Introduction to Neural Computation

Prof. Michale Fee
MIT BCS 9.40 — 2018

Lecture 19
Neural Integrators

Short-term vs long-term memory

- Long-term memory

  Can last a lifetime

  Large capacity—can hold many memories

  Mechanism: physical changes in neurons and synapses

- Short-term memory

  Lasts tens of seconds

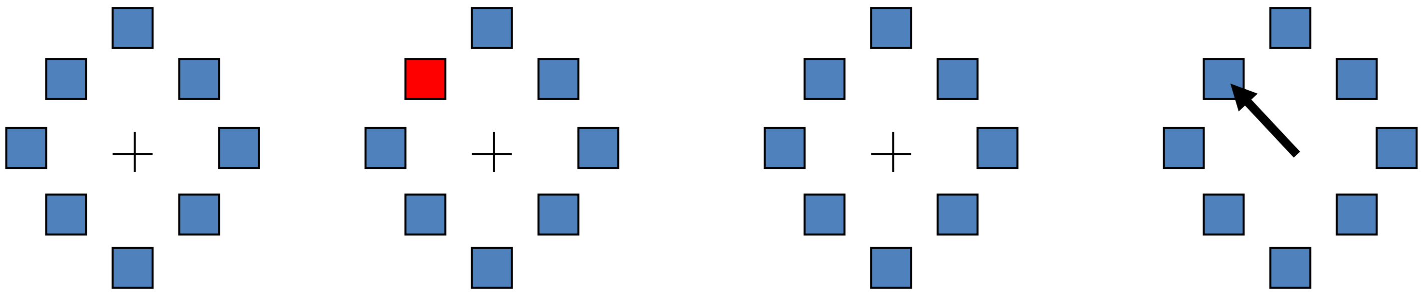
  Small capacity—only can hold a small number at any time

  Mechanism: persistent firing in a population of neurons

Short-term memory

- Persistent firing is the neural correlate of short-term memory

Delayed Saccade Task:



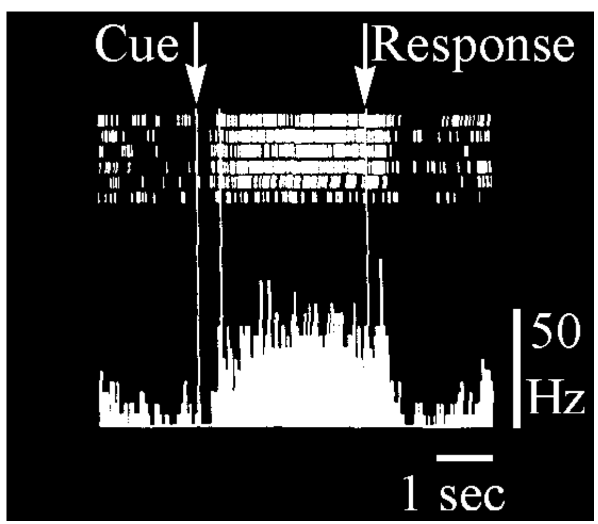
1) Fixate cross

2) Cue turns on

3) Delay period

4) Response:
--Saccade to memorized location

Neural activity in prefrontal cortex:

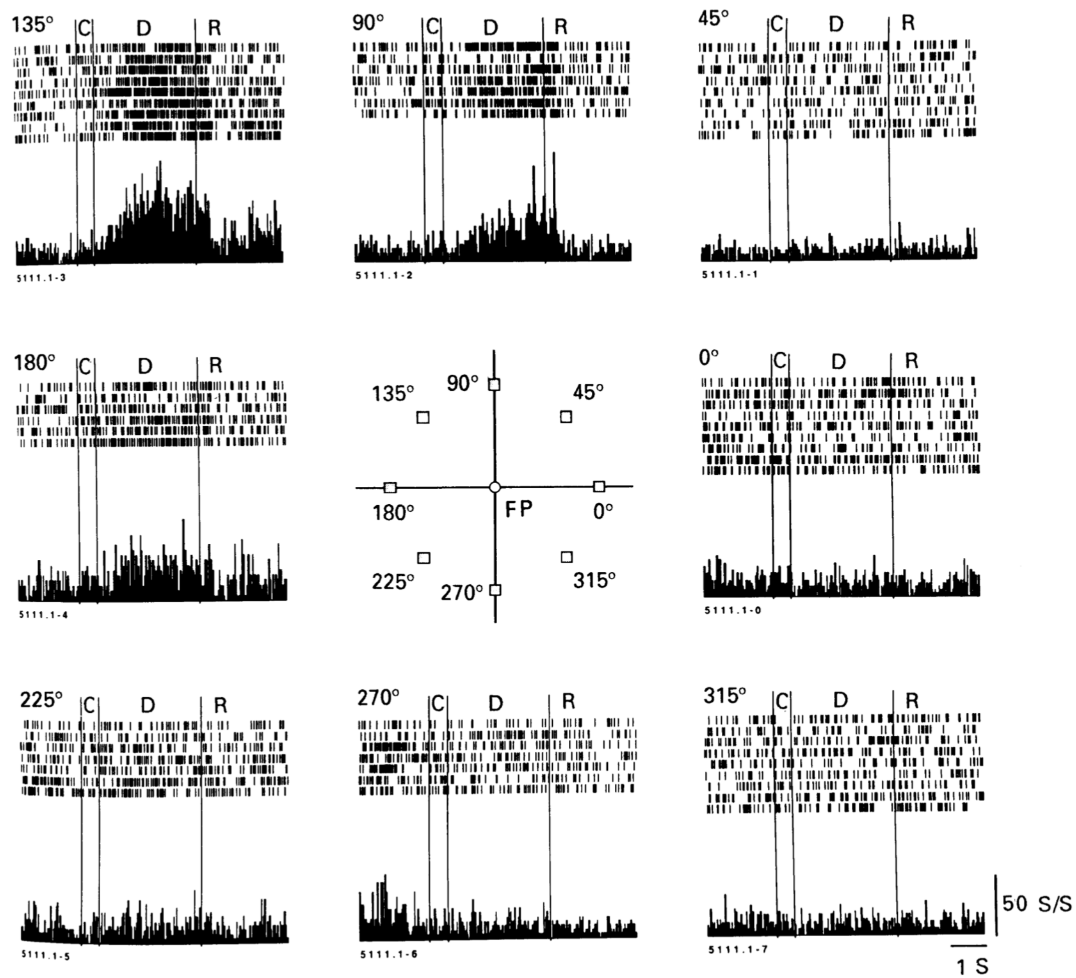


Funahashi, Goldman-Rakic (1991)

Short-term memory

- Delay activity is selective for remembered cue location

Single neuron response to different memorized target locations:

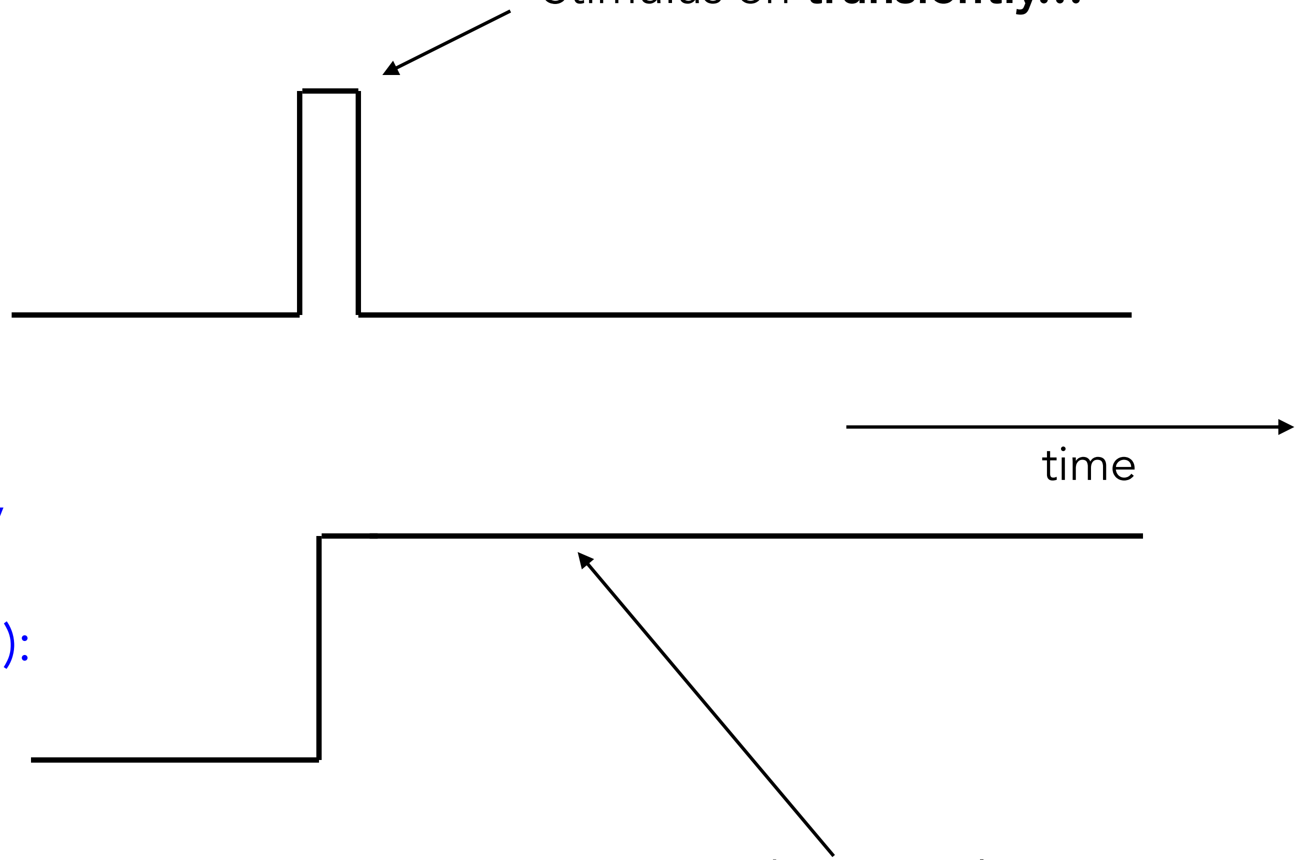


Short-term memory

- Persistent activity is the neural correlate of short-term memory

Stimulus on **transiently…**

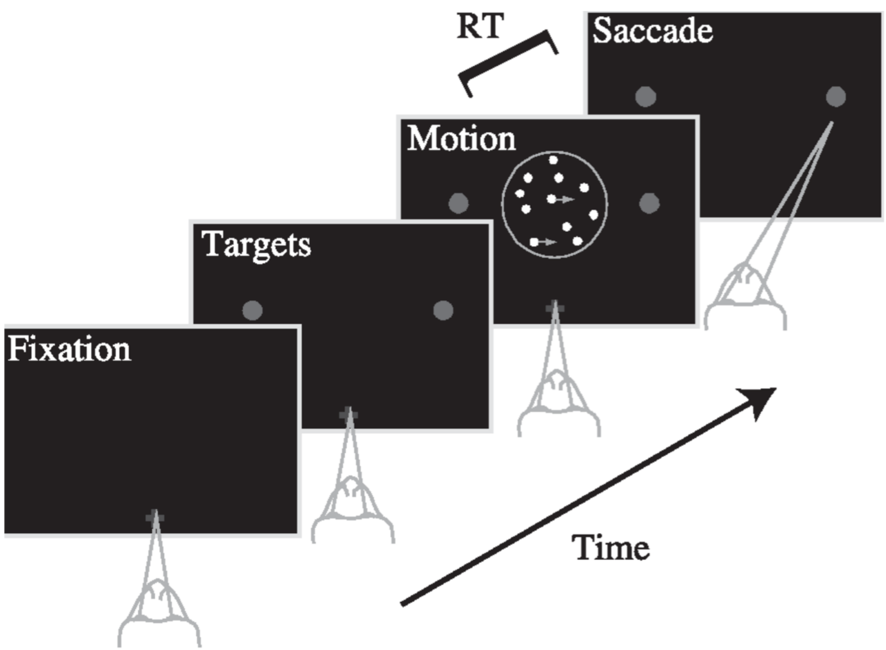
Stimulus:

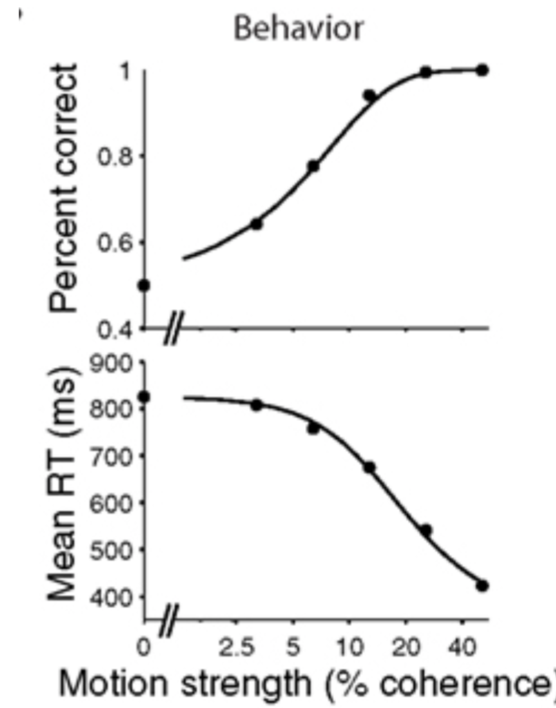


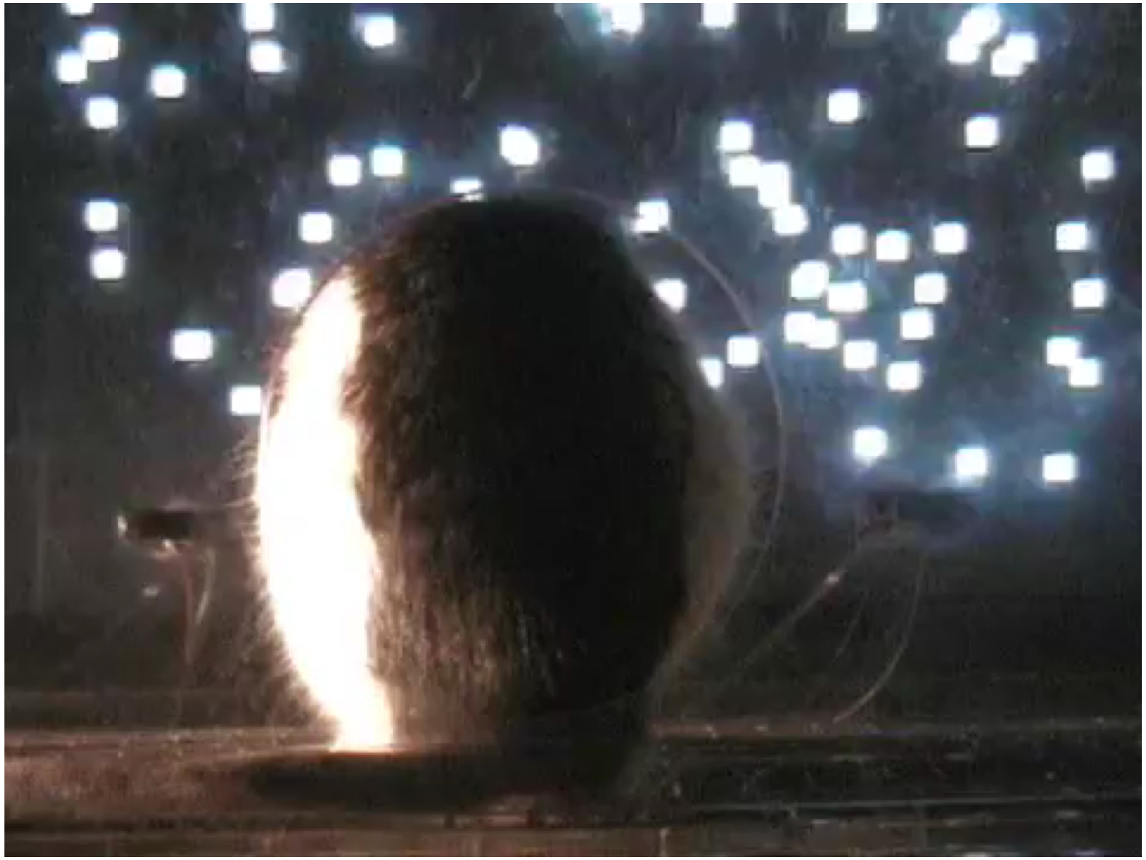
Neural activity (firing rate in spikes/second):

…but neural activity **persists** for ~10’s of seconds after stimulus turns off

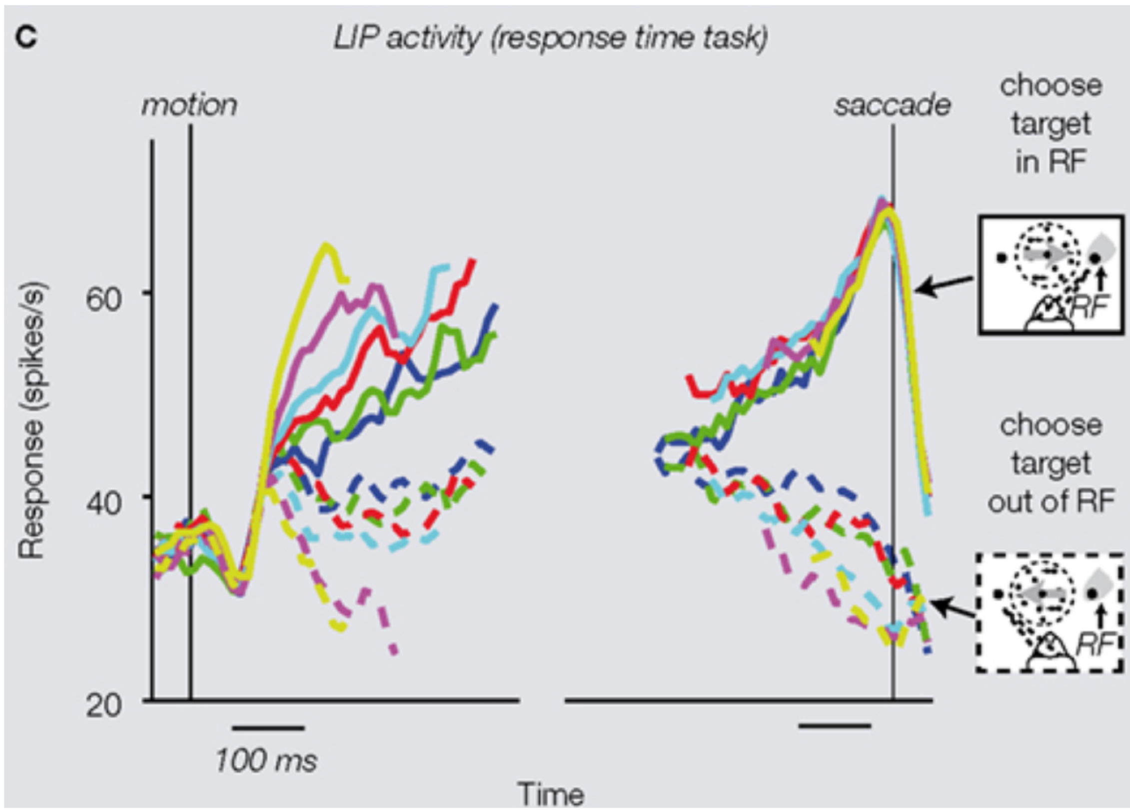
Evidence accumulation for decision-making







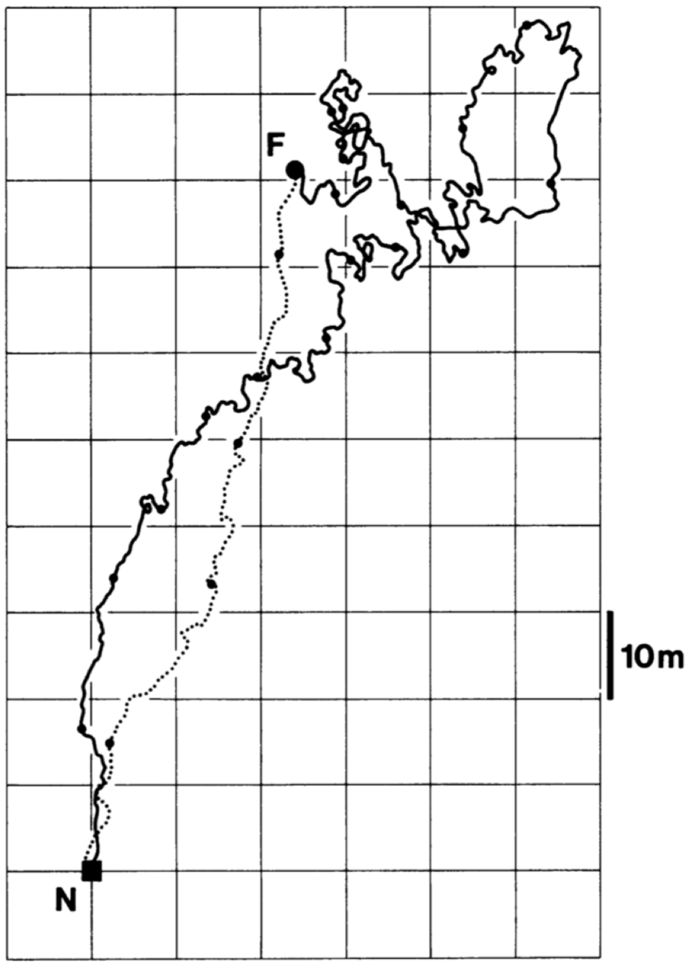
Evidence accumulation for decision-making

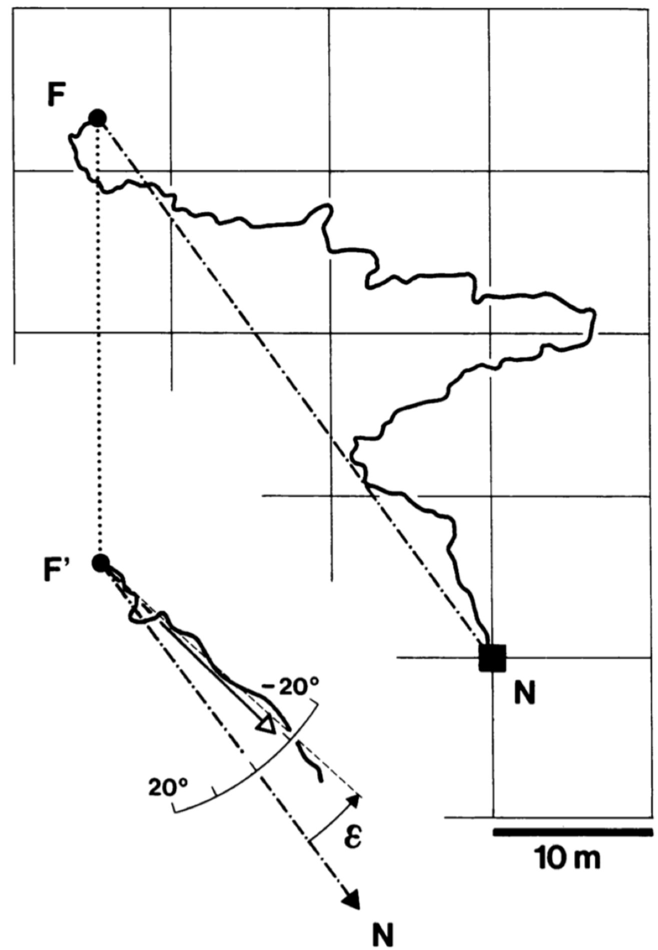


Other Integrators

desert ants

-Navigation by path integration:

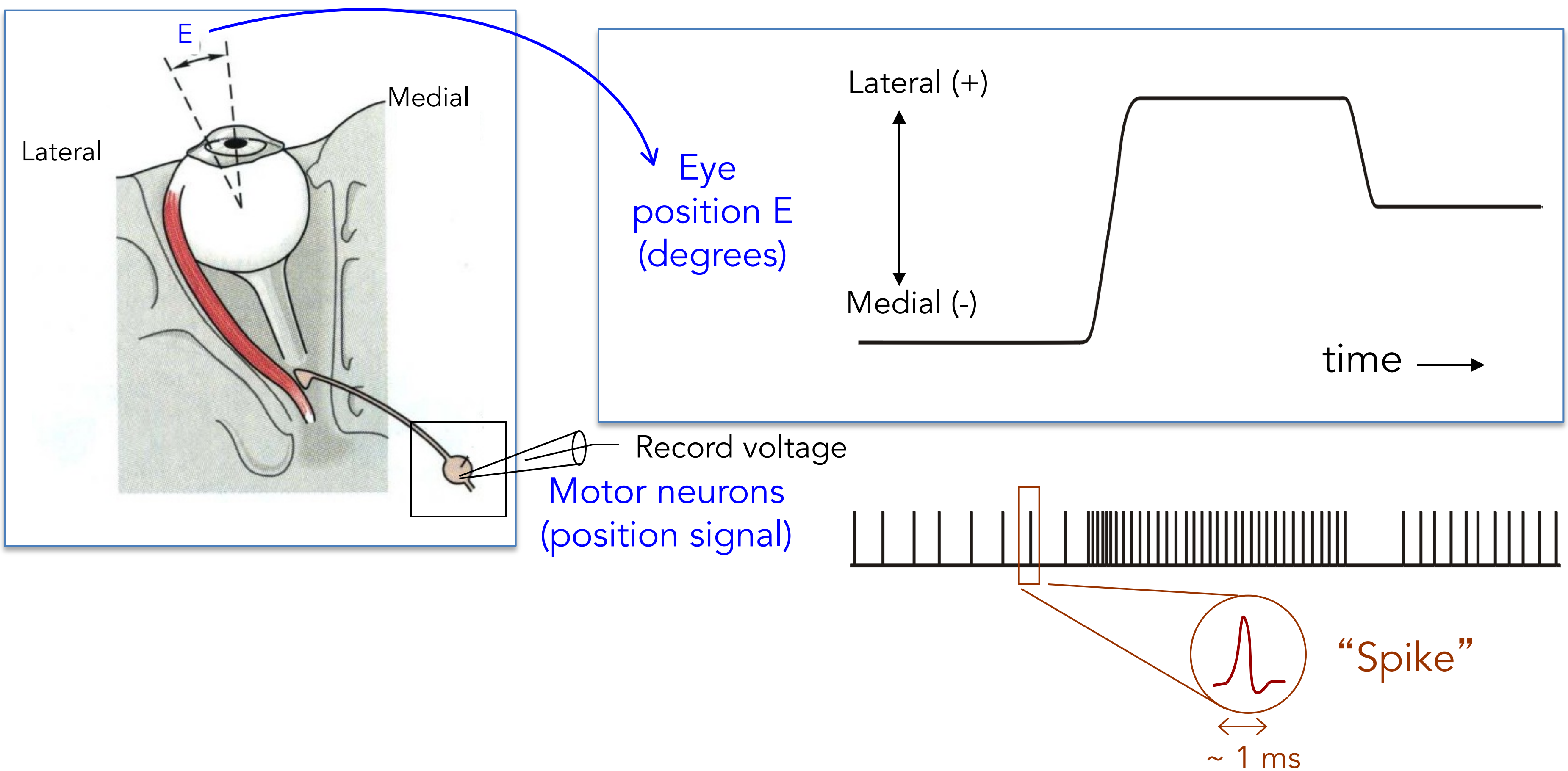




(Müller & Wehner, 1988)

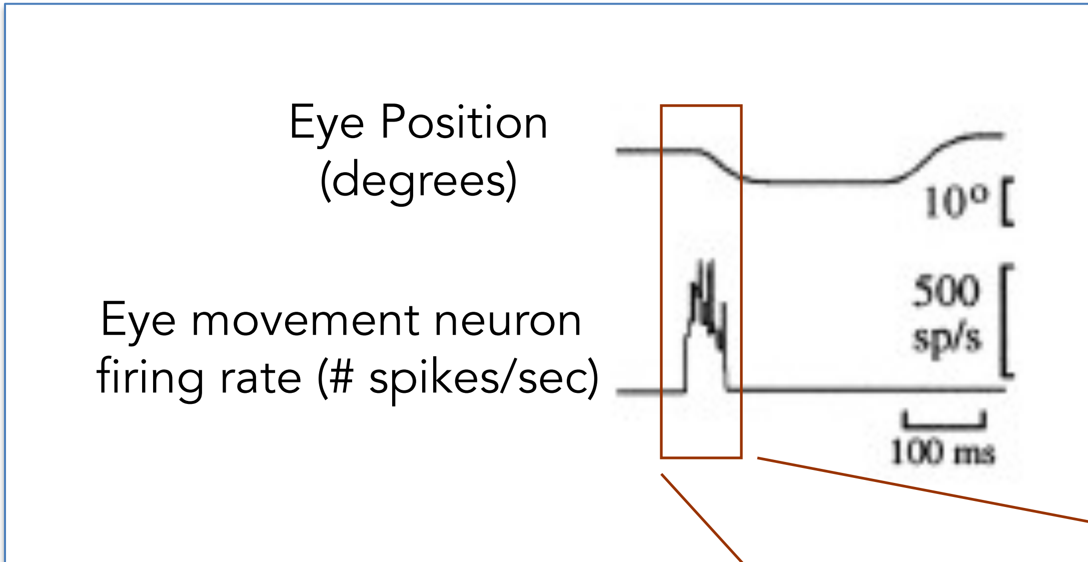
Short-term memory in the eye-movement system

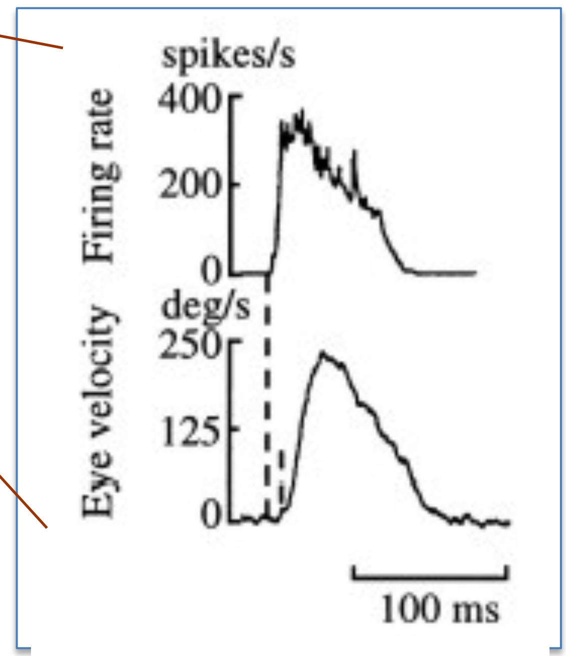
The eye-movement system



Saccade burst generator neurons

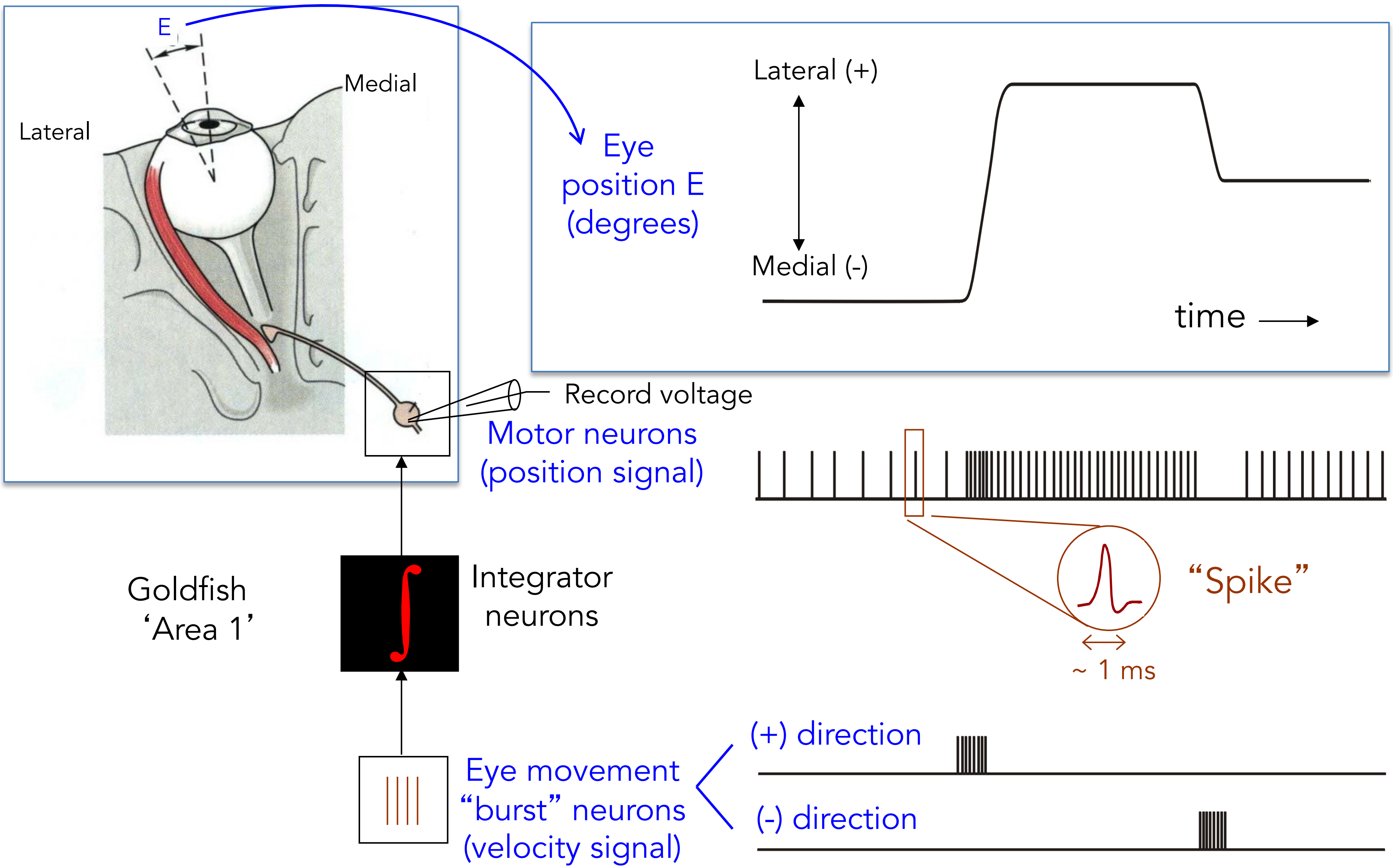
(adapted from Yoshida et al., 1982)



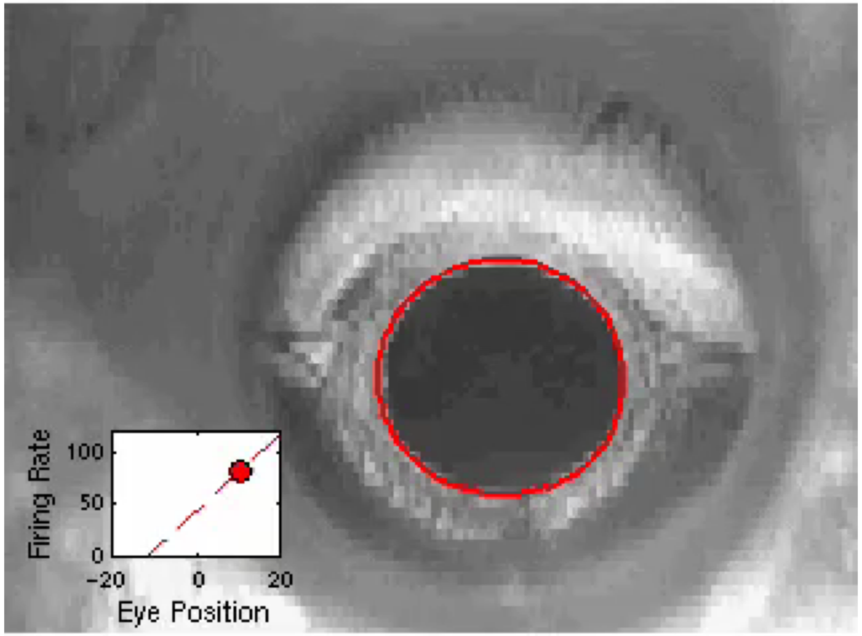


Burst neurons code for eye velocity

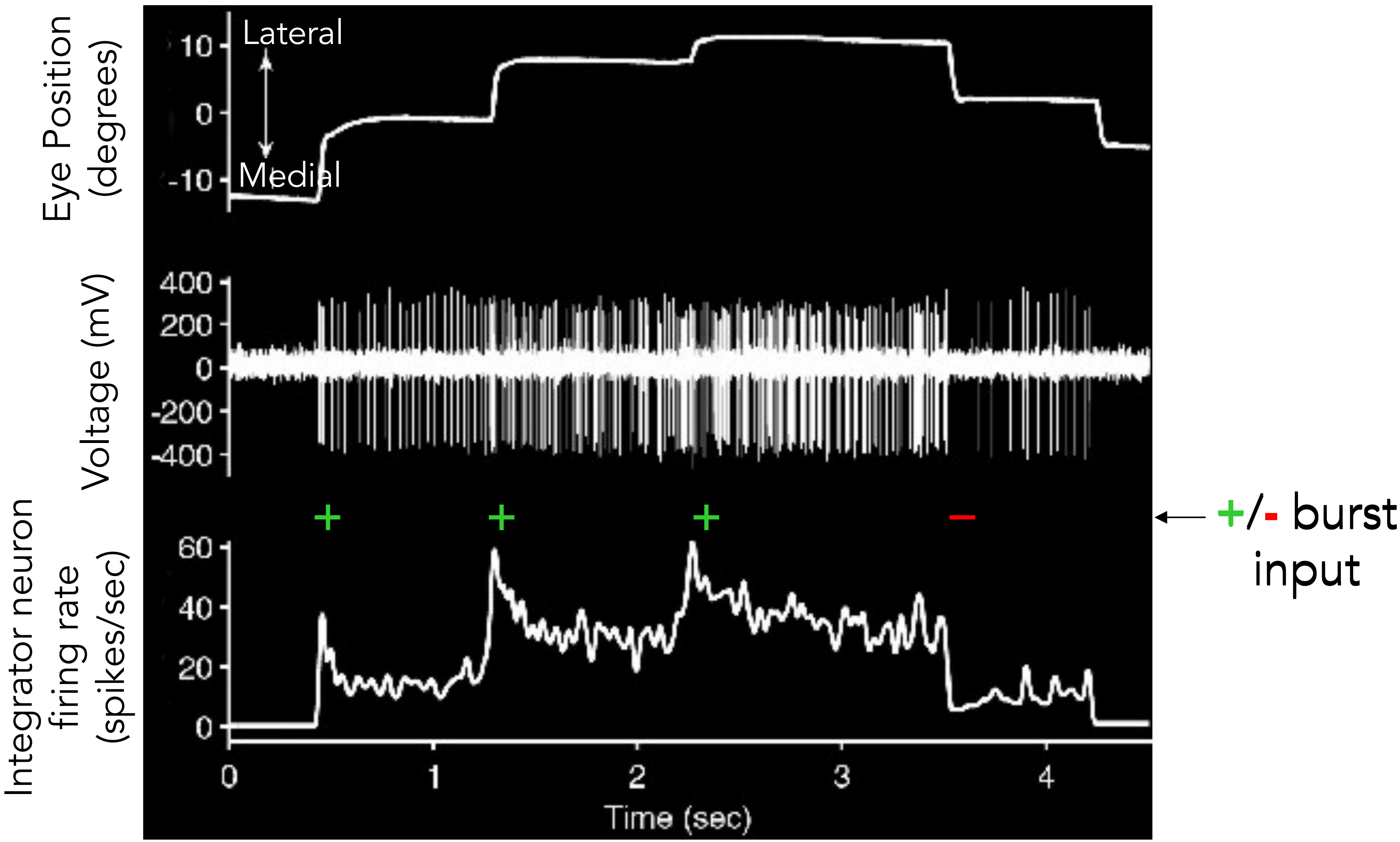
The eye-movement system



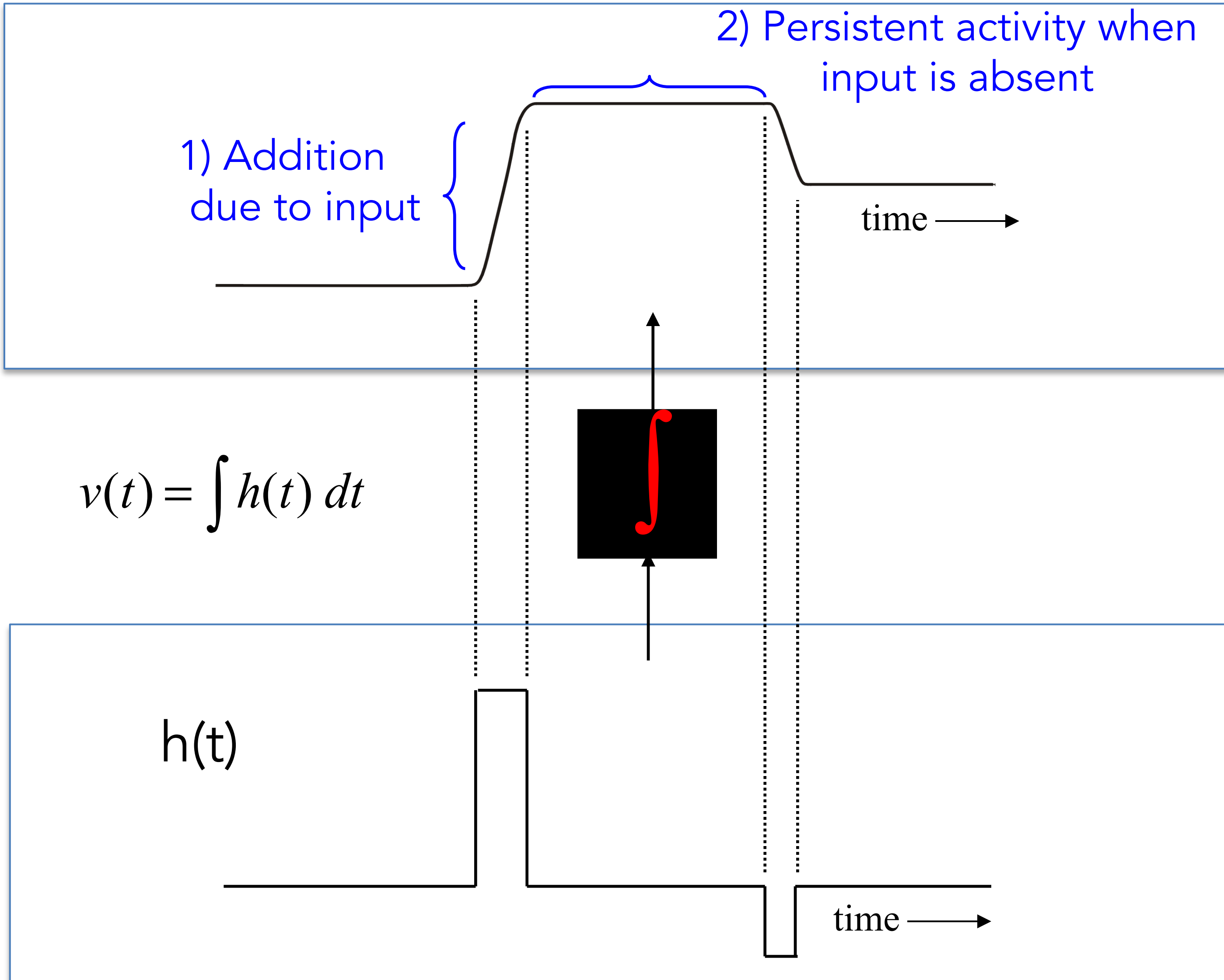
Integrator neurons



Integrator neuron carry an eye-position signal

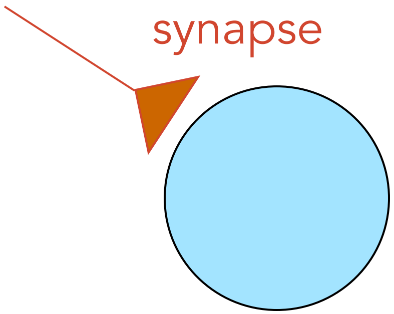


How neurons integrate



Basic model of a neuron

synapse



= firing rate of neuron

$$\tau_n \frac{dv}{dt} = -v + Input$$

change in firing rate over time

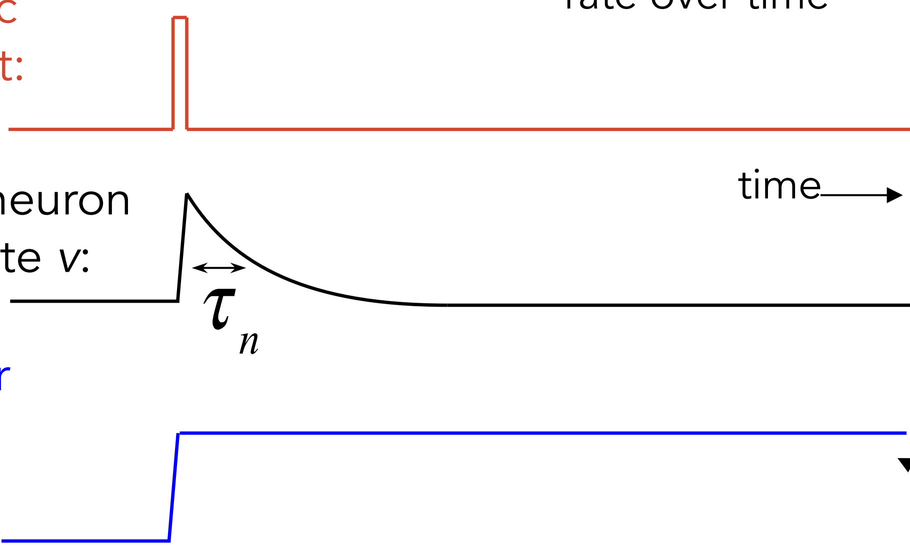
intrinsic decay

synaptic input

Synaptic input:

Single neuron firing rate *v*:

But, motor neuron firing rate:

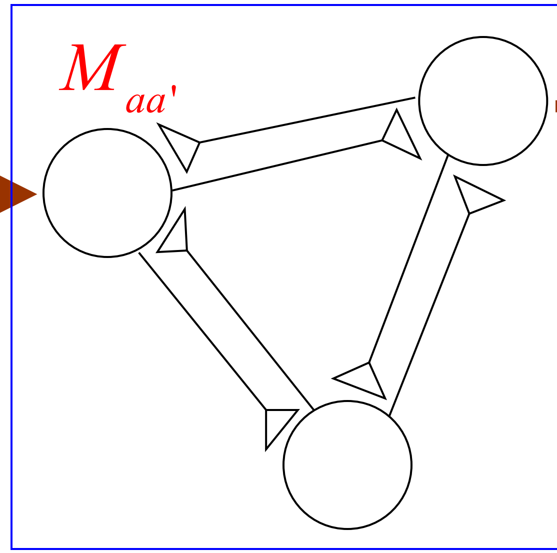


$$\frac{dv}{dt} = 0$$ when *Input* = 0

Network mechanism of persistence

Eye movement command input (burst)

Integrator circuit



Output to eye motor neurons

$M_{aa'}$ = input to neuron a from neuron a′

$$\tau_n \frac{dv_a}{dt} = \cancel{-v_a + \sum_{a'} M_{aa'} v_{a'}} + \mathit{burst\ input}$$

~~Leak~~ ~~Feedback from other neurons~~

If Feedback balances Leak

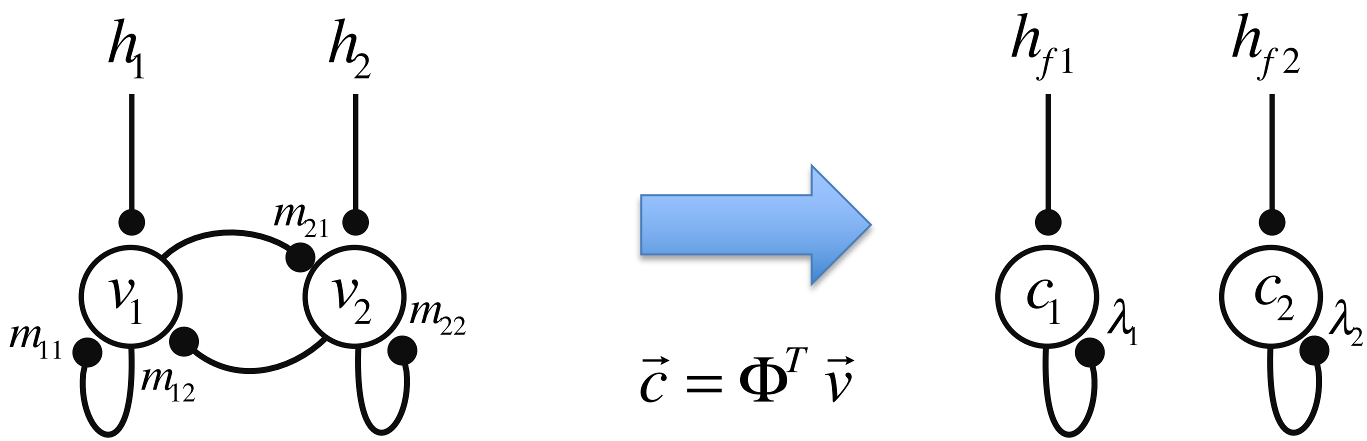
$$r = \frac{1}{\tau_{neuron}} \int (\mathit{burst\ input})\, dt$$

Recurrent networks

- We saw how the behavior of a recurrent network can be described if M is symmetric.

$$M = \Phi \Lambda \Phi^T$$

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \Phi = \left[\, \hat{f}_1 \,\middle|\, \hat{f}_2 \,\right]$$



Network mechanism of persistence

Eigenvectors:

- Most have eigenvalue << 1: rapid exponential decay after burst terminates
- One has eigenvalue ≈ 1:

Equation for component along this eigenvector:

$$\Rightarrow \quad \tau_n \frac{dc_1}{dt} = -c_1 + \lambda_1 c_1 + burst\ input$$

Between bursts

$$\frac{dc_1}{dt} = \left(\frac{\lambda_1 - 1}{\tau_n}\right) c_1$$

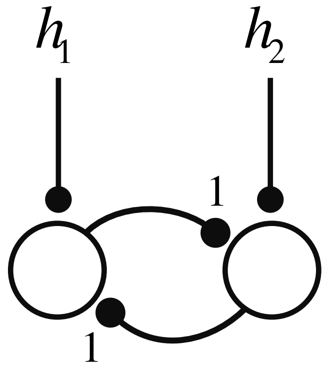
If $\lambda_1 = 1$ $\longrightarrow$ Perfect integrator!

feedback balances leak

$$c(t) = \frac{1}{\tau_{neuron}} \int (burst\ input)\ dt$$

Integrating network

- Now let's look at a case where two output neurons are connected to each other by mutual excitation with synaptic strength of one.



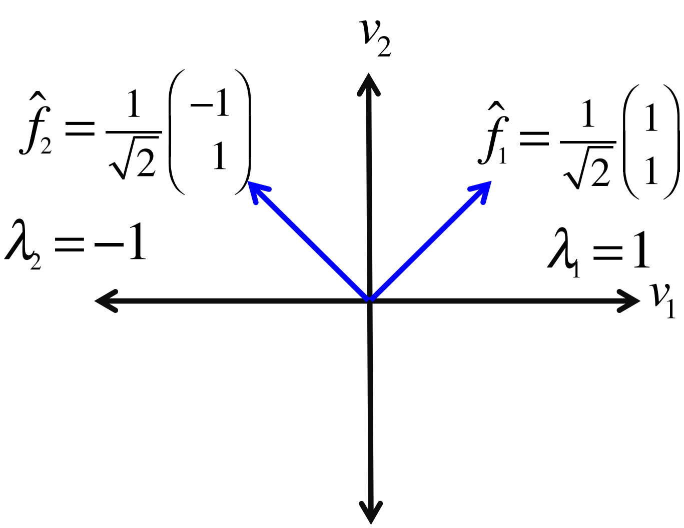
What is the weight matrix?

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$M\Phi = \Phi\Lambda$$

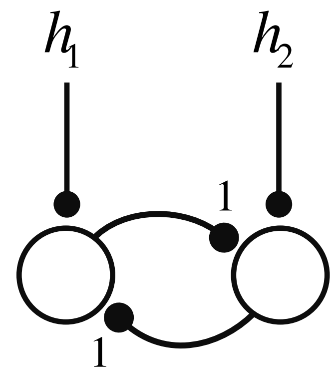
$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

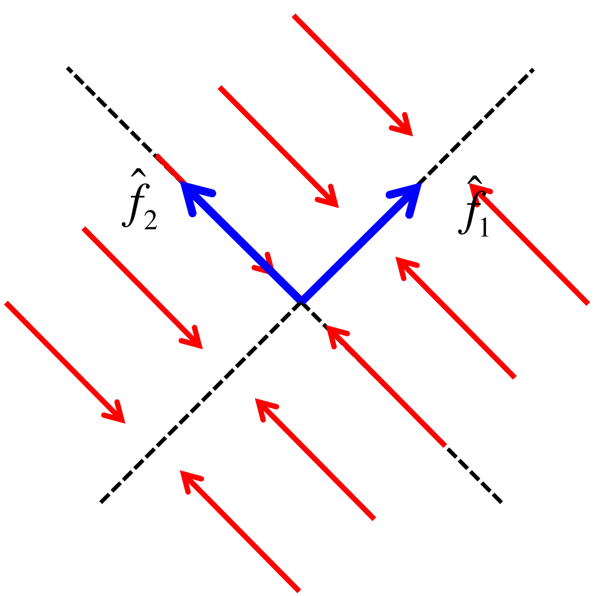
$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

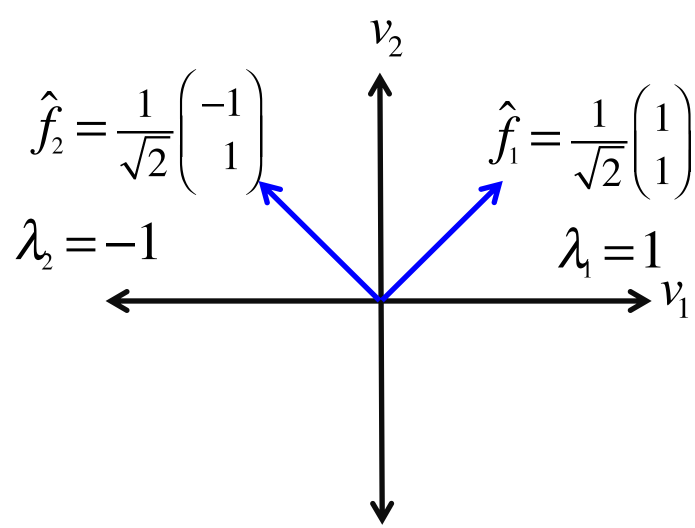


Recurrent networks

- If the input is parallel to the eigenvectors, then only one mode is excited.

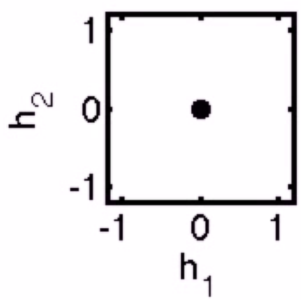


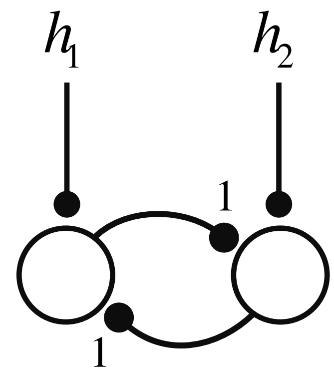


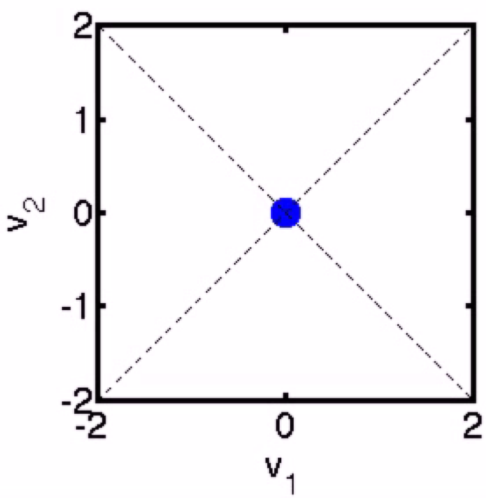


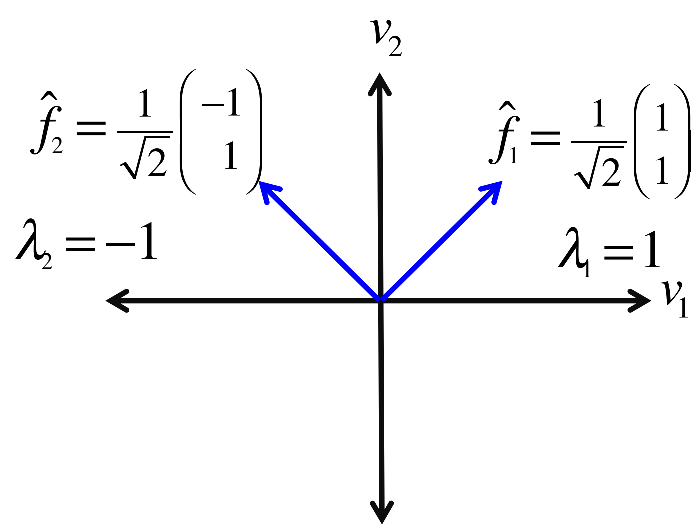
Recurrent networks

- If the input is parallel to the eigenvectors, then only one mode is excited.



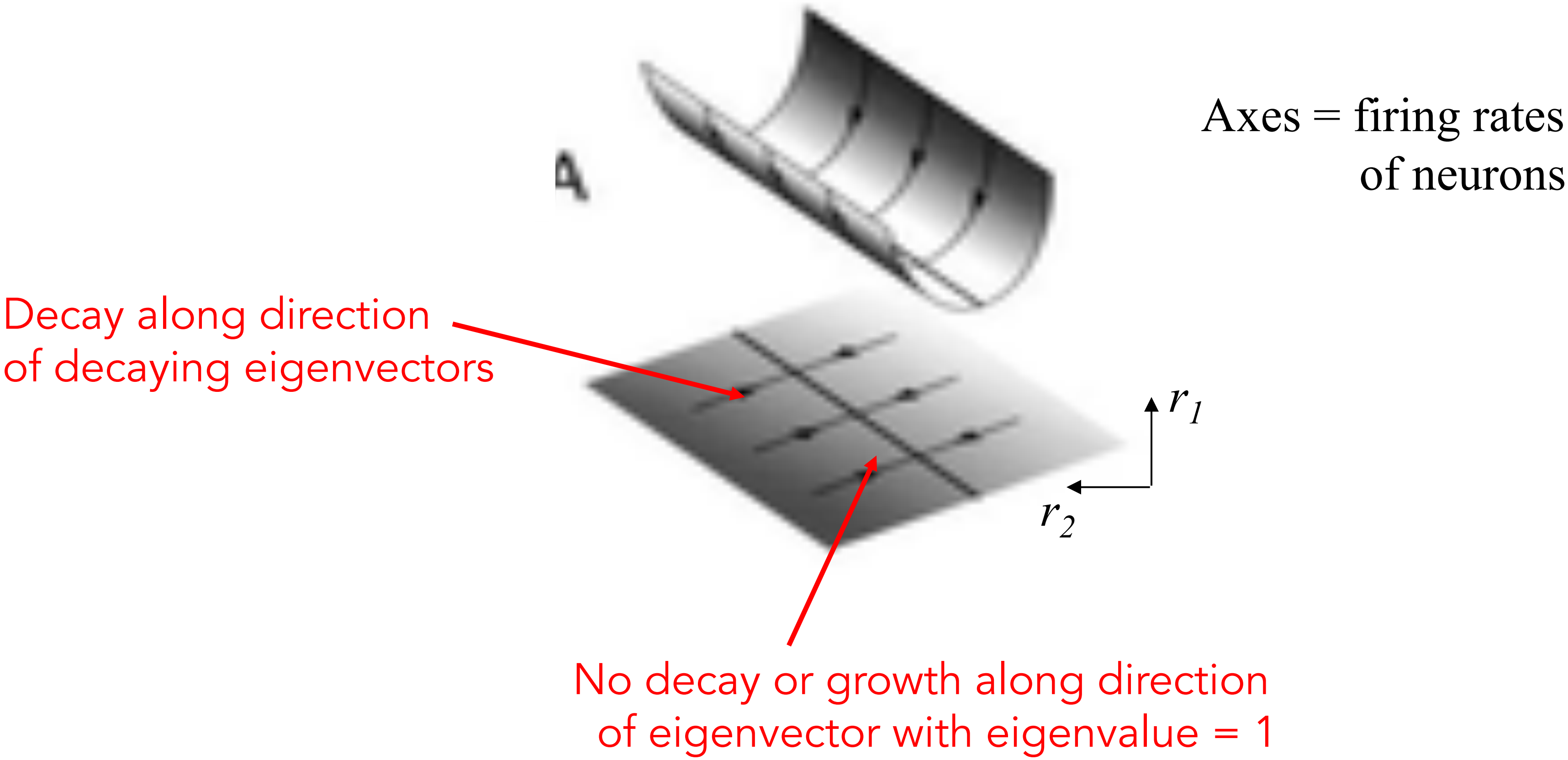






Geometric interpretation

- Line attractor picture of neural integrator

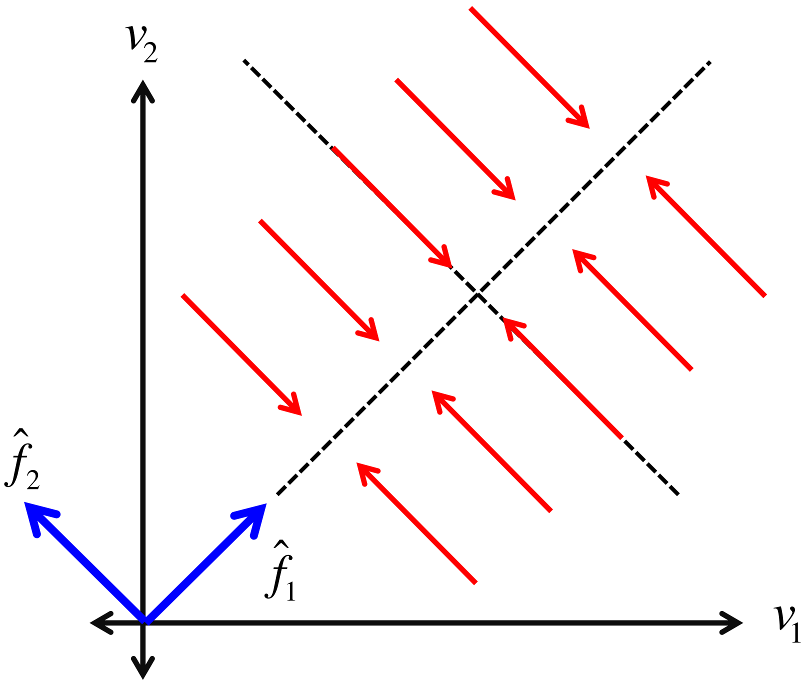


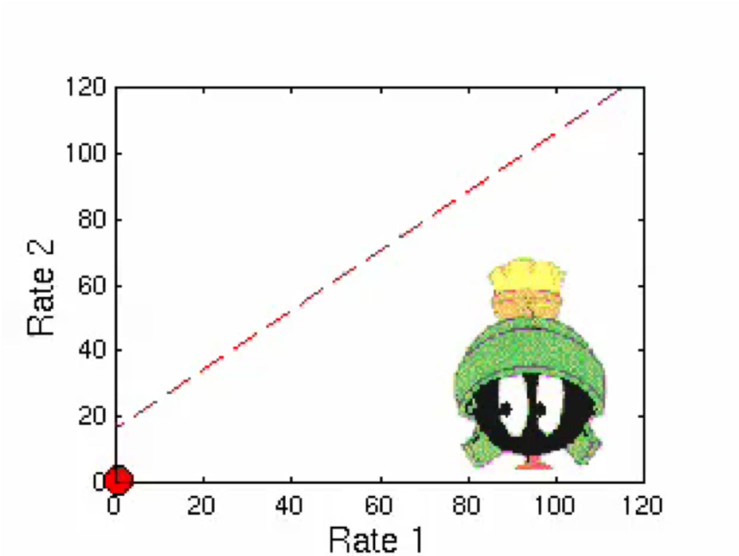
"Line Attractor" or "Line of Fixed Points"

Geometric interpretation

- Line attractor picture of neural integrator

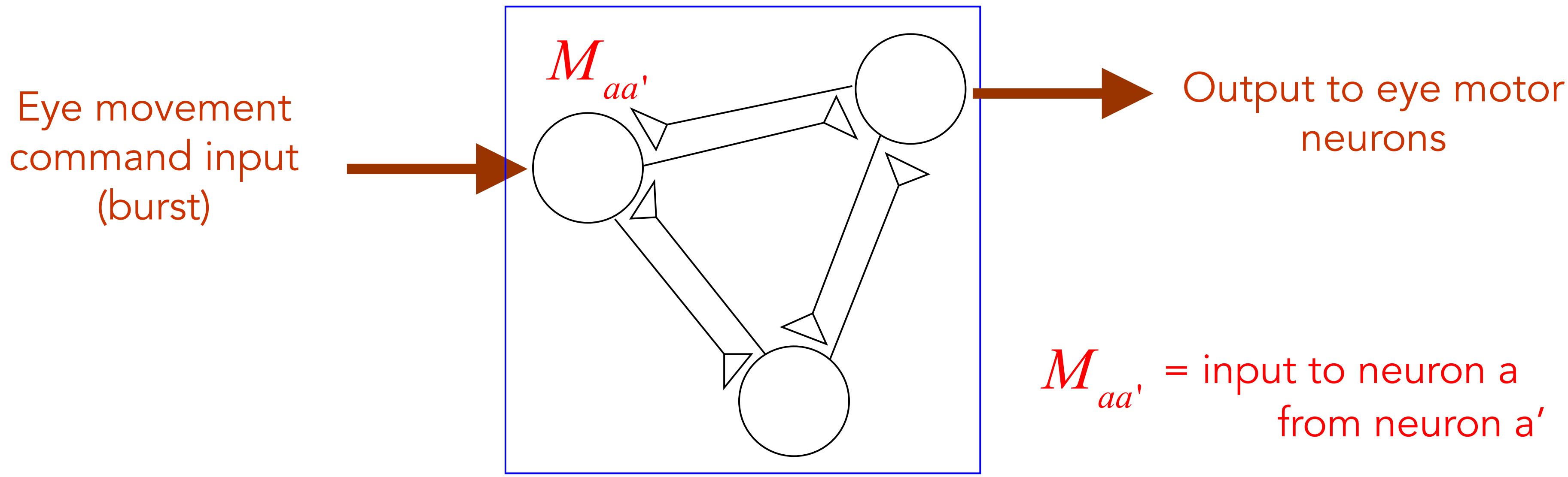
[Geometrical picture of line attractor]

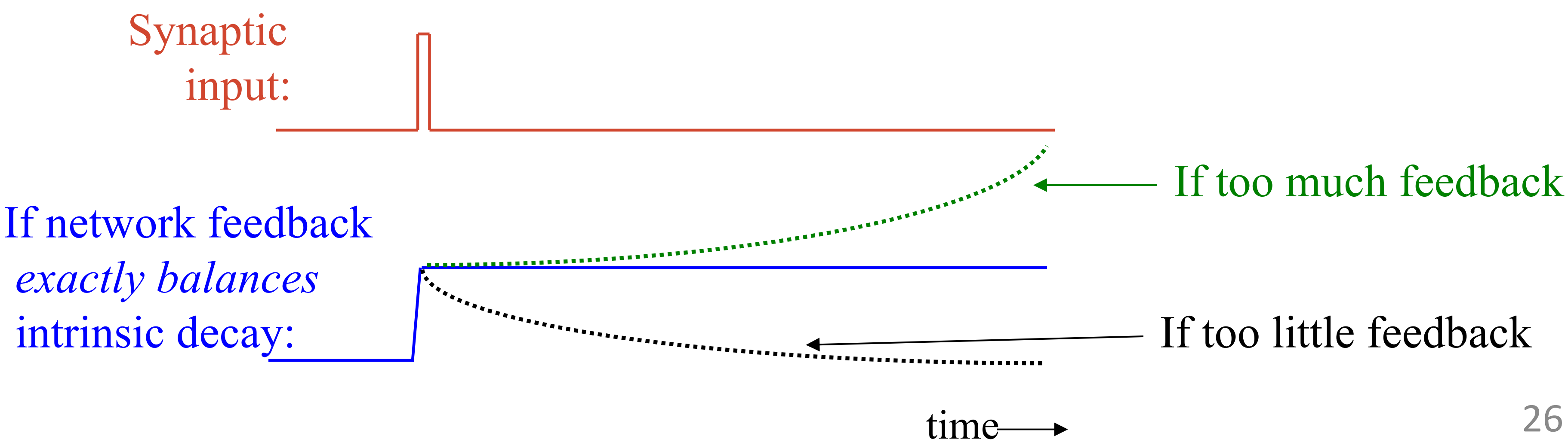




Perfect, leaky, and unstable integrators

- Network requires precise tuning of feedback strength



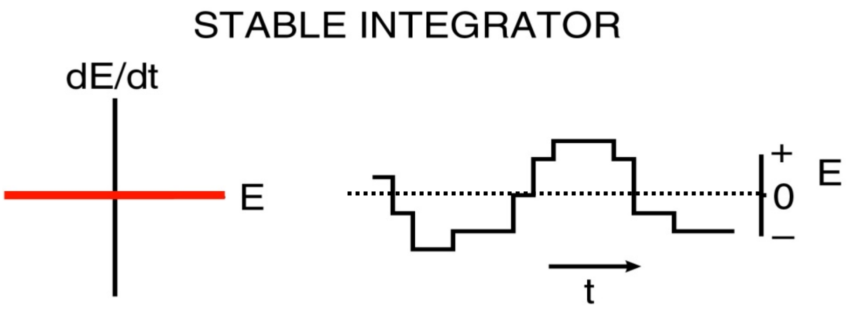


Perfect, leaky, and unstable integrators

Between bursts: $$\boxed{\frac{dc}{dt} = kc, \quad \text{where} \quad k = \frac{\lambda - 1}{\tau_n}}$$

$\lambda = 1$: Perfect integrator

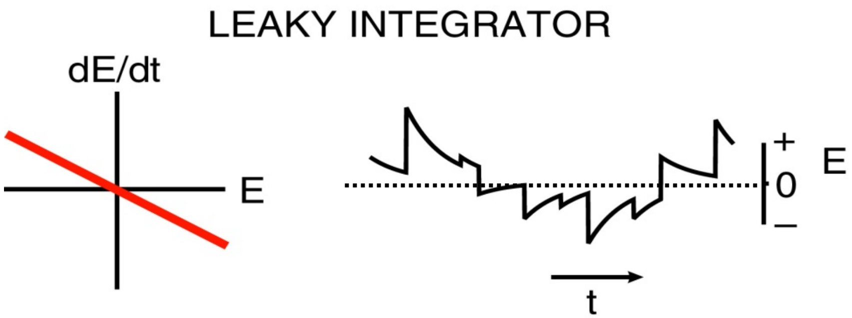
$$c(t) \sim constant$$



$\lambda < 1$: Leaky integrator

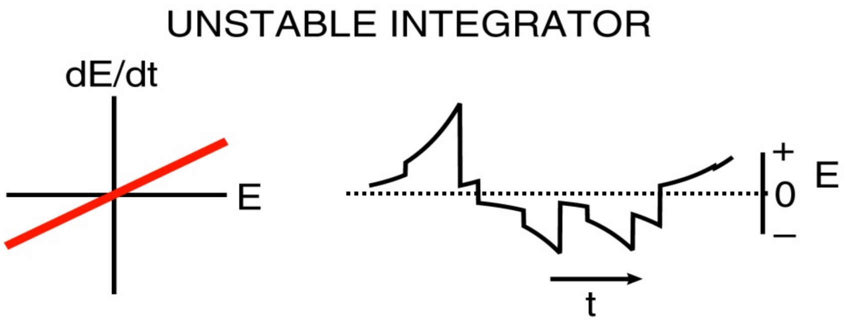
$$c(t) \sim e^{-|k|t}$$

$$\tau_{leak} = \frac{1}{|k|} = \frac{\tau_n}{1-\lambda}$$



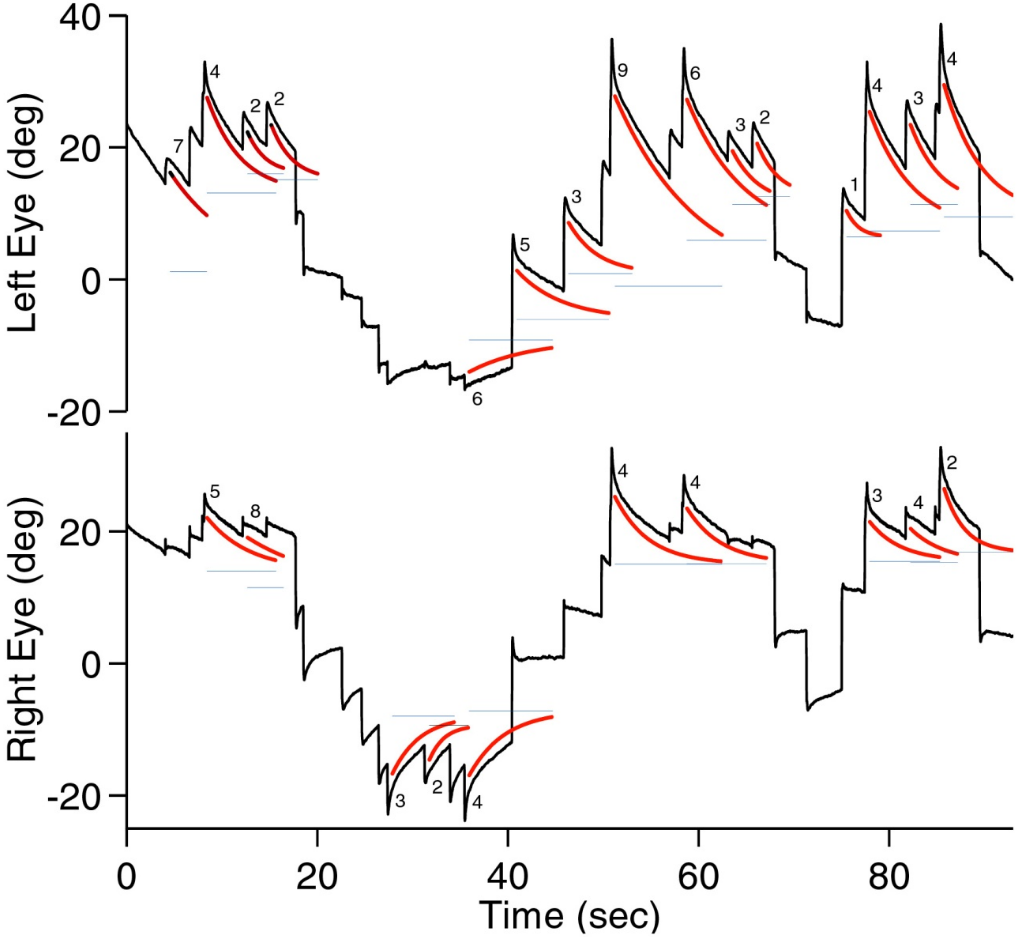
$\lambda > 1$: Unstable integrator

$$c(t) \sim e^{+|k|t}$$



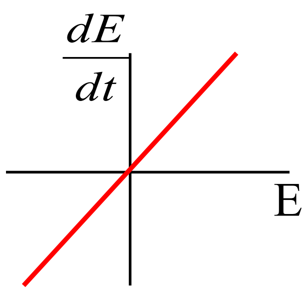
leaky integrator

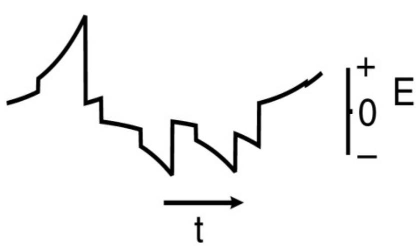
- Experiment: reduce feedback in the integrator circuit with local anesthetic



unstable integrator

- Human patient with unstable congenital nystagmus





Unstable neural integrator!

Robustness of the integrator

Integrator equation: $$\frac{dc}{dt} = \frac{(\lambda - 1)}{\tau_n} c \; + \; burst\ input$$

Experimental values:

Single isolated neuron: $\tau_n \approx 10 - 100 \text{ ms}$

Integrator circuit: $\tau_{network} = \dfrac{\tau_n}{|1-\lambda|} \approx 30 \text{ sec}$

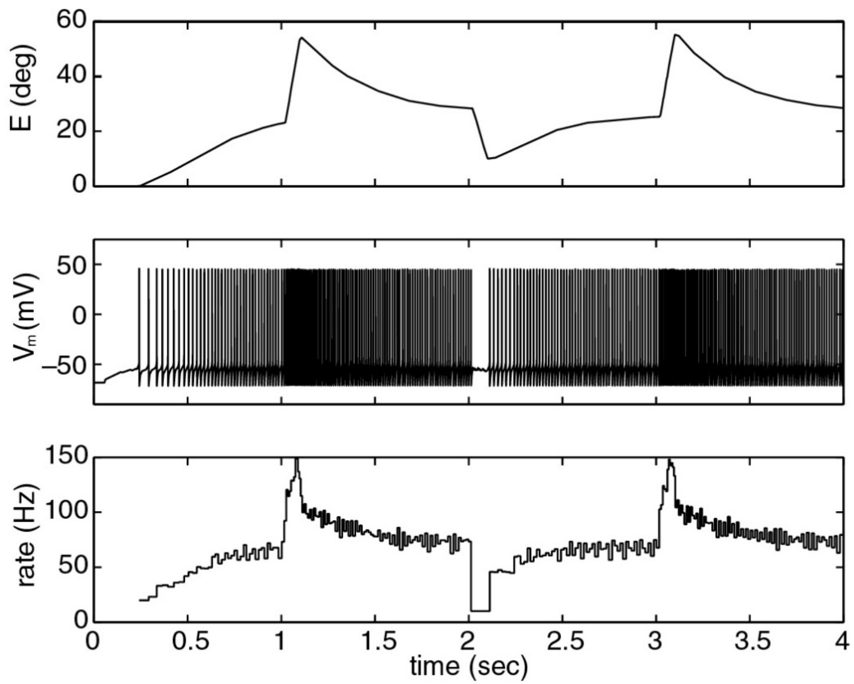
Synaptic feedback λ must be tuned to accuracy of:

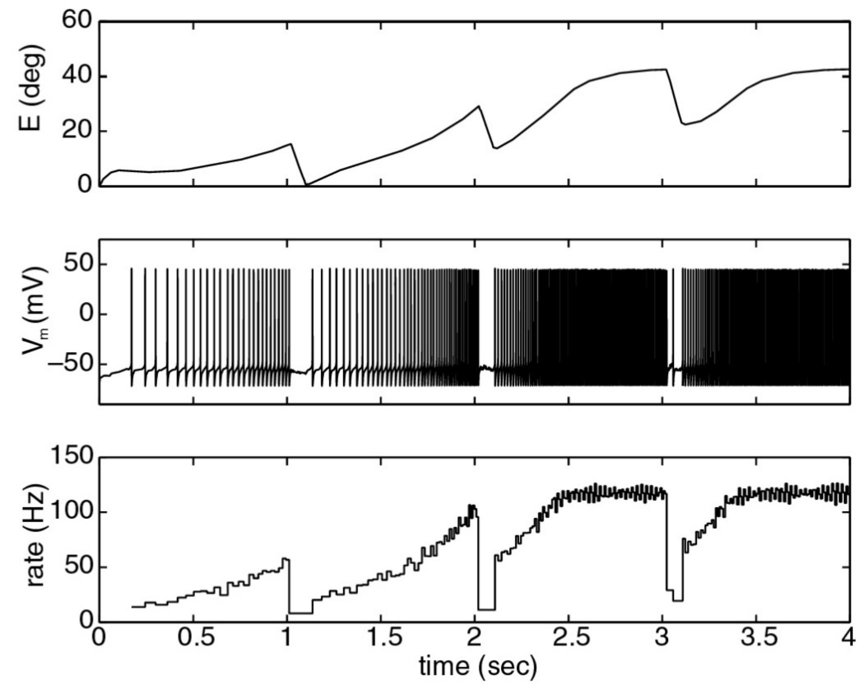
$$|1-\lambda| = \frac{\tau_n}{\tau_{network}} \approx 0.3\%$$

Robustness of the integrator

- Results with spiking network model

(Seung et al., 2000)





Leaky integrator (synaptic weights decreased 10%)

Unstable integrator (synaptic weights increased 10%)

Part III: Learning to Integrate

- How to accomplish fine-tuning of synaptic weights?

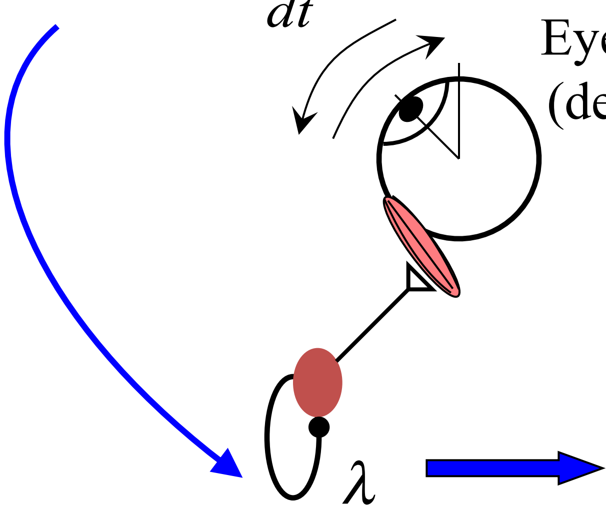
  ↳ IDEA: Synaptic weights learned from "image slip"

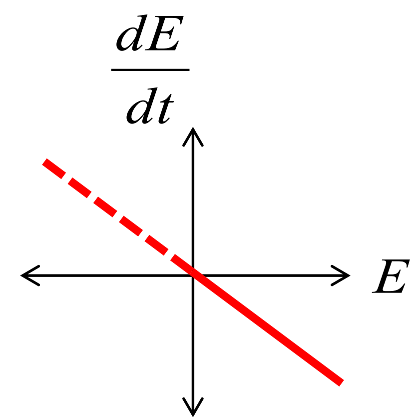
  (Arnold & Robinson, 1992)

- Imagine we have a leaky integrator

Image slip = $-\dfrac{dE}{dt}$

Eye slip (decay) $= \dfrac{dE}{dt} < 0$

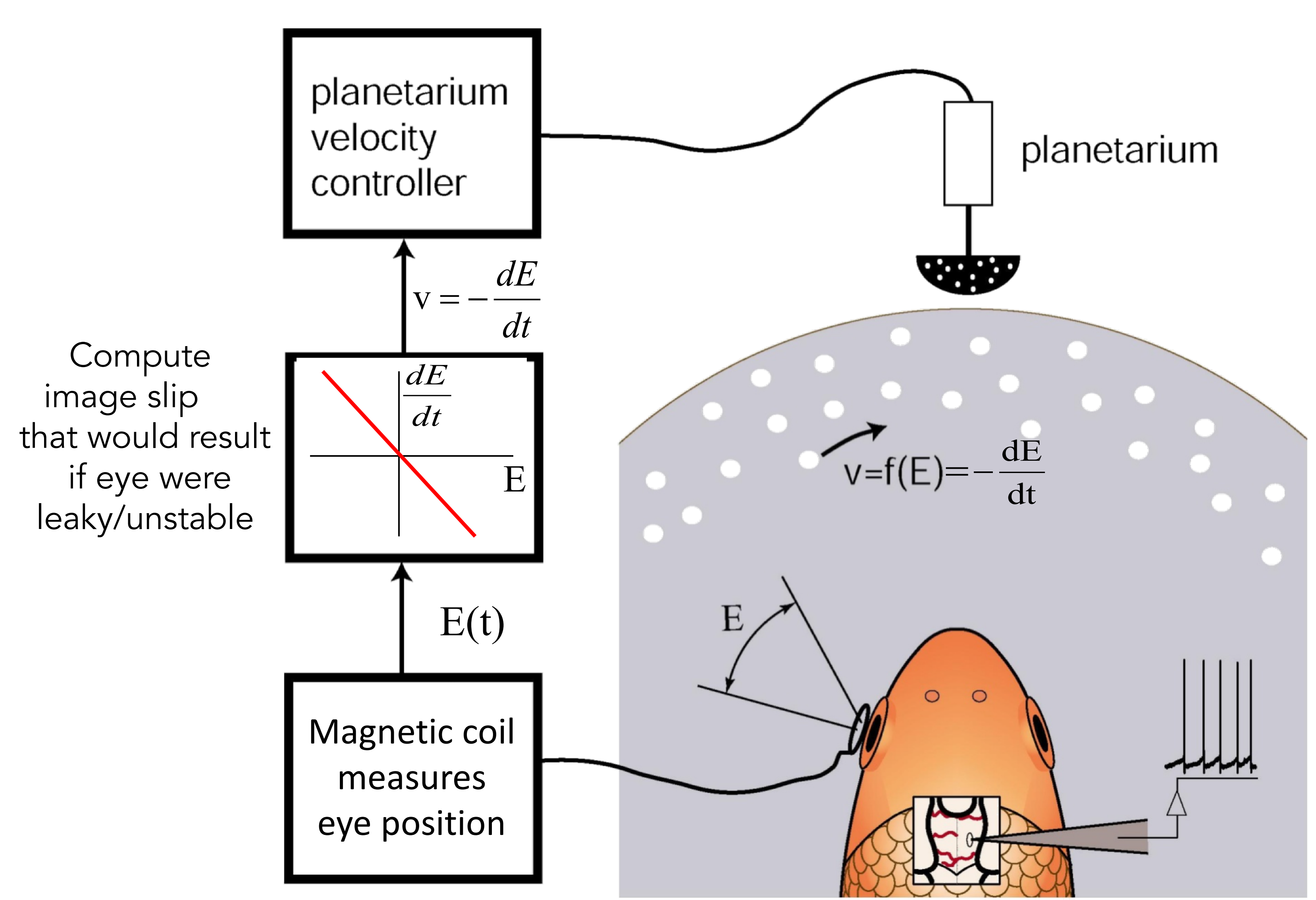




Need to turn up feedback → $\boxed{\Delta\lambda \propto -\dfrac{dE}{dt}}$

Learning to Integrate

- Experiment: Give feedback as if integrator is leaky or unstable



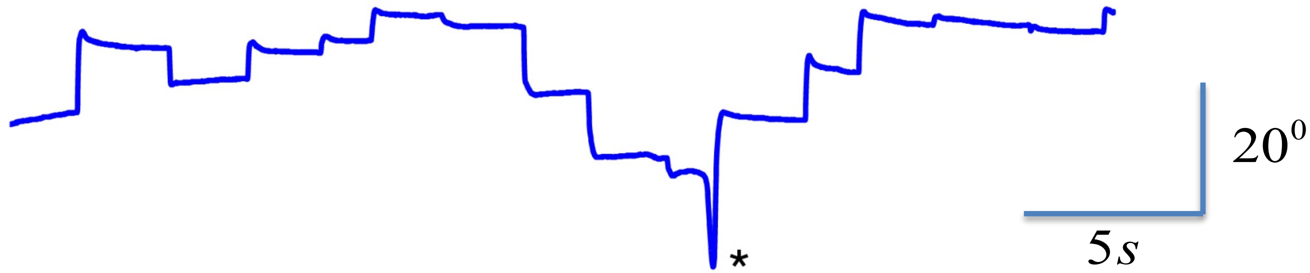
Learning to Integrate

- Experimental setup for tuning integrator

Learning to Integrate

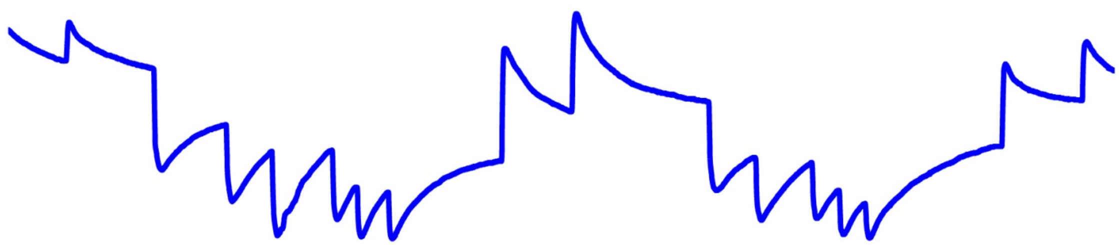
- Integrator can be trained to become leaky or unstable

Control (in dark):



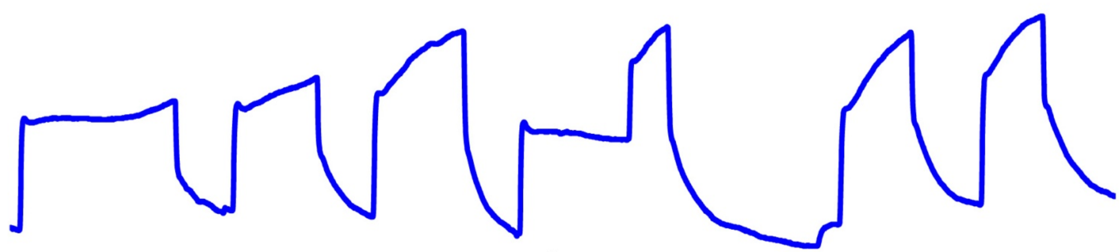
Give feedback
as if unstable

⟶ Leaky:



Give feedback
as if leaky

⟶ Unstable:



Summary and open questions

I. Goldfish do integrals!

$$\underbrace{Eye\,Position}_{\text{Integrator neurons}} = \int \underbrace{Eye\,Velocity}_{\text{burst input}}\,dt$$

II. How goldfish do integrals: neural mechanism

-Network feedback balances leakiness of neurons

-But…model is less robust than real integrator

III. Goldfish learn to do integrals!

-Integrator compensates for image slip

-How and where does learning occur?

Synapse modification? Intrinsic neuronal modification?

-Is visual feedback the only learning signal?

Acknowledgements

UC Davis

Mark Goldman

MIT

Sebastian Seung

N.Y.U.

Bob Baker

Princeton University

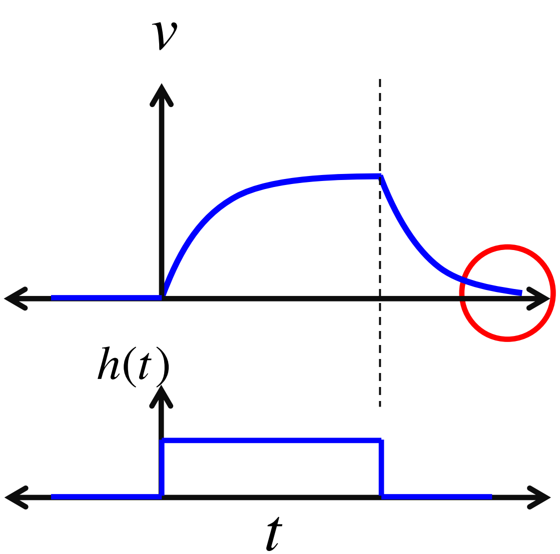
David Tank
Guy Major (Cardiff Univ.)
Emre Aksay (Cornell Med.)

Recurrent networks

- The behavior of the network depends critically on λ

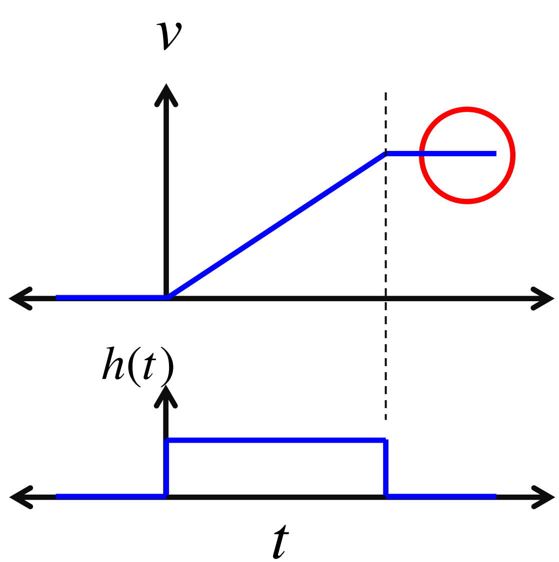
$\lambda < 1$

Exponential relaxation



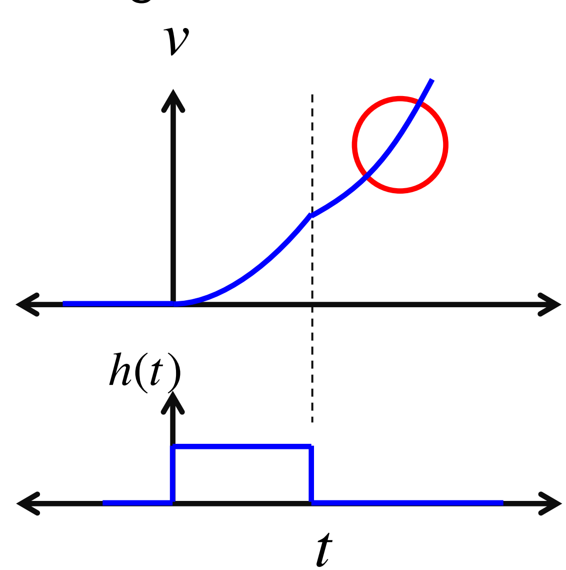
$\lambda = 1$

Integration



$\lambda > 1$

Exponential growth



With zero input…
relaxation back to zero

With zero input…
persistent activity!

MEMORY!