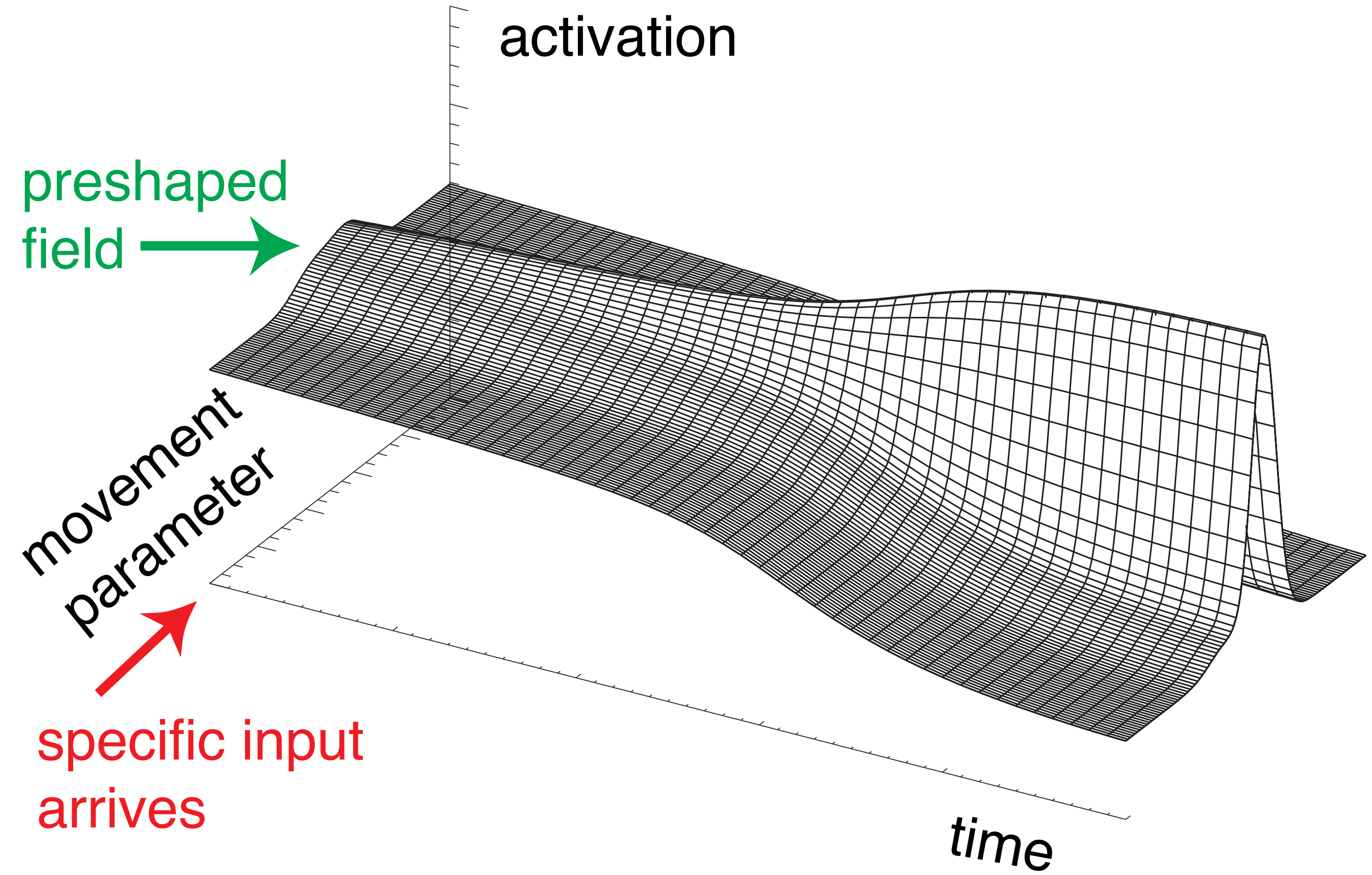


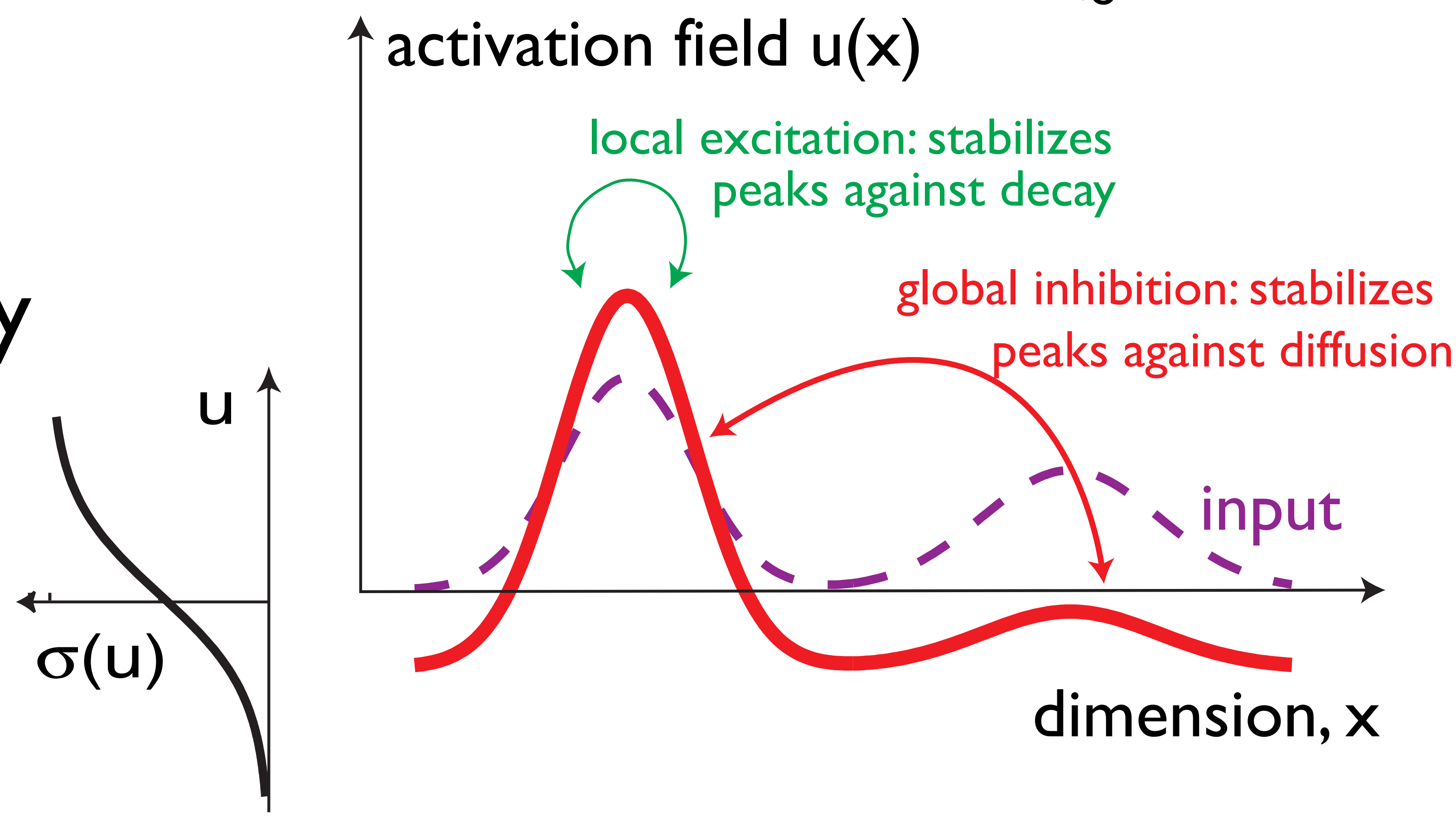


=> simulation

The dynamics activation fields

- field dynamics combines input
- with strong interaction:
  - local excitation
  - global inhibition
- => generates stability of peaks





Amari equation

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

where

- time scale is τ
- resting level is $h < 0$
- input is $S(x,t)$
- interaction kernel is

$$w(x - x') = w_i + w_e \exp\left[-\frac{(x - x')^2}{2\sigma_i^2}\right]$$

- sigmoidal nonlinearity is

$$\sigma(u) = \frac{1}{1 + \exp[-\beta(u - u_0)]}$$

=> simulations

Summary: DFT

- attractor states
  - input driven solution (sub-threshold)
  - self-stabilized solution (peak, supra-threshold)
- instabilities
  - detection instability (from localize input or boost)
  - reverse detection instability
  - selection instability
  - memory instability