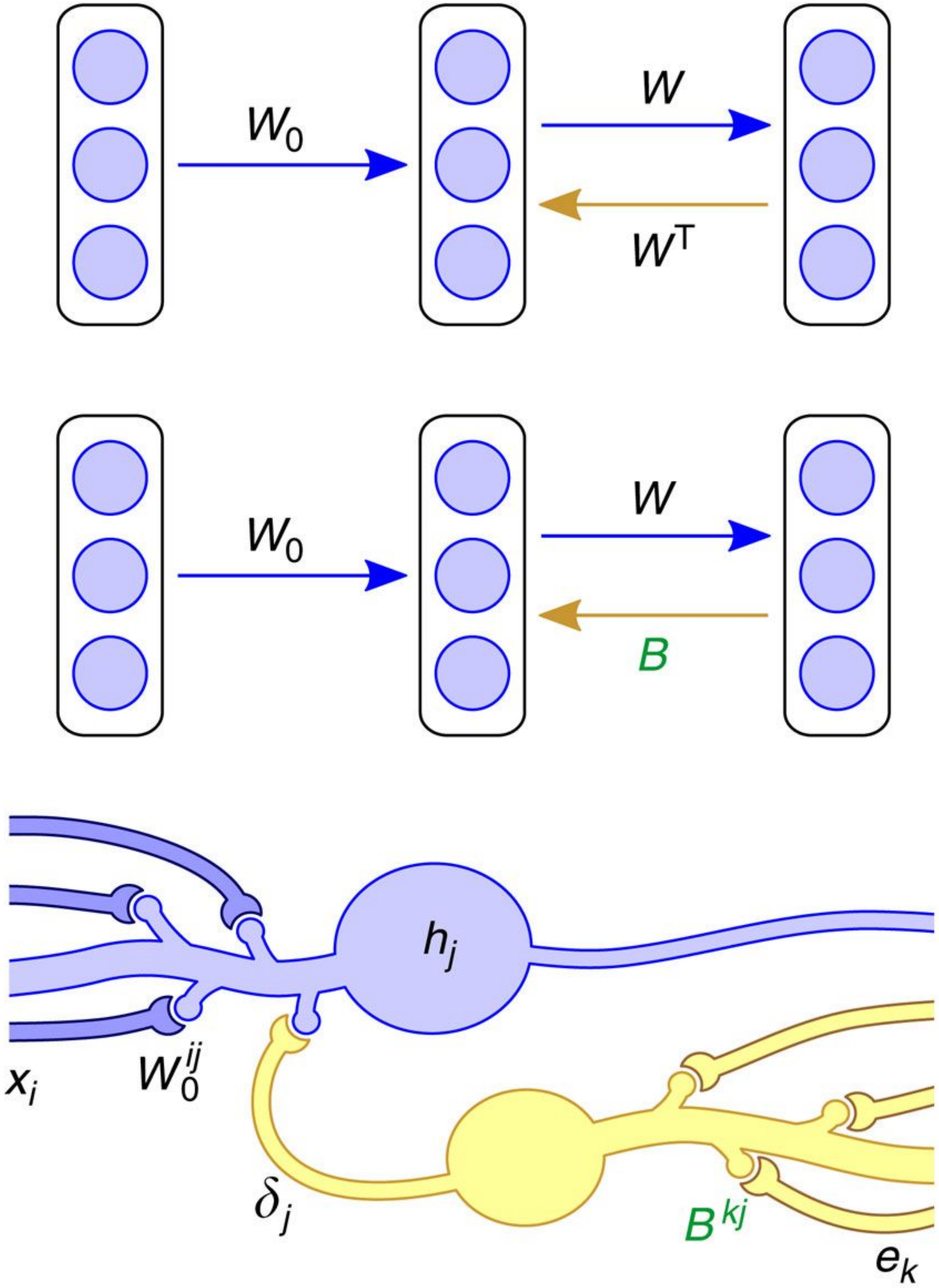
welcome to Week 6

Computation and the Brain
Fall 2019

Biologically plausible ANNs

[Lillicrap et al. 1914]:
constant (non-plastic)
random synaptic weights
in a **backwards** synapse
suffice for some learning!
(and this is biologically plausible)

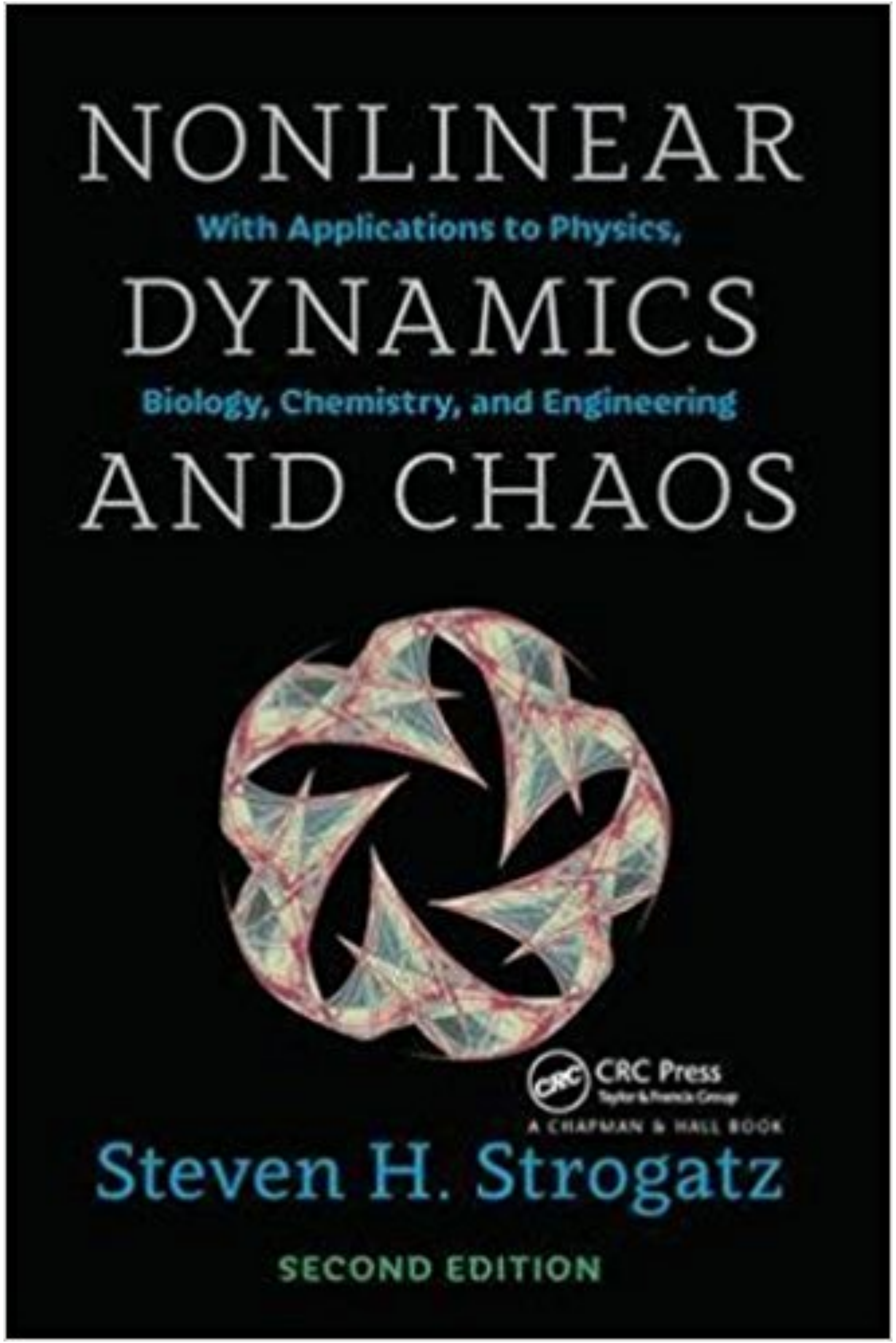


Biologically Plausible ANNs: NNevolution, the experiment

Repeat for T generations:

- Generate P genotypes and the corresponding ANNs
- Evaluate each on a minibatch
- Find the ***fitness*** f(A) of each allele A (say, 1/4 minus the average square error over all genotypes that have it)
- Update the gene probabilities $p_A \rightarrow \sim p_A (1 + \varepsilon f(A))$

Experiments: competitive with signSGD



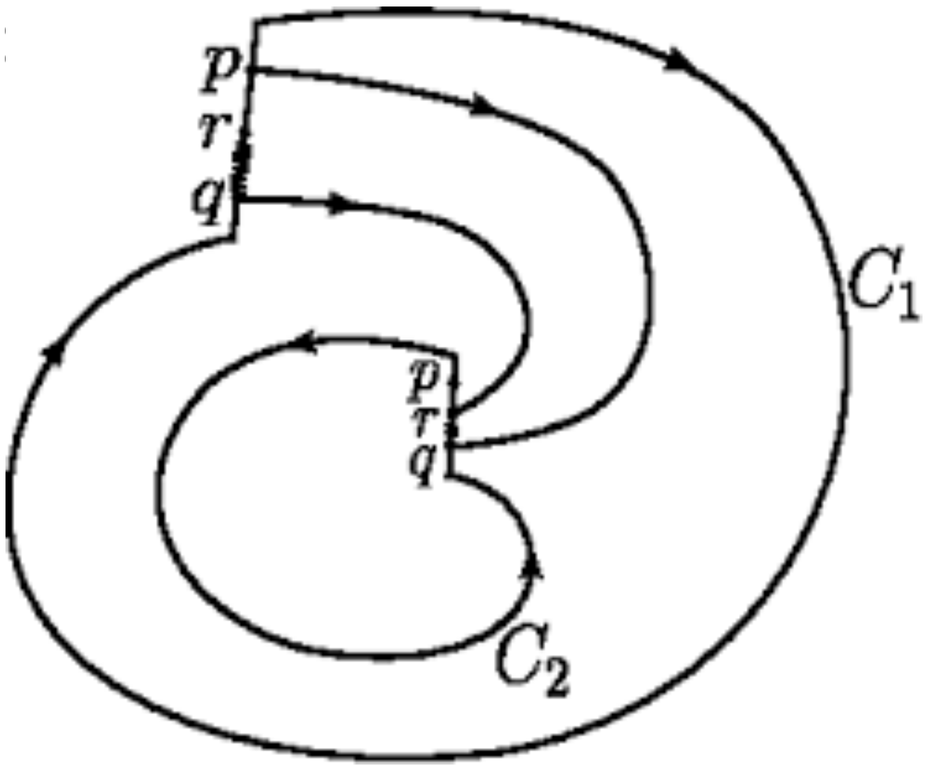
Excellent and accessible book!

One and two dimensional ODEs, summary

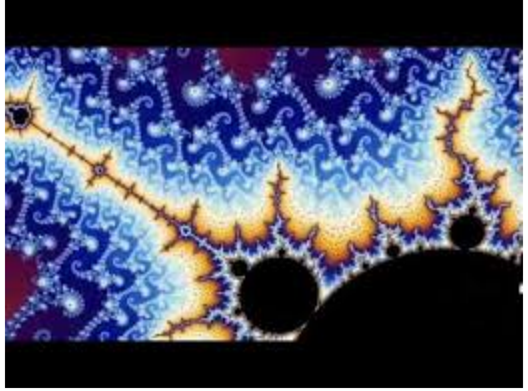
- In 1D dynamical systems, the limit behavior is equilibrium (or growth): there can be no cycles
- In 2D? **Poincaré – Bendixson theorem**:

In 2D the limit behavior is either stationary (equilibrium) or periodic (cycles)

- There can be no chaos here, the flow "restrains itself"



What is Chaos?



- Exponentially small perturbations in parameters and initial conditions lead to qualitatively different behaviors
- A seemingly periodic behavior repeats forever, except that the system never exactly cycles (Lorenz)
- An attractor is strange (fractal-like)
- In discrete time: there cycles of all kinds of periods (but a cycle of period three is enough….)
- The system cannot be solved (or understood…) in any satisfactory way

The fundamental theorem of dynamical systems (cont.)

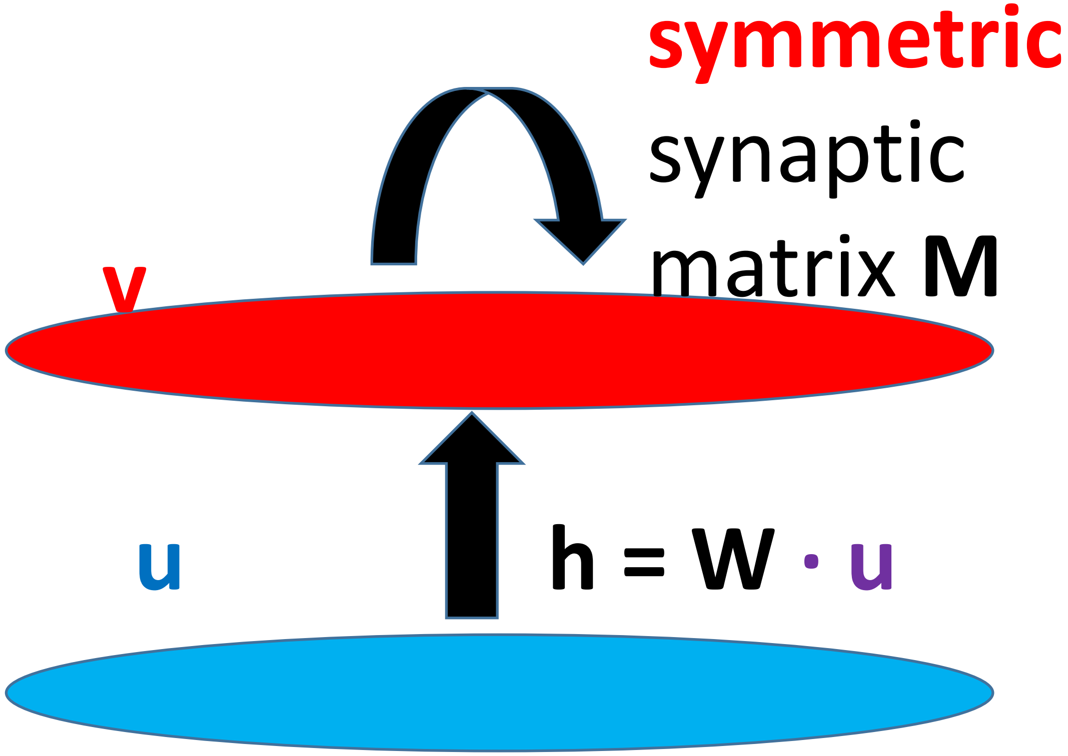
- **Theorem** [Conley 1984]: The domain of any dynamical system ***(both continuous- and discrete-time)*** can be decomposed in the chain recurrent components (CRC) and the transient parts. There is a Lyapunov function that drives any transient point towards the CRCs.
- In other words "if you squint a little, chaos goes away and the dynamics converges to equilibria or "cycles"…"

Suppose F is *linear* and M *symmetric*

$$\tau \cdot dv/dt = h + (M - I) \cdot v$$

M has positive eigenvalues $\{\lambda_k\}$ with orthonormal eigenvectors v_k

eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$



Suppose that $\lambda_2 << \lambda_1 \approx 1$

We can ignore all other terms, and we have the asymptotic solution

$\mathbf{v}^* \approx (\mathbf{e}_1 \cdot \mathbf{h}) \cdot \mathbf{e}_1 / (1 - \lambda_1)$

So, this circuit takes the feedforward input h and projects it on $\mathbf{e}_1$, amplifying it by a large number

(If two eigenvalues of M are close to one, it projects on a plane…)

Suppose $\lambda_1 = 1$

Recall : $\tau \cdot dc_1/dt = \quad n - (1 - \lambda_1) \cdot c_1(t) + \quad e_1 \cdot h$

So, the circuit ***integrates*** the feedforward input's projection on $\mathbf{e_1}$

NB: Integration means memory. (Why?)

A system like this seems to be at work in the brain stem of verterbrates, remembering eye position

ReLU FFR networks

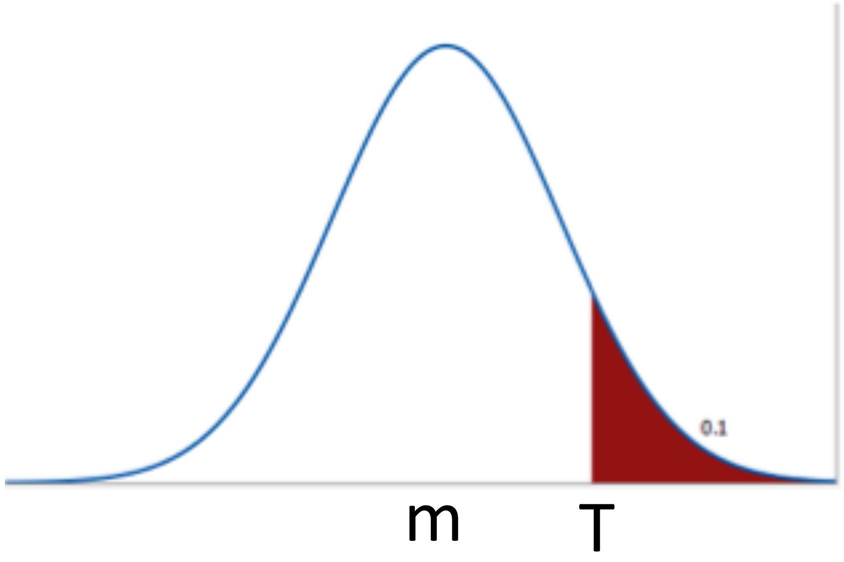
$\boldsymbol{\tau \cdot dv/dt = -\, v + [h + M \cdot v - T]_+}$

Subtract a vector of thresholds **T** and set to zero if negative

Can model simple and complex cells!

E – I balance

Notation: GT(T, m, var) is the Gaussian tail with parameters m and var above threshold T



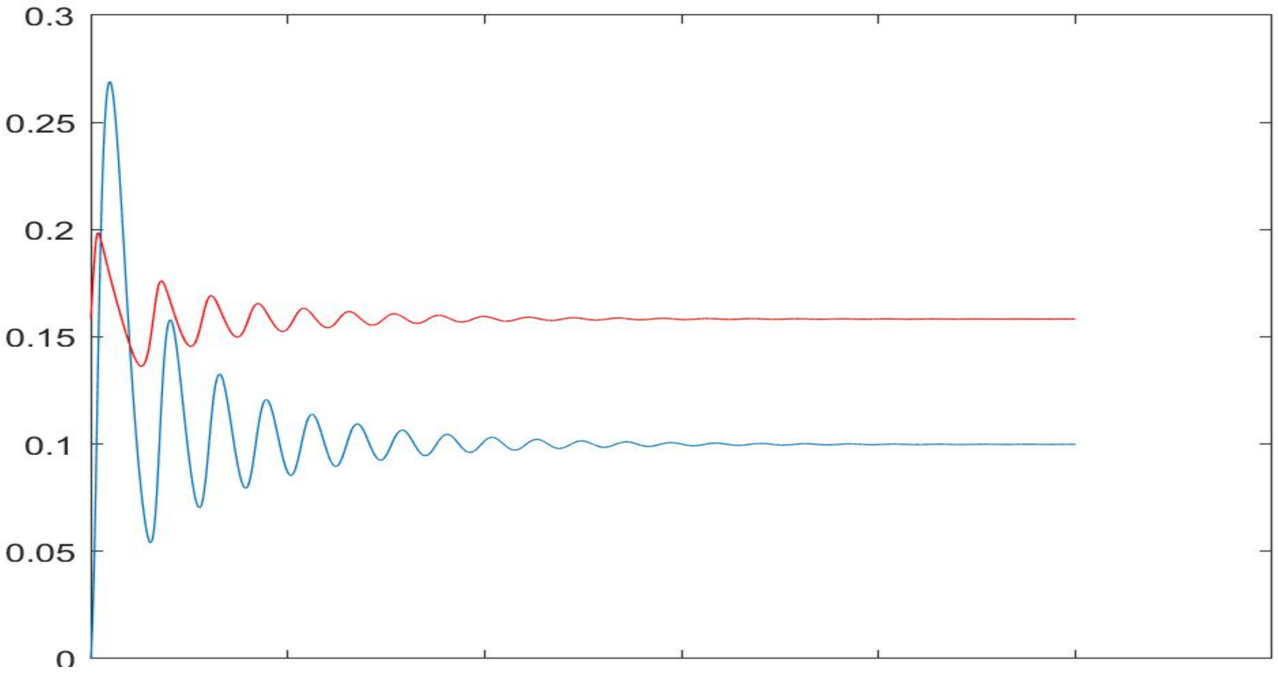
Nonlinear ODE

$\tau_E \dot{E} = GT(T_E, np(E - I), np(1\text{-}p)(E + I)) - E$

$\tau_I \dot{I} = GT(T_I, npI, np(1\text{-}p)I) - I$

E – I balance (cont.)

If τ_E is sufficiently larger than τ_I, an E – I balance will be reached after a few up and down oscillations



QUESTIONS?

- What are some advantages of backprop-free learning? Can we use the extra degree of freedom in the feedback weights for some other goal, for example sparsity? Is feedback alignment more suited for non differentiable non-linearities? Can being gradient-free make the model more robust to adversarial attacks, which are often gradient based?
- After Lilicrap et al 2016 paper, was there any experimental work to study whether learning happens in the brain by means of a constant random matrix?
- Are there theories about emergent behavior in the brain that could lead to consciousness and memory? How do we think memories and the impression of consciousness are created from a structural point of view? Would a fully mapped connectome allow us to probe this question?
- Concretely what form does learning take in the brain? Mostly Long-Term Potentiation? How does synaptic pruning connect to this? Do we think pruning occurs when a connections weakens below a certain threshold?
- Why were there such great advances in the study of the visual cortex but not in other brain regions?
- How practical is using random matrices instead of the transpose? I'm convinced by my testing and the paper but I'm still confused how it's possible. Is training slower in any way? Does it scale as well as backprop and Jacobians?
- While it's promising that small neural networks can encode polynomial functions, this doesn't immediately apply learnability, either from the perspective of sample complexity or computation. Can sample complexity bounds be better for these functions than the generic ones given by the pseudodimension of networks of appropriate size? If the network is shallow, can we obtain convergence results for gradient descent?

- While it’s promising that small neural networks can encode polynomial functions, this doesn’t immediately apply learnability, either from the perspective of sample complexity or computation. Can sample complexity bounds be better for these functions than the generic ones given by the pseudodimension of networks of appropriate size? If the network is shallow, can we obtain convergence results for gradient descent?
- While we can’t easily map all of the connections between neurons in the brain, is it possible to identify all of the connections of a given neuron, or at least select one uniformly at random? This would allow to perform random walks through the connectome, which could yield insight into distributional patterns of connectivity (such as frequency and size of cycles).
- WIth the seemingly personalized pruning of connections after birth discussed in the lecture, even if we could map all connections in one person’s brain, how would this generalize to a human connectome? Is there a way to say two neurons in different brains are serving the same purpose?
- Similar question from someone else: How do we know if we’ve found a canonical pattern in cells? During an experiment one can only sample few rats/patients, how can we generalize to a population?
- How do you feel about connectomics?
- Is there any hope that we will be able to deconstruct the brain so we can fully understand and replicate it… or is it the best path to just use the brain as inspiration to build something completely different?
- Given the connection pattern of neurons (eg. sparse, correlated) can we build a mathematical model to formalize the connection graph? Can we use tools for social networks to model the connection graph?
- I am interested in digging deeper into inhibition, and in exploring the “10x more backward connections”. Do inhibitory neurons have similar learning properties to other neurons we’ve studied so far? How does inhibition affect synaptic weights and learning in general?
- Based on readings, there are underlying organisations within the neural structure that we have yet to know about. Whilst the field is in its early stage I wonder what the role of genetics is on how neuron connections change. Babies go through pruning: how much of this is from DNA, how much of this is from environment? If a baby does not move for months would the neurons still go through the same changes?

- In the video Lichtman mentioned that adults have sparse mental connections because they have learned to have only the connections that will do certain tasks. What happens if someone learns something the wrong way? How hard is it to add back connections?In the inhibitory motifs example, for feed forward inhibition, why does it matter that the LIN is inhibiting the TC when the RGC goes directly to the TC as well?
- I was trying to apply the Johnson Lindenstrauss Lemma to the number of wordpieces in recent language models. Systems with 32000 wordpieces and 128 dimensions per wordpiece have been found to work well but JLL would give epsilon about 0.8 for this combination. What's a common value of epsilon one could expect? Could it be that epsilon < 1 is all that is needed for the network to be able to tell all words apart?
- There seems to be a lot of funding going into mapping the connectome of larger animals,s o most people must believe it is highly relevant for something. First, is this really backed by C elegans connectome, i.e. does the connectome of the worm say something about each individual worm, or is it just a cool tool that its connectome (.... question cut off by mistake in submission procedure)
- Do connections play a larger role in the brain's functioning than neuron type?
- How could advances in connectomics help with the development of brain inspired technology?
- Why don't more papers take the statistical approach that the Song et al. paper does? It seems much more convincing to me than the conjecture about brain function.
- What are some ways we can use canonical patterns in deep learning and would that drastically change the way models are currently trained/tested?
- Do you think the feedback connections are an important part of learning or do you think it is primarily something like STDP/LTP/LTD?
- Has anyone tried the pruning of neural networks described with human muscle neurons in any sort of neural network like machine learning architecture?
- The number of connected neurons of a child is larger than that of an adult, while the strength is less. How does the human do graph sparsification? Can we use the spectral graph theorem to explain it?
- Lichtman showed us automated techniques for slicing + imaging, but what methods are used to make sense of vast data? What kind of models are used to distinguish types of neurons/cells, identify anomalies, count cell bodies, etc.

A discrete-time neural system: Hopfield net

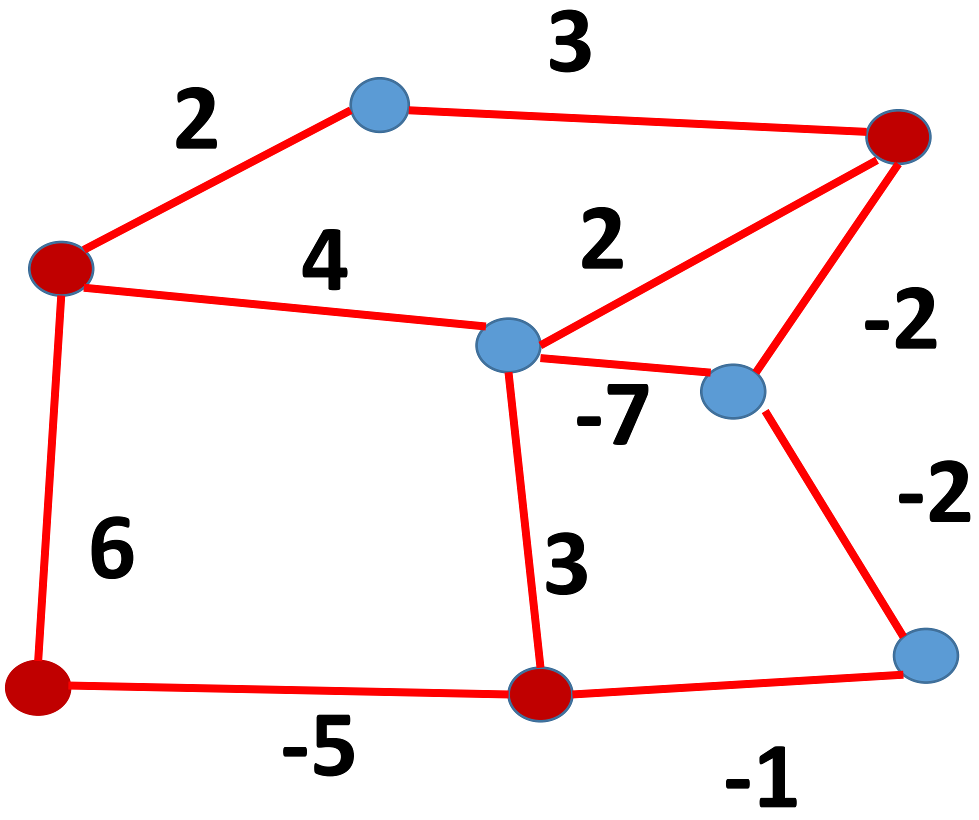
Nodes have two values: **+1**, **-1**

Symmetric synaptic weights

Node i is ***happy*** if $\sum_j v_i v_j w_{ij} \geq 0$

Algorithm/dynamical system:

while there is
an unhappy node,
flip it



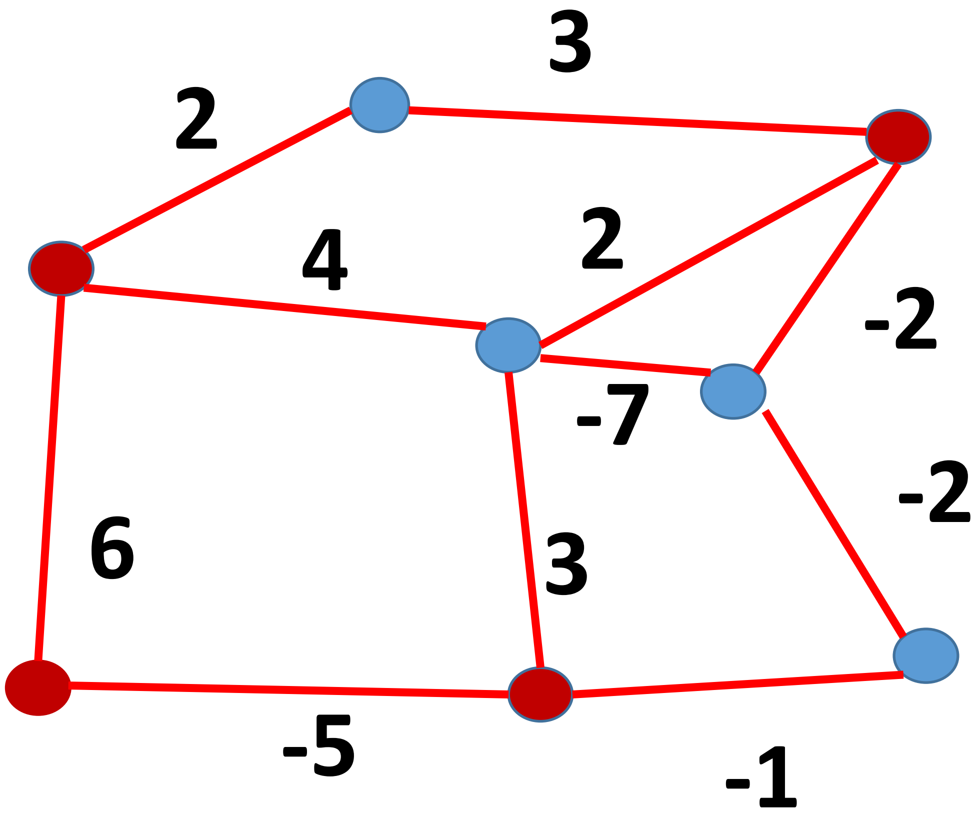
A discrete-time neural system: Hopfield net

Theorem [Hopfield 1982]: Dynamical system converges

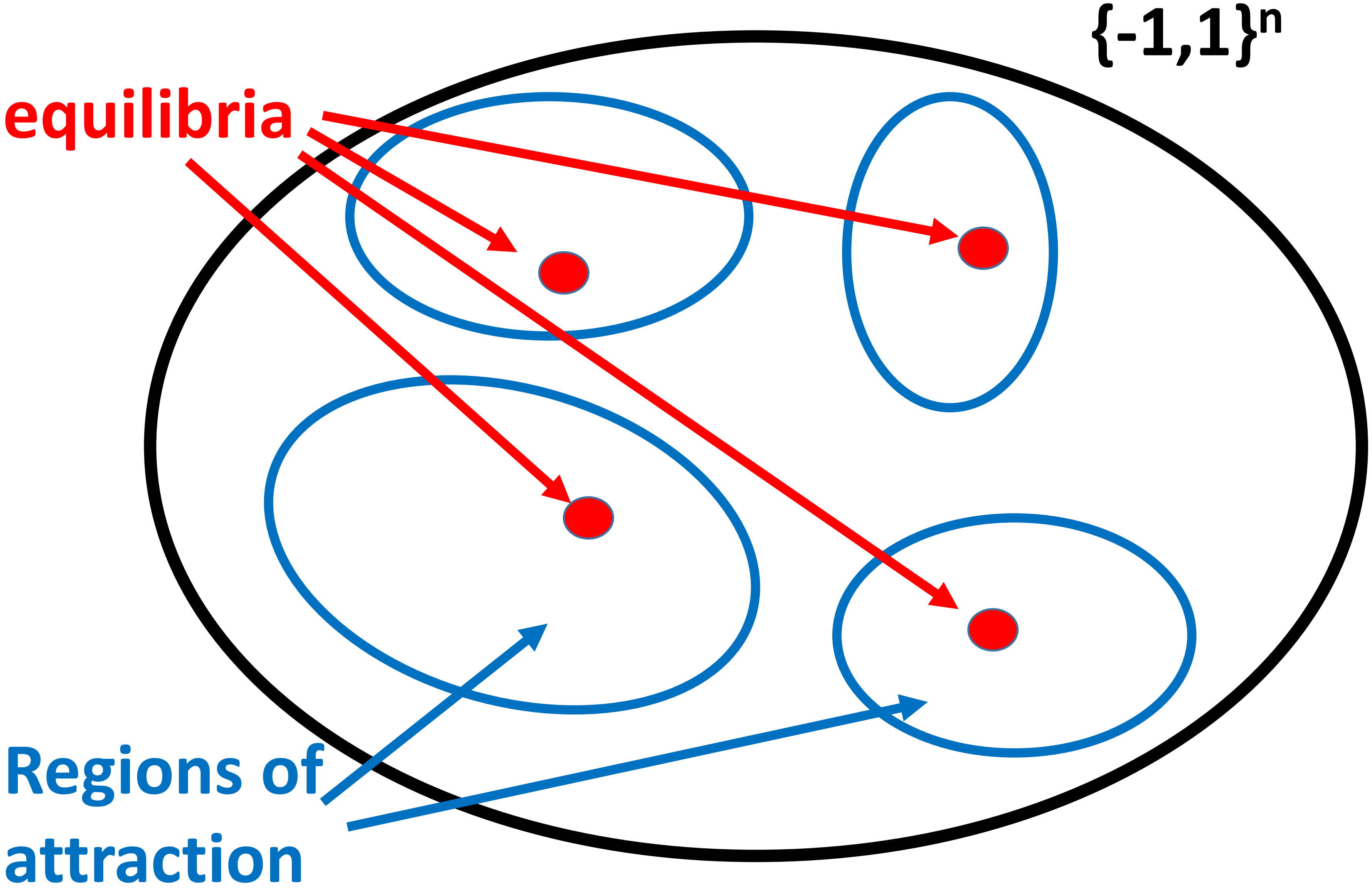
Proof: **Lyapunov** function

$$\chi(v) = \Sigma_{i,j}\, v_i\, v_j\, w_{ij}$$

always increases



Goal: Pattern completion



How do you train a Hopfield net so that it pattern completes?

- Given a set of desired memories $M_1, \ldots, M_m \in \{-1, +1\}^n$
- Set the **weight** of edge a-b to $\boldsymbol{\Sigma_k M_k^a \times M_k^b}$
- That is, for every memory k we bias the weight in the direction "that memory k wants the weight to be".
- Question: does it work?

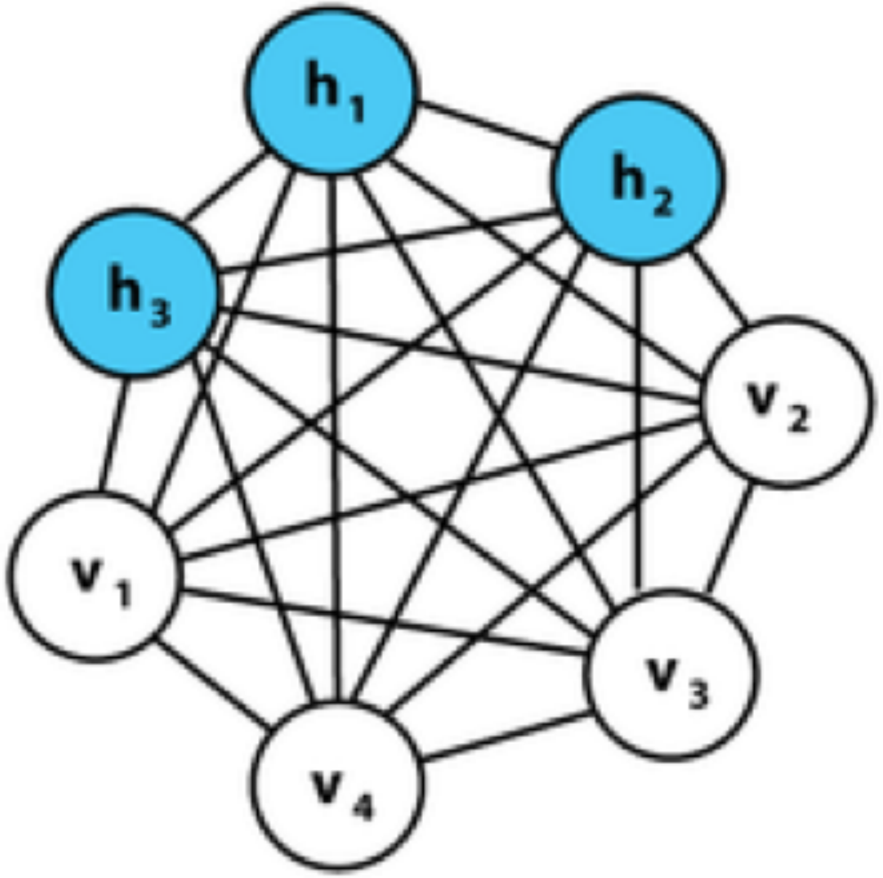
Theorem: with n nodes, 0.138n random patterns can be stored with probability of erroneous retrieval < 0.4%

Do Hopfield nets work?

- Spurious memories: if **M** is stored, **-M** is also stored and retrieved
- Also, if **M, M’, M’’** are stored, so are **± M ± M’ ± M’’**
- Finally, if you store **M** K times, it will be retrieved L >> K times more often than other memories
- ***Fun:*** [Hopfield and Tank 1984] show that Hopfield nets solve the traveling salesman problem with 10 cities

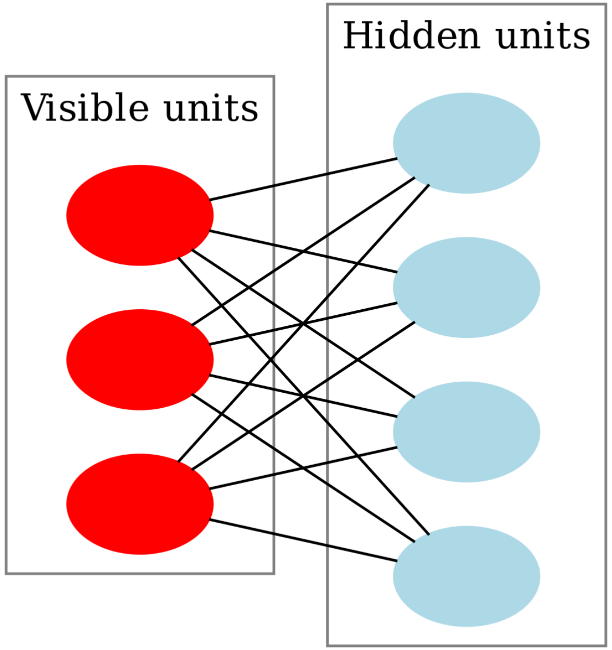
Boltzmann Machines

- Some of the nodes are **hidden**
- The visible nodes receive the training data by clamping
- After each training data vector is input, the whole network is left free to run until it stabilizes



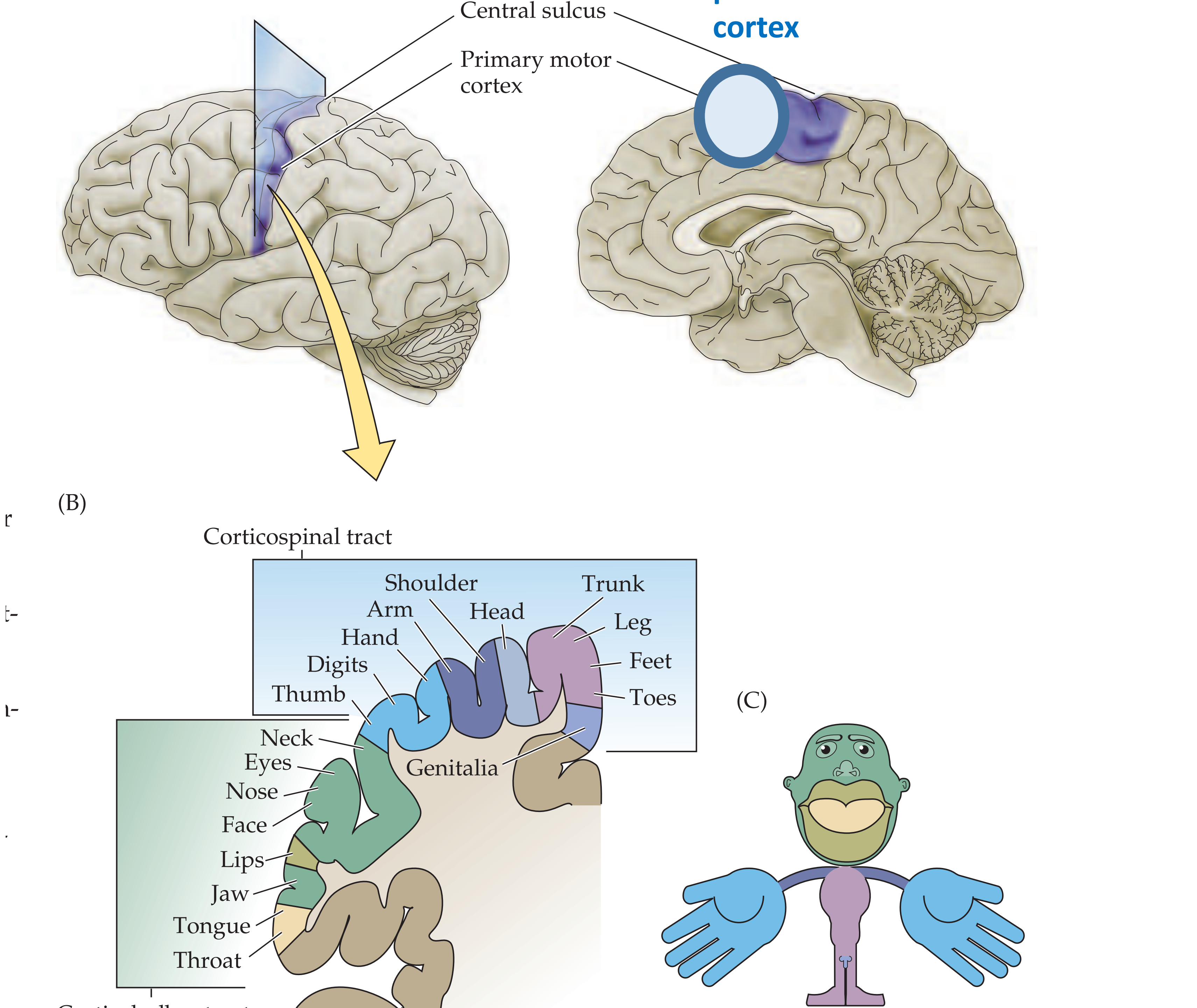
Boltzmann Machines (cont.)

- With some engineering, learning can happen
- Restricted Boltzmann machine: graph is bipartite [Hinton 2005]
- Can be stacked to form a deep net…

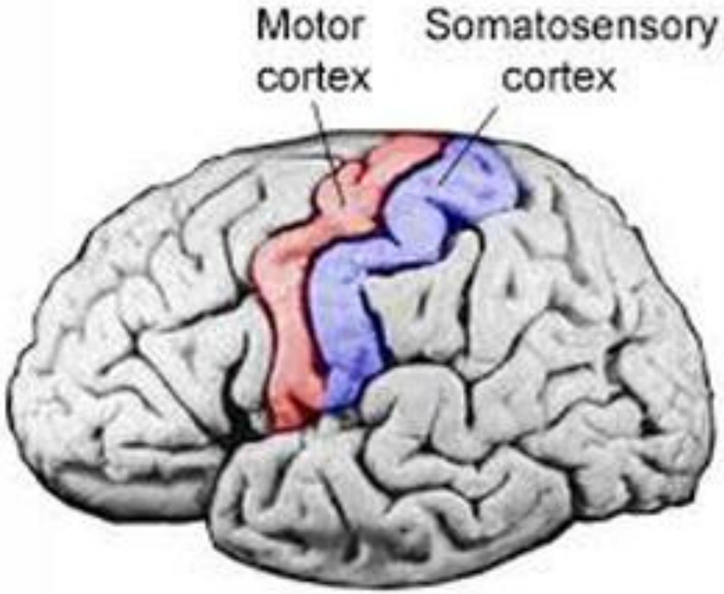


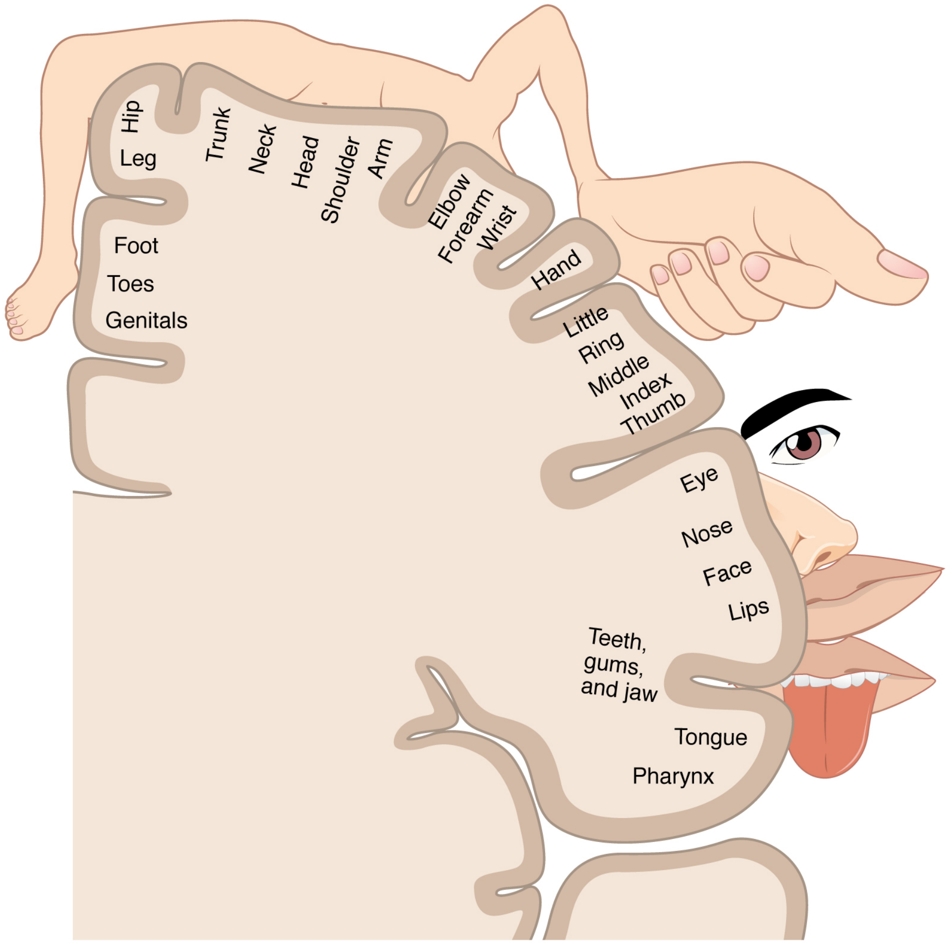
- History: Hopfield nets were the precursors of recurrent nets and LSTMs…

Next:
The Motor Cortex



Compare with: the somatosensory humunculus





Back to the motor system:

- Motor maps date back to the 1960s and were found to be basically right. ***But*** the complete picture is much more complicated and under construction…
- Motor talent is reflected in motor cortex areas. E.g., violinists, cellists, and classical guitarists appear to have exaggerated left digit areas, presumably developed in parallel with their motor talent. This means the map is plastic…

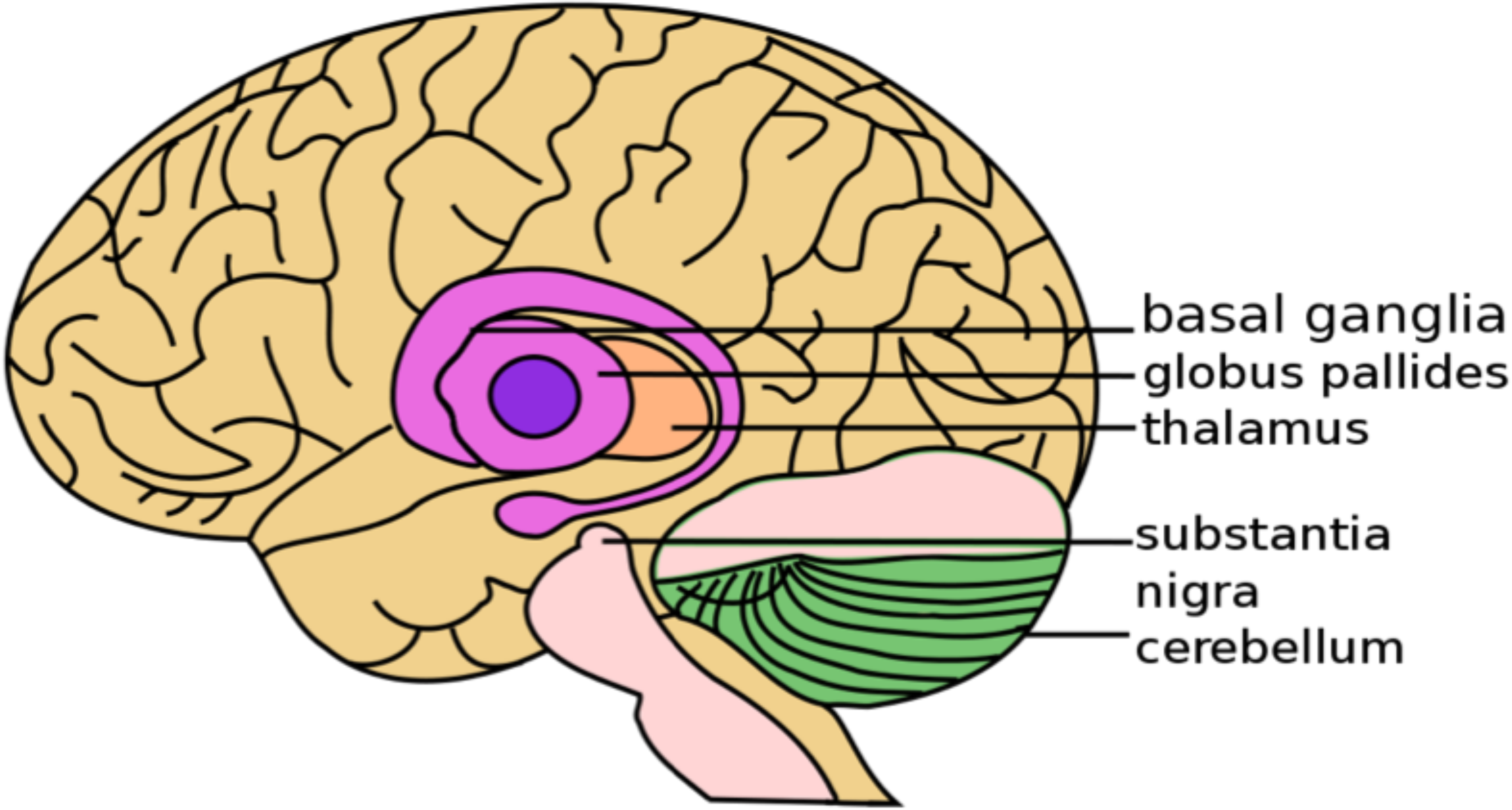
What do we know (about the motor cortex) (and when did we know it)

- 1890s: Motor cortex discovered: electrical stimulation causes limb motion
- 1900s: The first "motor maps"
- 1905: Premotor cortex
- 1960s-2000s: Sophisticated recording and electrical stimulation technology allows extensive experimentation
- Theories of motor cortex reflect emerging theories of the visual cortex (simple and complex cells, coding of movements etc.)

Back to the motor system (cont):

- Premotor cortex receives input from many brain areas, provides input to the motor cortex, but also directly to the spinal cord, and seems to be implicated in conditional motion
- The roles of the motor cortex and the premotor cortex are not understood. They can both elicit motion in the absence of the other
- Motor and premotor cortex are just the remote control center of a complex and very extensive system, involving also the brain stem, spinal cord, nerves in muscles, **basal ganglia,** and much more

Basal Ganglia and Related Structures of the Brain



M1, M2, M4, etc?

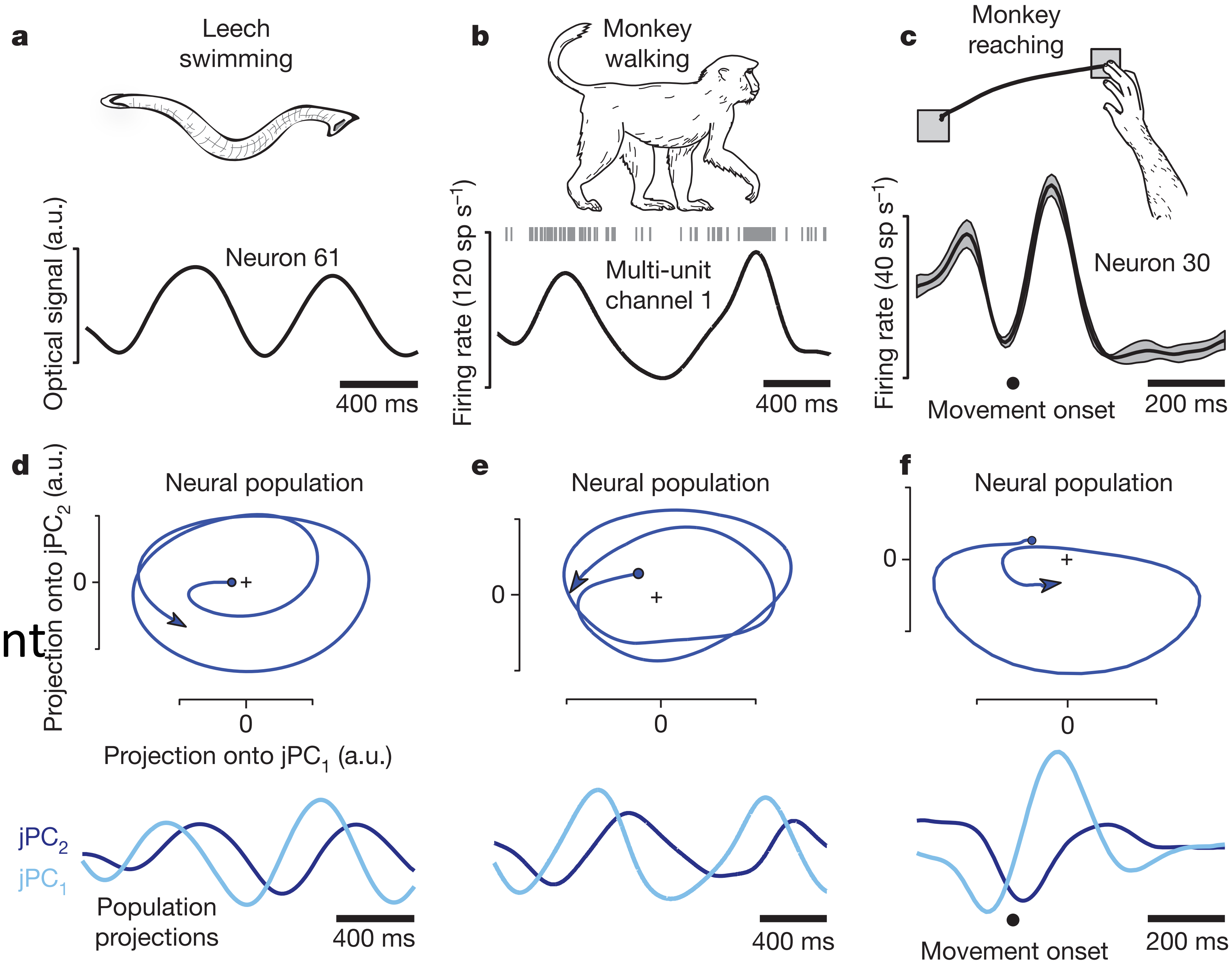
- Classical view: a neuron in the motor cortex encodes the motion being enacted
- But this coding appears to be complicated (polyphasic)
- That is, its firing rate is a function of many motion parameters simultaneously, such as direction, speed, state of muscles, positions of bones, etc.

$$r_i^t = f_i(par_1, par_2, \ldots , par_{30})$$

Opposing view [Churchland et al. 2012]

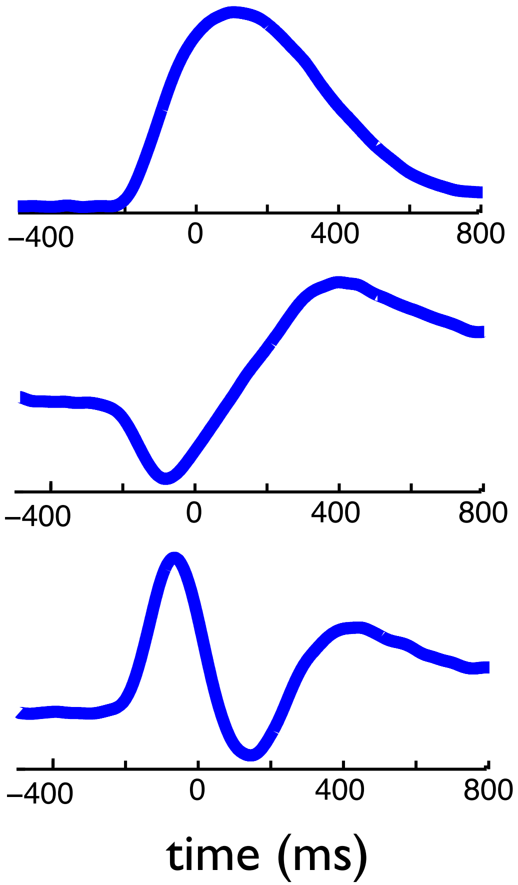
- A neuron in the motor cortex is a row (eqn) of a dynamical system, whereby the motor cortex generates and controls movement: $\dot{\mathbf{r}}^t = \mathbf{f}(\mathbf{r}^t, \mathbf{u}^t)$
- So, neural responses reflect underlying dynamics, and encode stimuli only incidentally
- Hypothesis: "movement generation across the animal kingdom involves rhythmic, oscillatory activity"
- **BUT** does the neural state rotate during all motion (even non-rhythmic movement like reaching)?

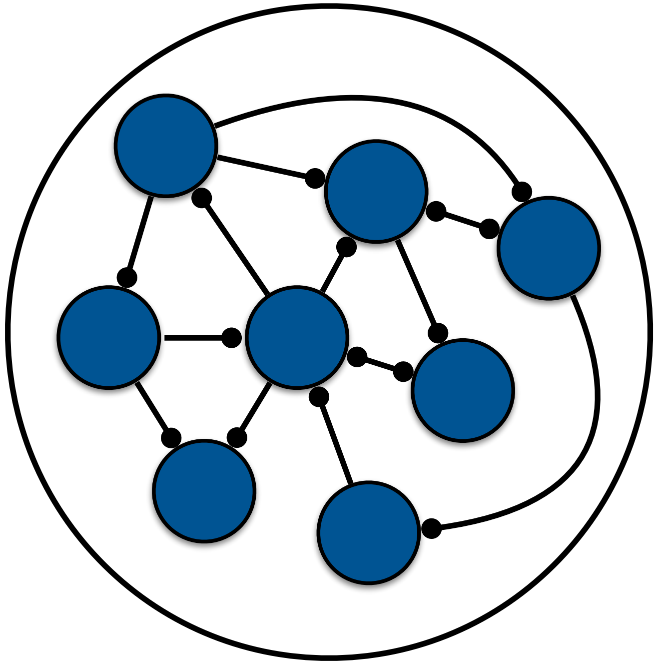
- Two animals
- Three motions
- Two of the motions are rhythmic, the third is not
- The all have significant cyclic projections of neural activity
- ***jPCA***



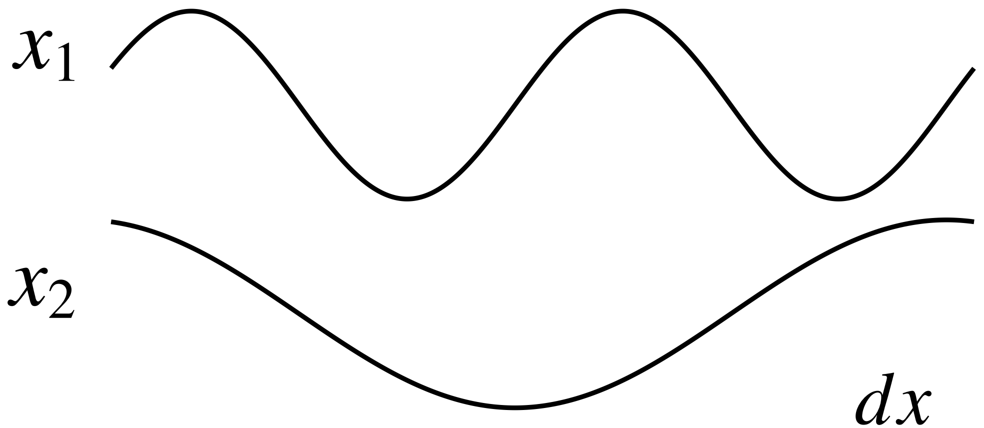
Dynamics

$x(t)$

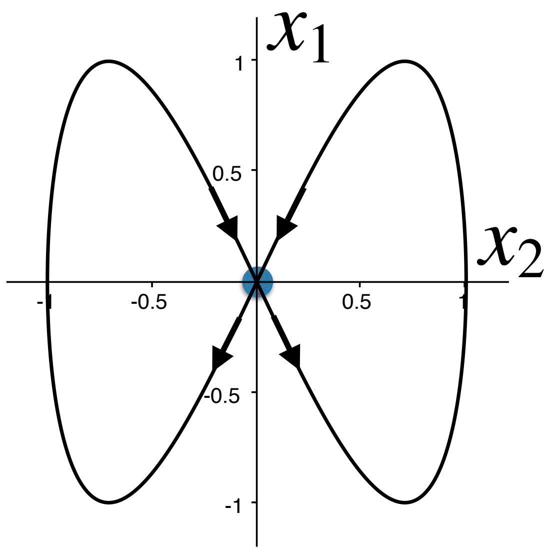




$$\frac{dx}{dt} = H(x)$$

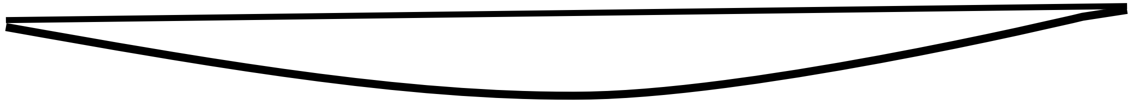


$$\frac{dx}{dt} \neq H(x)$$

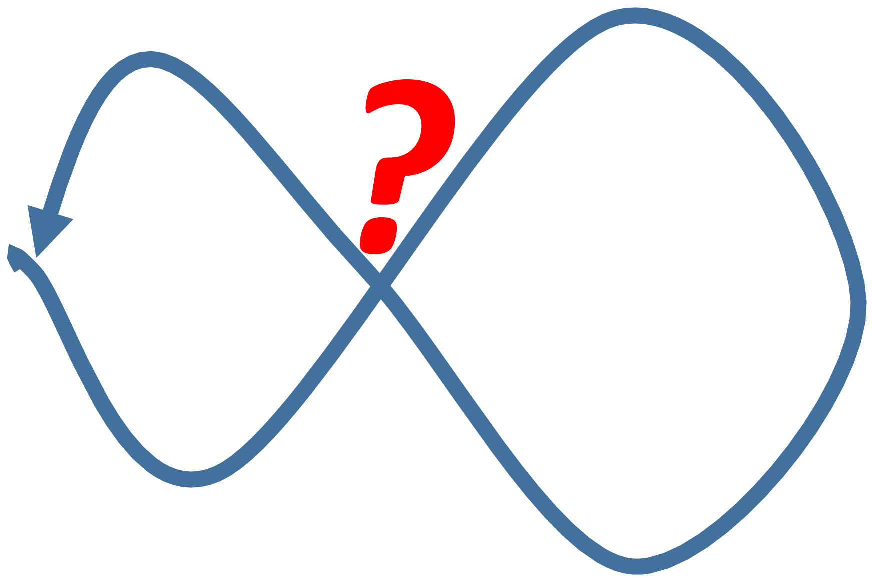


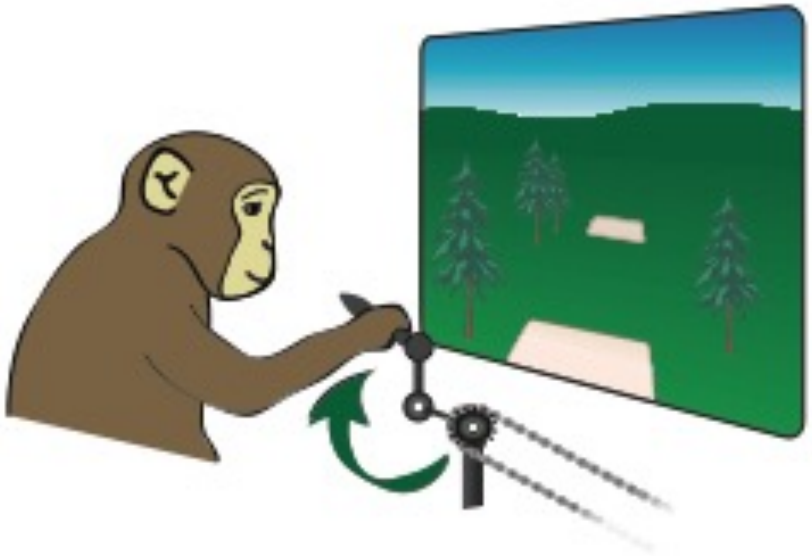
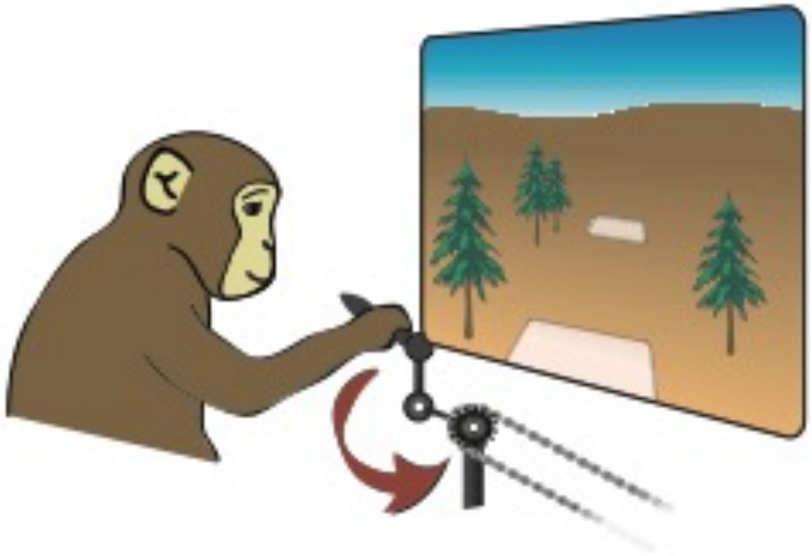
tangled

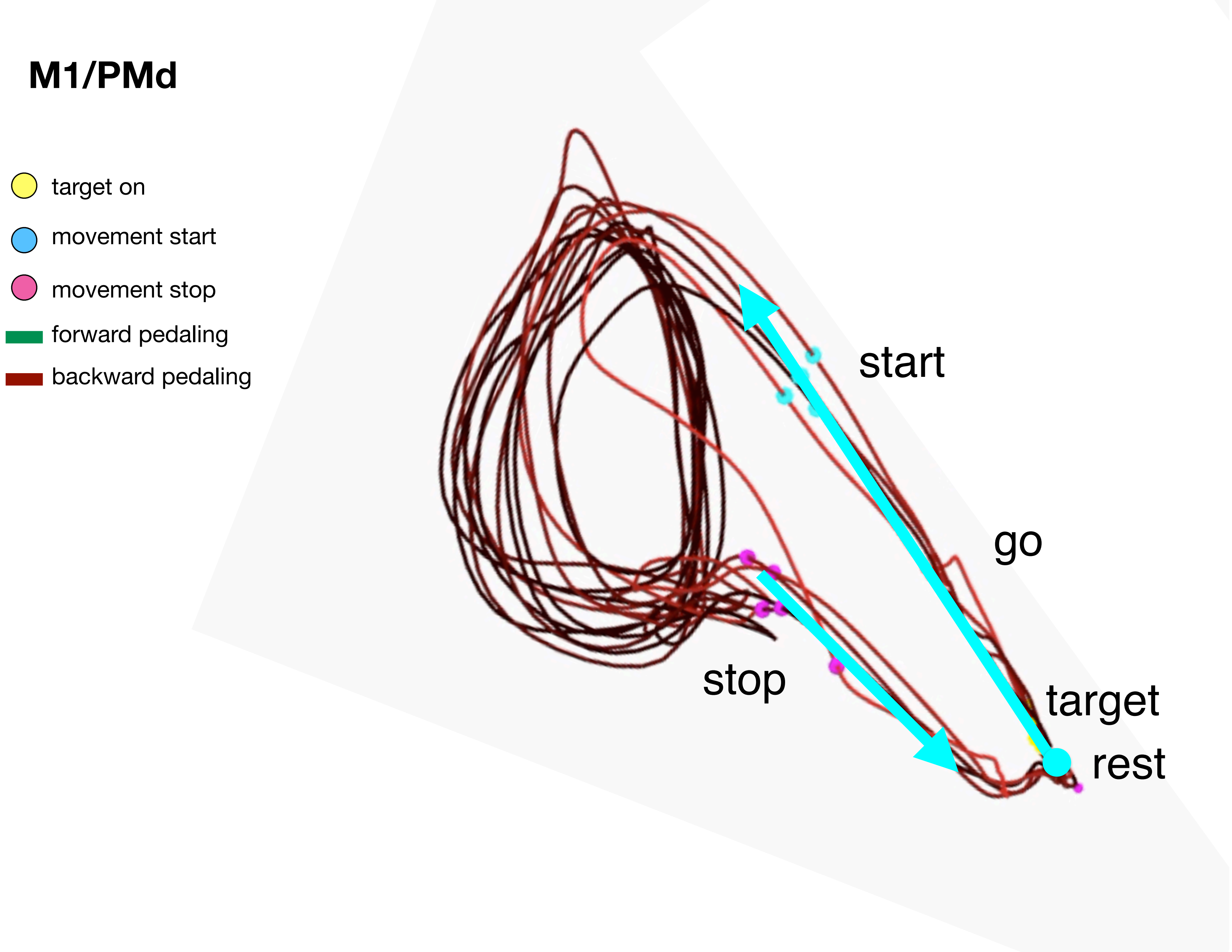
Untangled! x_3



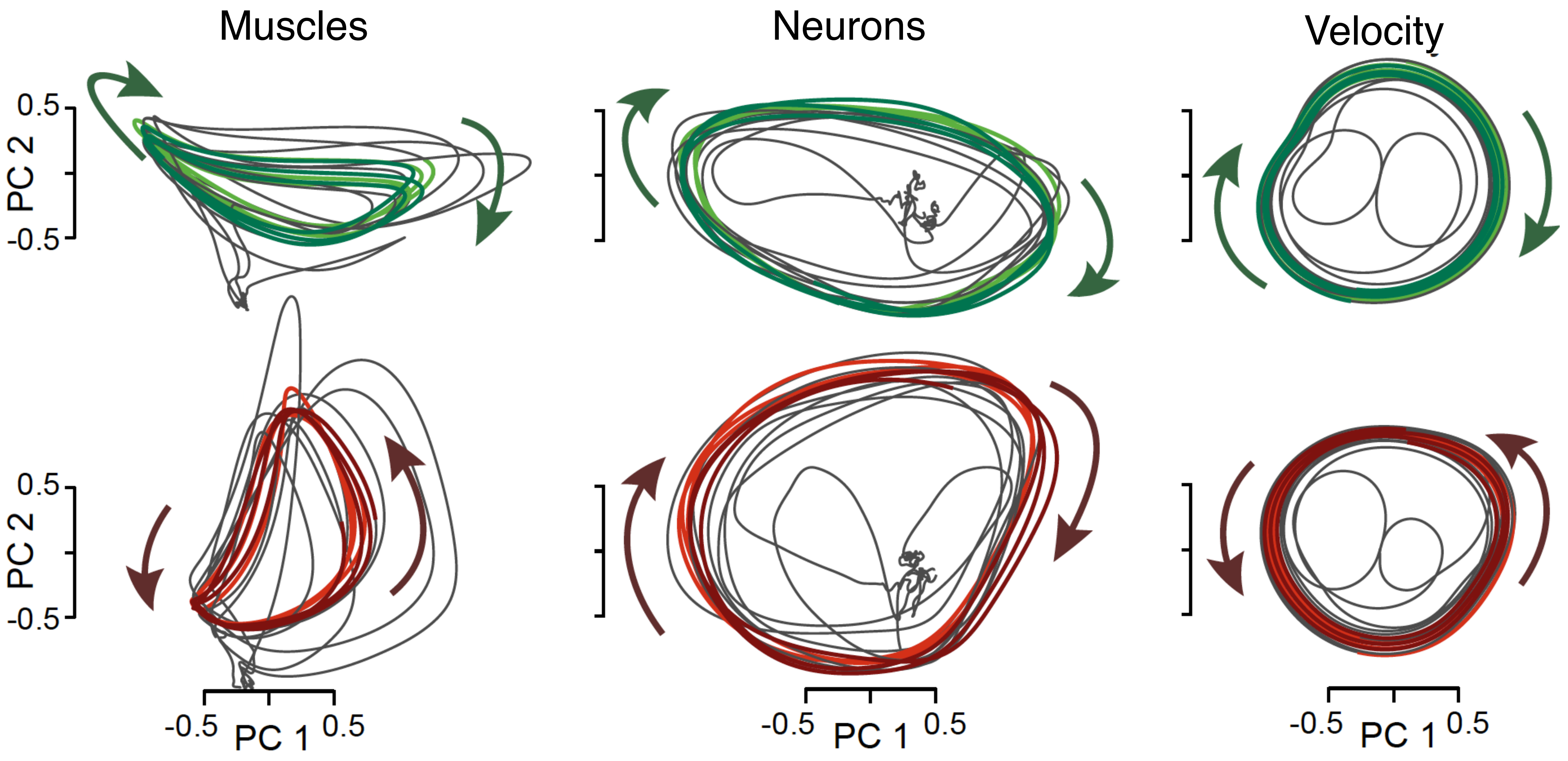
Fact: to represent a periodic motion you need more than two dimensions!





M1/PMd
target on
movement start
movement stop
forward pedaling
backward pedaling
start
go
stop
target
rest

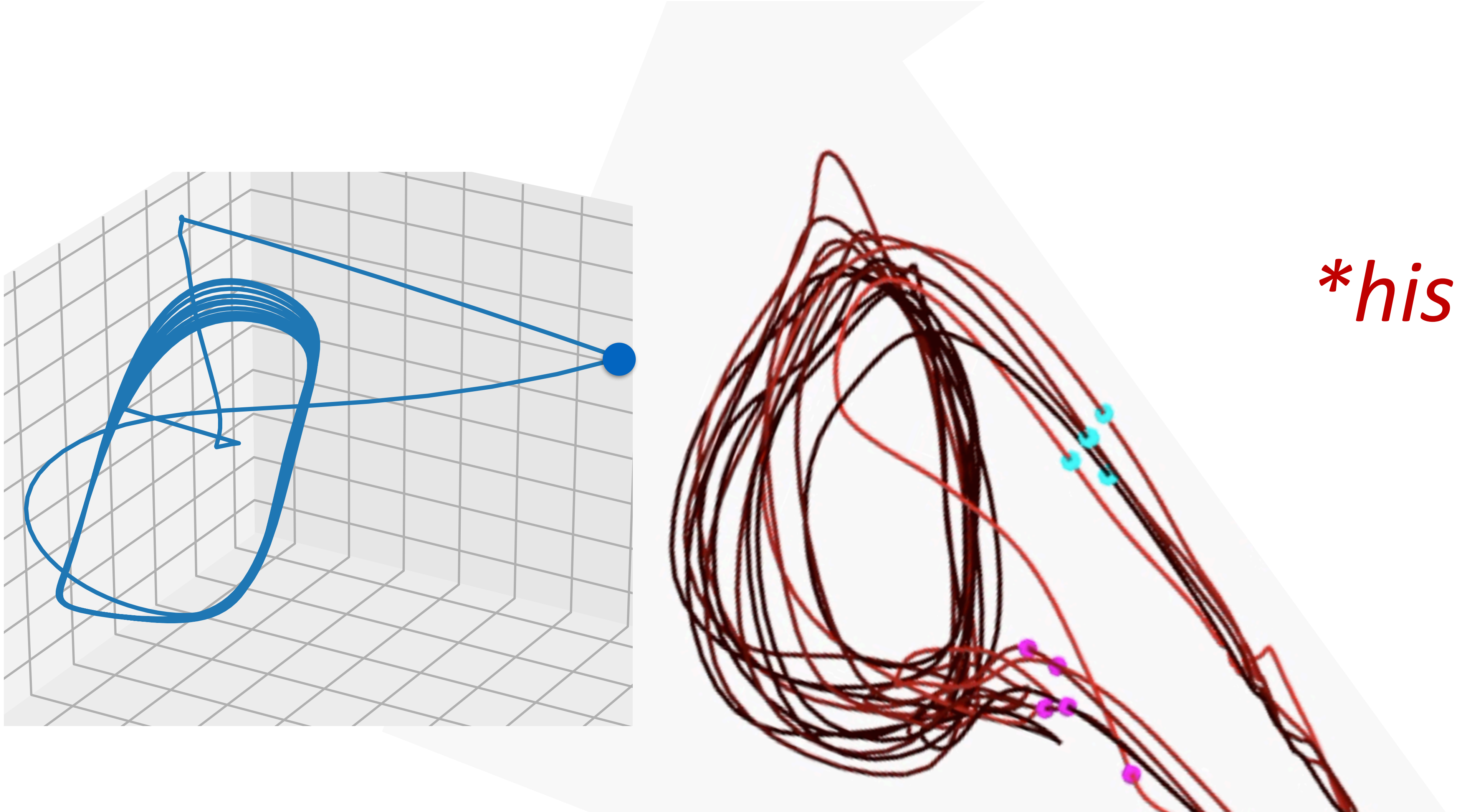


Muscles
Neurons
Velocity
0.5
PC 2
-0.5
0.5
PC 2
-0.5
-0.5
PC 1
0.5
-0.5
PC 1
0.5
-0.5
PC 1
0.5

Puzzling: always *"clockwise"*
in the neuron…

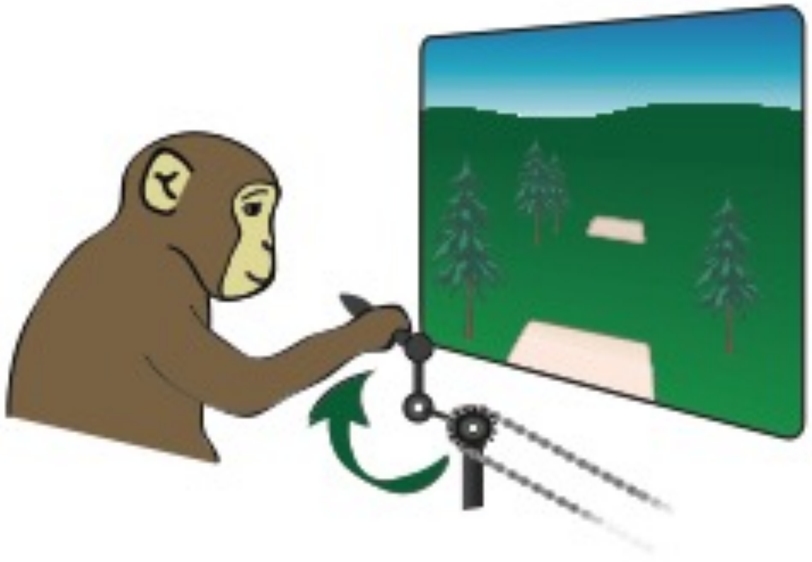
[Precise Estimates of Single-Trial Dynamics in Motor Cortex using Deep Learning Techniques]()

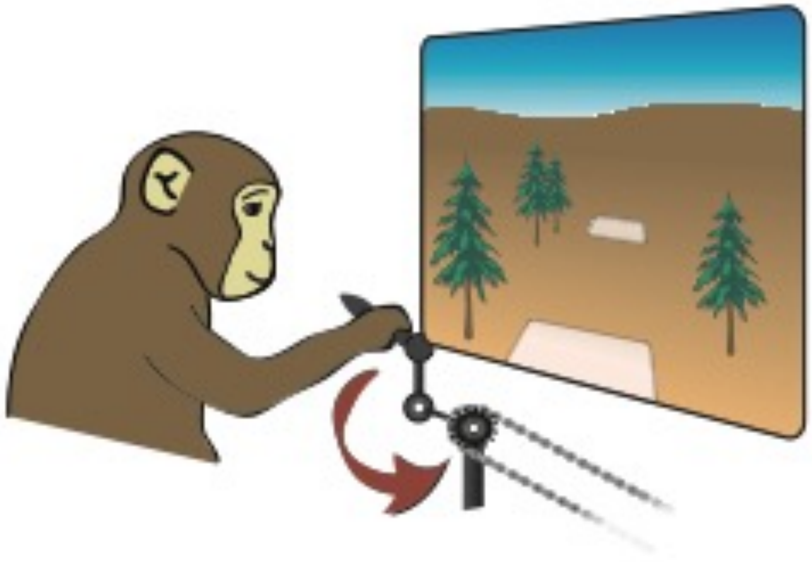
…Mark Churchland, ***Larry Abbott****, …

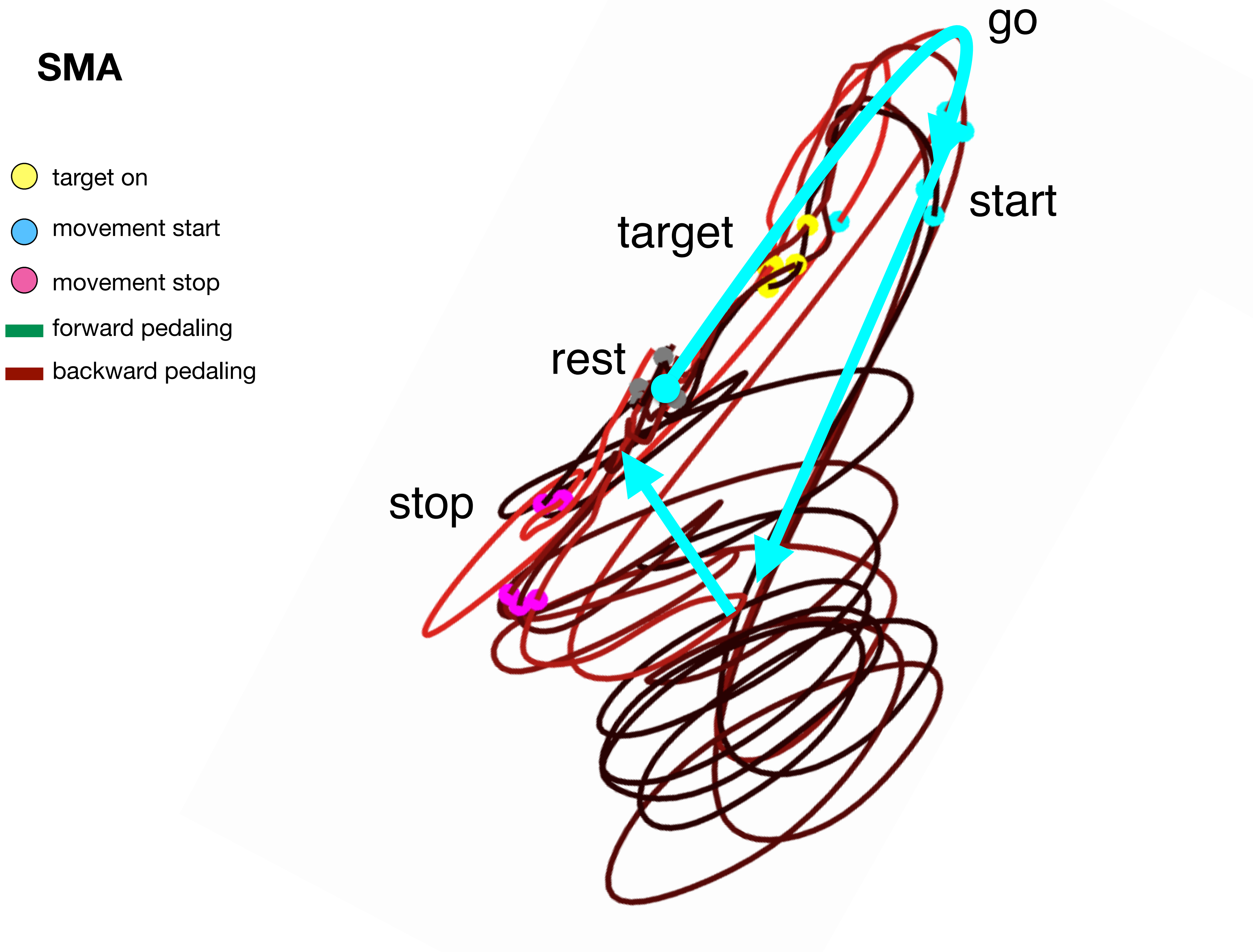


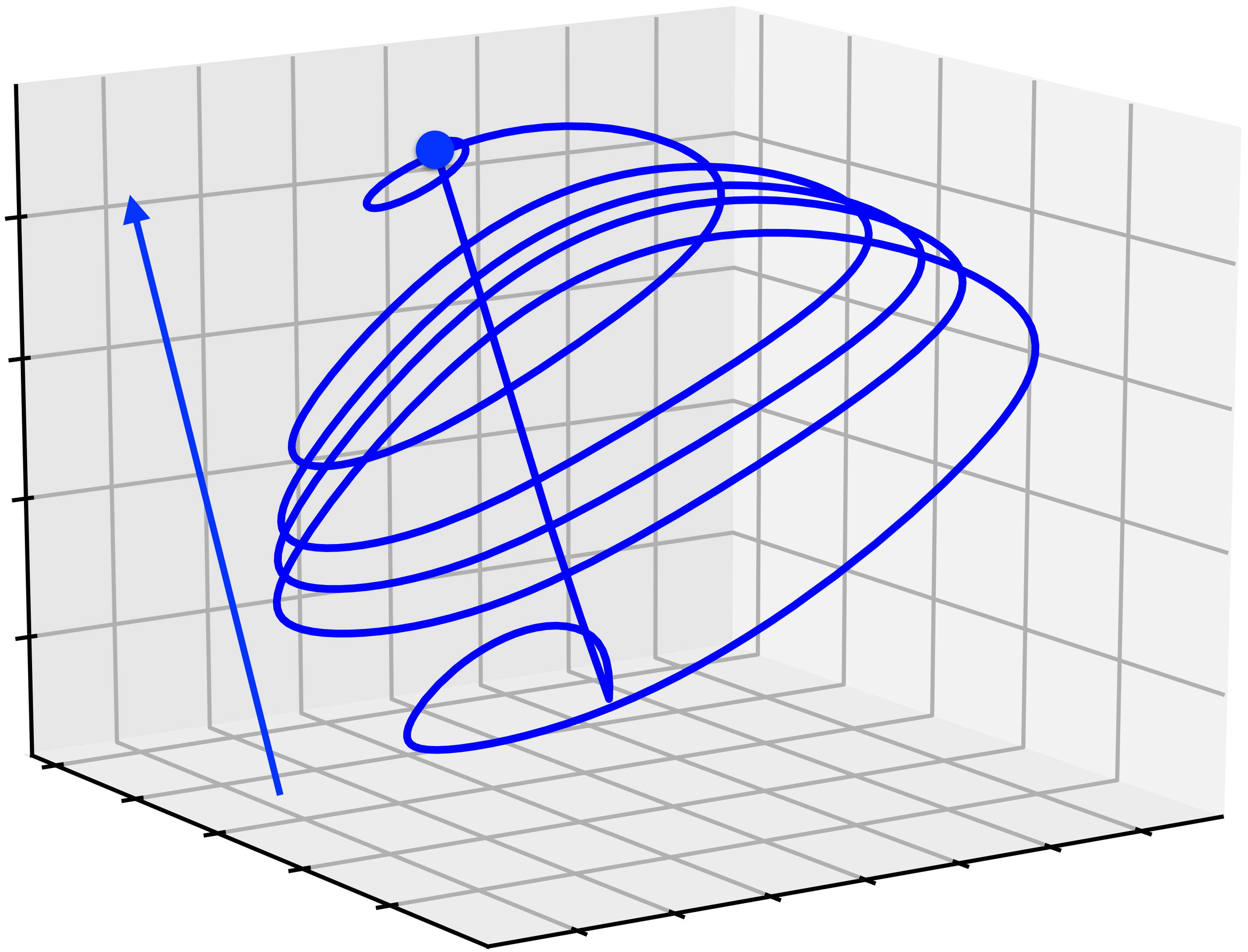
***his slides**

If the target is close,
the monkeys get smarter…



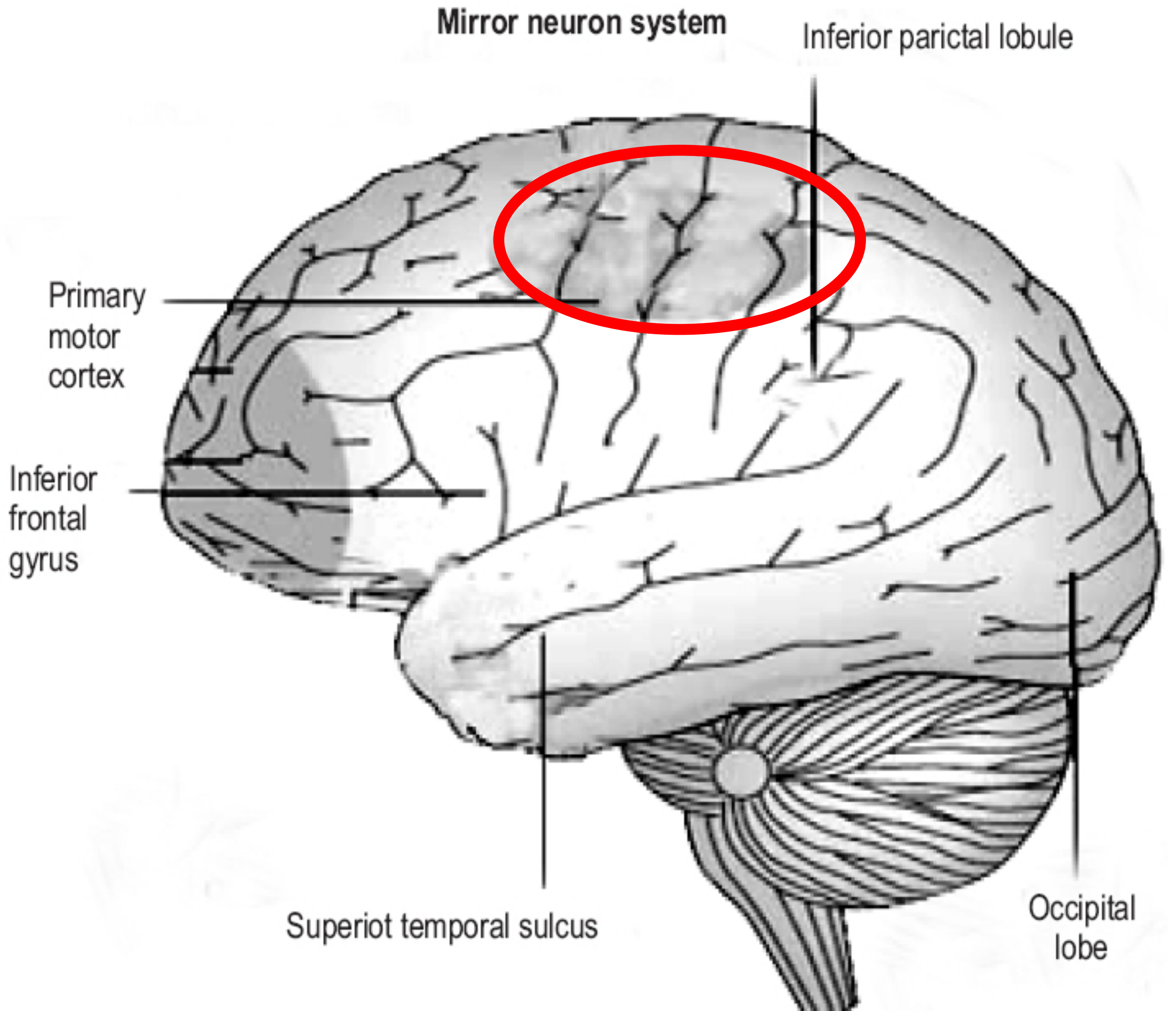



SMA
target on
movement start
movement stop
forward pedaling
backward pedaling
go
start
target
rest
stop



Also in the motor cortex: the mirror neuron system (MNS)…

…and its connections to language…

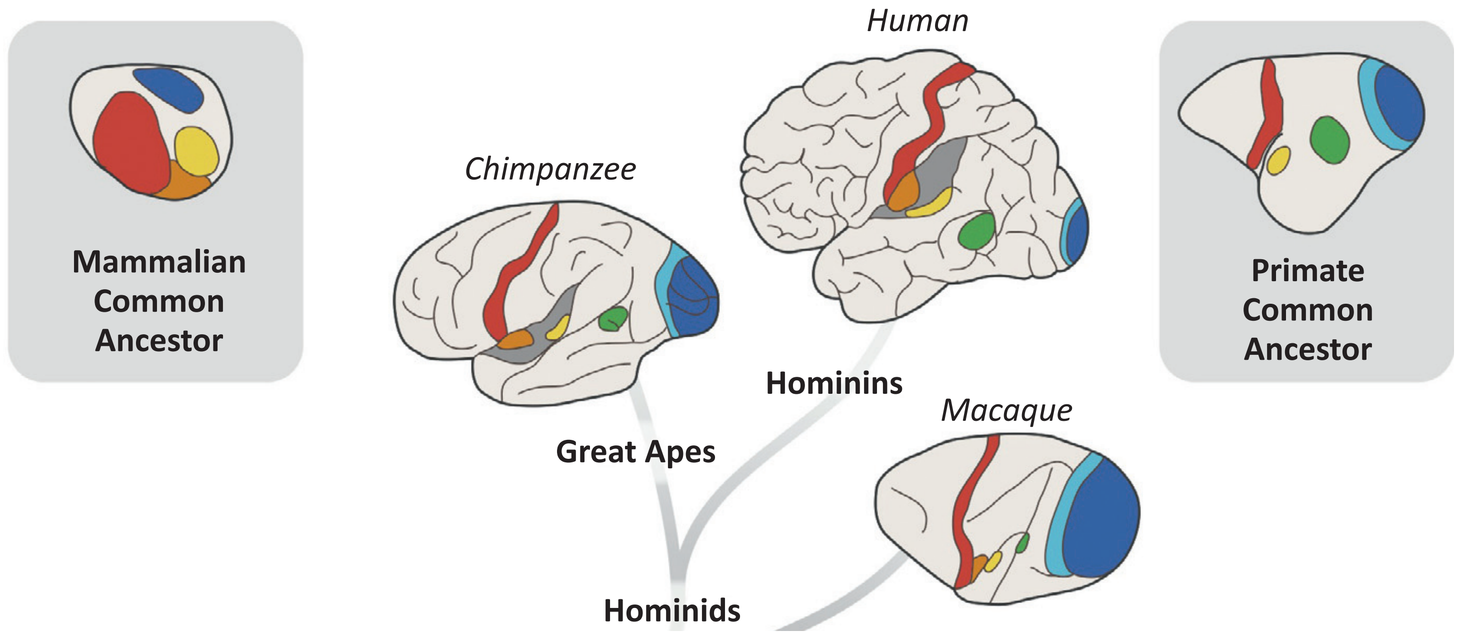


Sooo…

- It seems that (pre?) motor cortex populations learn motions (most likely during childhood)
- The 3D summary of the population dynamics for motions seems to be basically (always?) cyclical and “clockwise” – even if the motions are not…

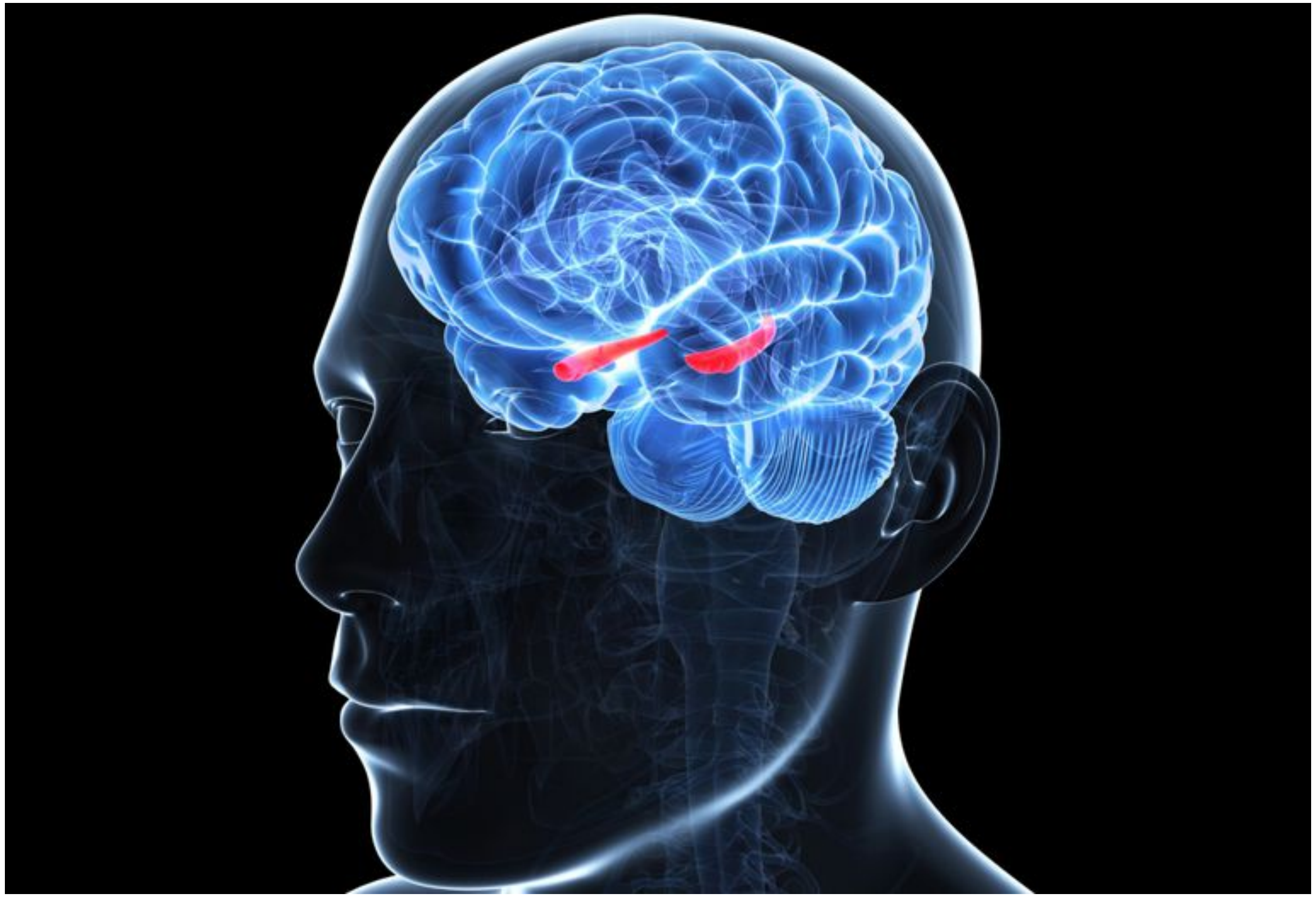
So much about the motor cortex…

- We have talked about:
- Vision
- Olfaction
- Motion
- Somatosensory (one slide…)
- Soon about audition
- ***Where else does computation happen in the Brain?***

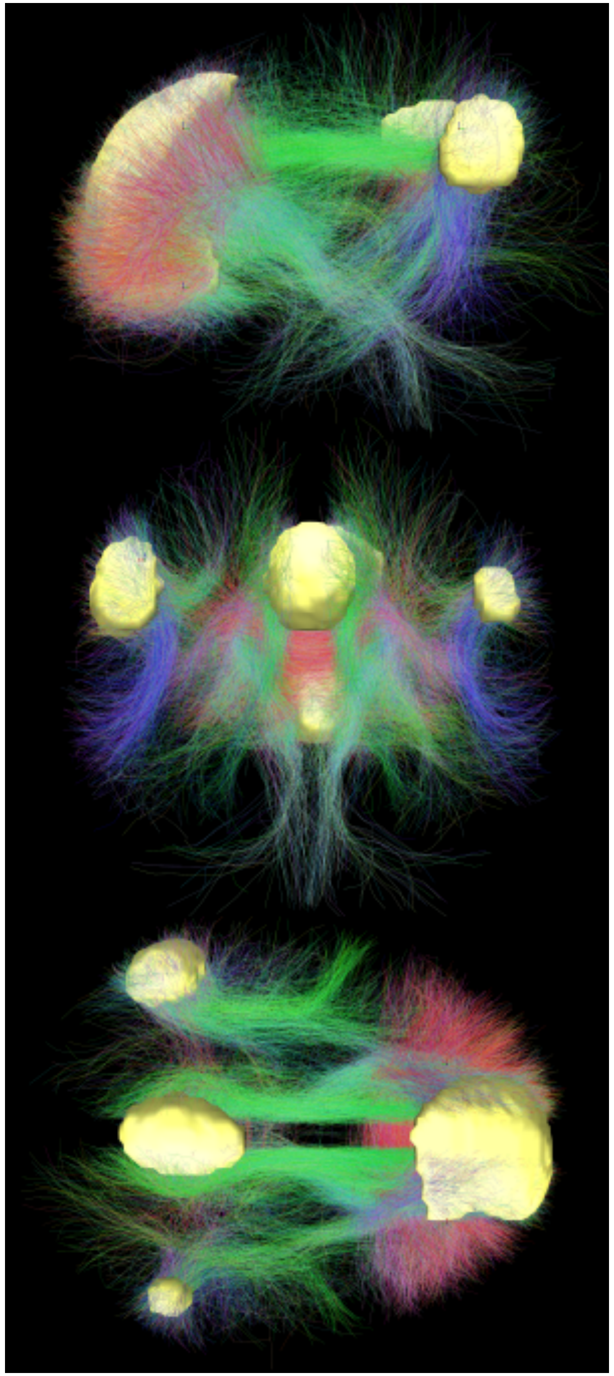


The rise of the ***association cortex***
(= neocortex excluding sensory and motor areas, and the DMN)

Hippocampus: the gateway of the AC



...and the default mode network (DMN)...



Oscillations in the Brain

frequency bands

delta (2–4 Hz)

theta (4–8 Hz)

low *gamma* (40–70 Hz)

high *gamma* (70–150 Hz)

default mode: 0.01 hz