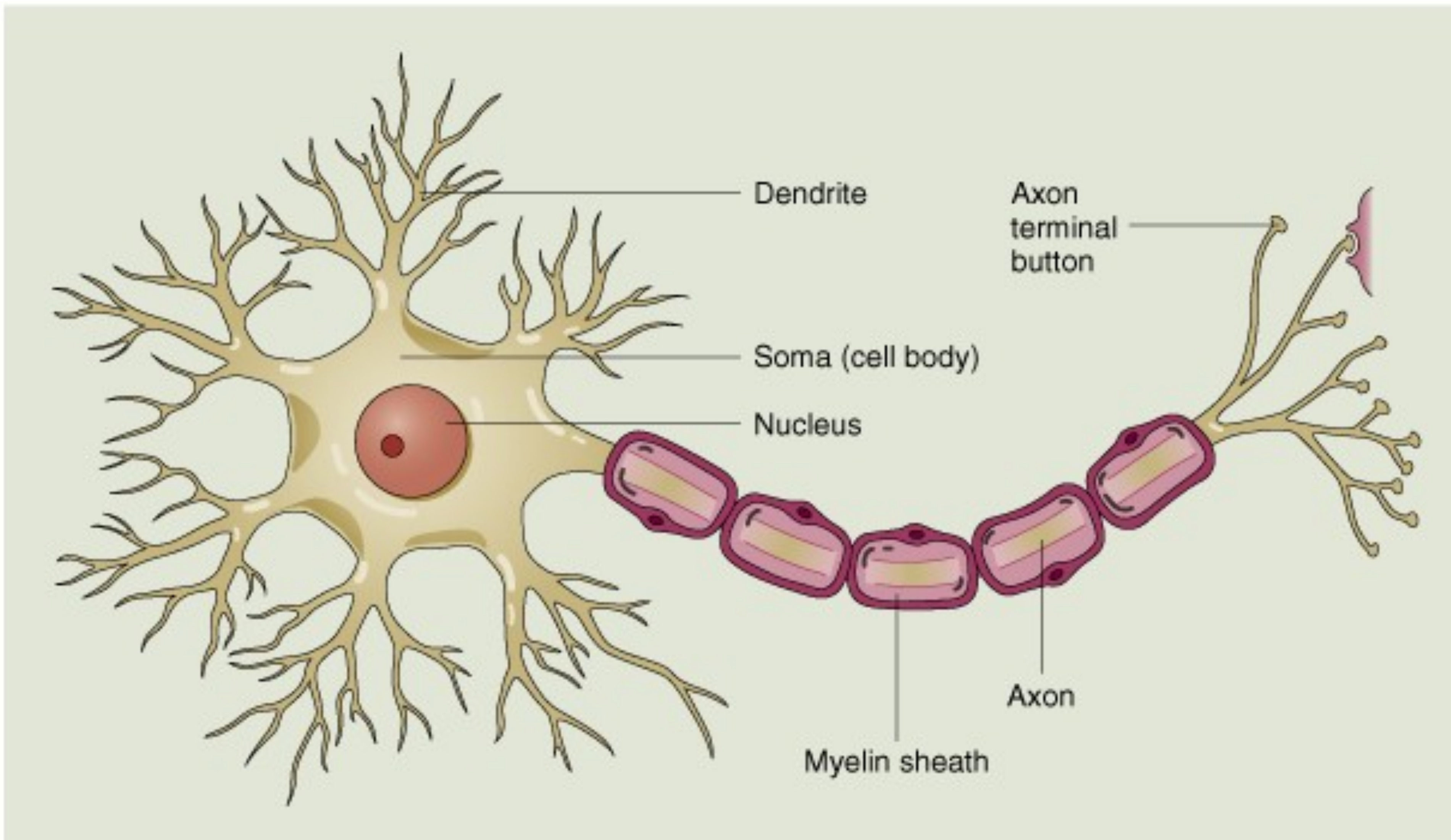
# Theoretical Neuroscience: Information theory and statistics

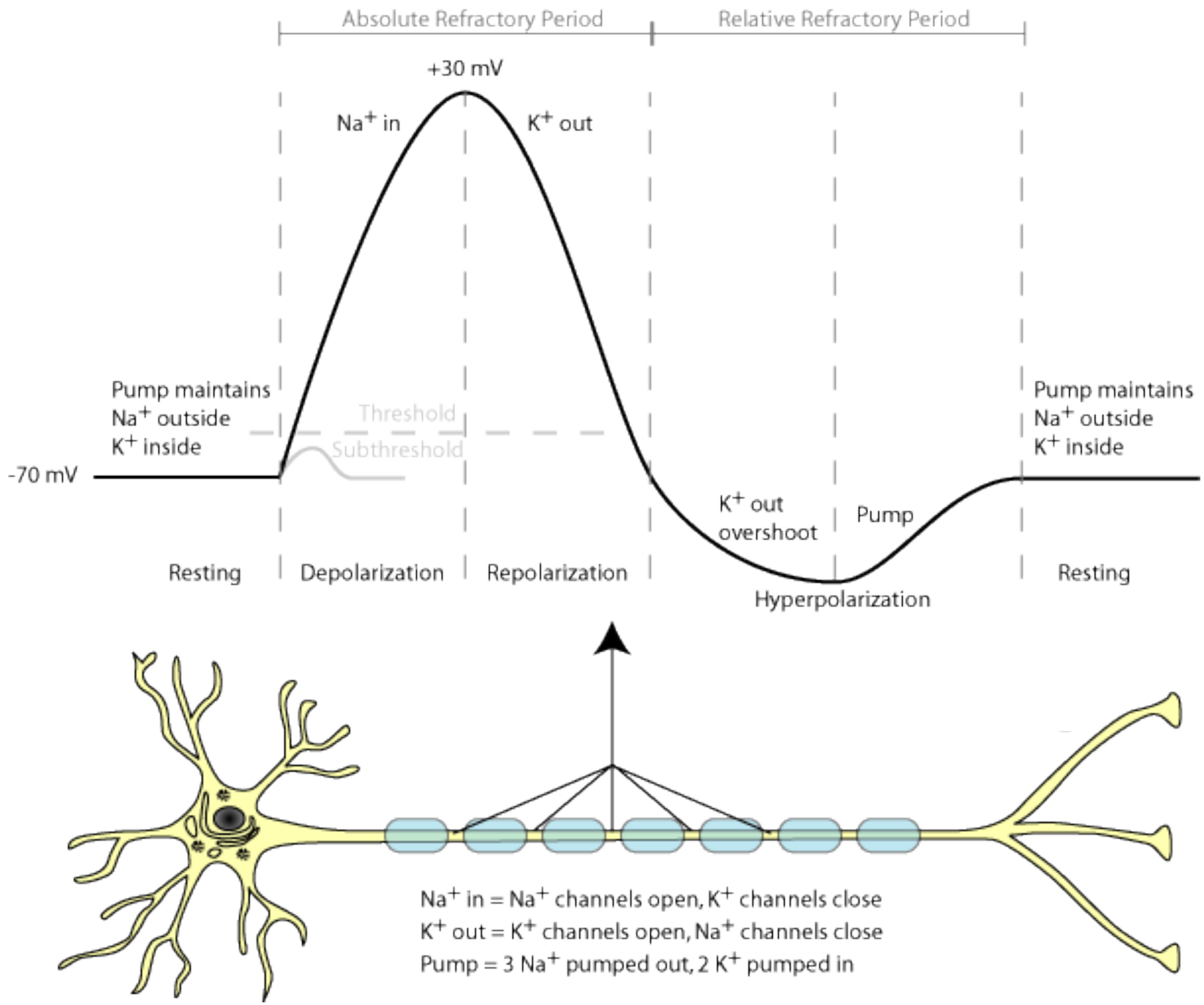
**Lecturer:**
Stephanie E. Palmer
Associate Professor
Dep't of OBA
Dep't of Physics
University of Chicago
sepalmer@uchicago.edu

$$
\begin{tikzcd}[column sep=small]
\text{stimuli} \arrow[r] & \text{neurons} \arrow[dr] & & & & \\
& & \text{neurons} \arrow[dr] & & & \\
& & & \text{neurons} \arrow[dr] & & \\
& & & & \cdots \arrow[dr] & \\
& & & & & \text{neurons} \arrow[r] & \text{behavior}
\end{tikzcd}
$$

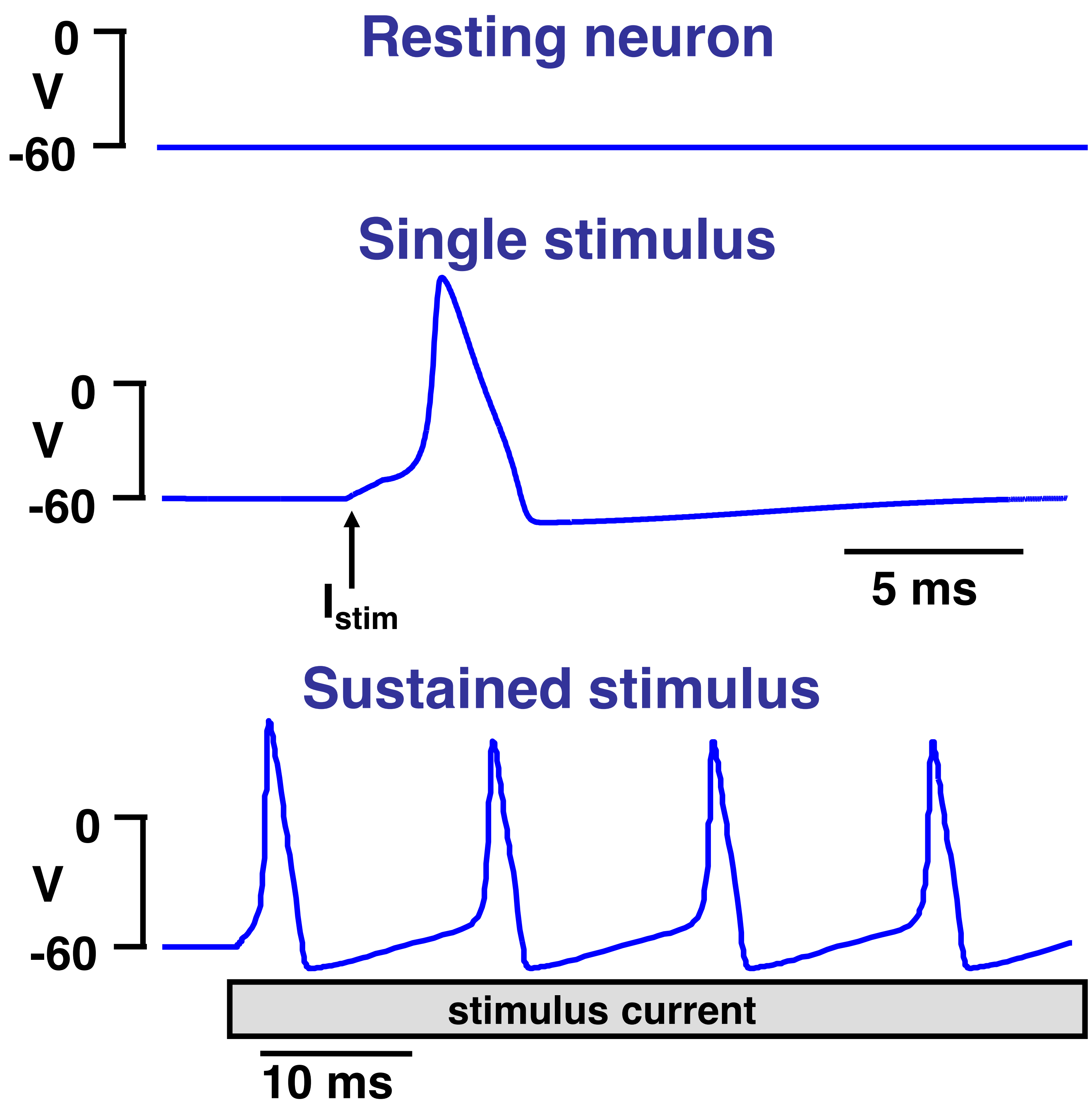
# Anatomy of a neuron:



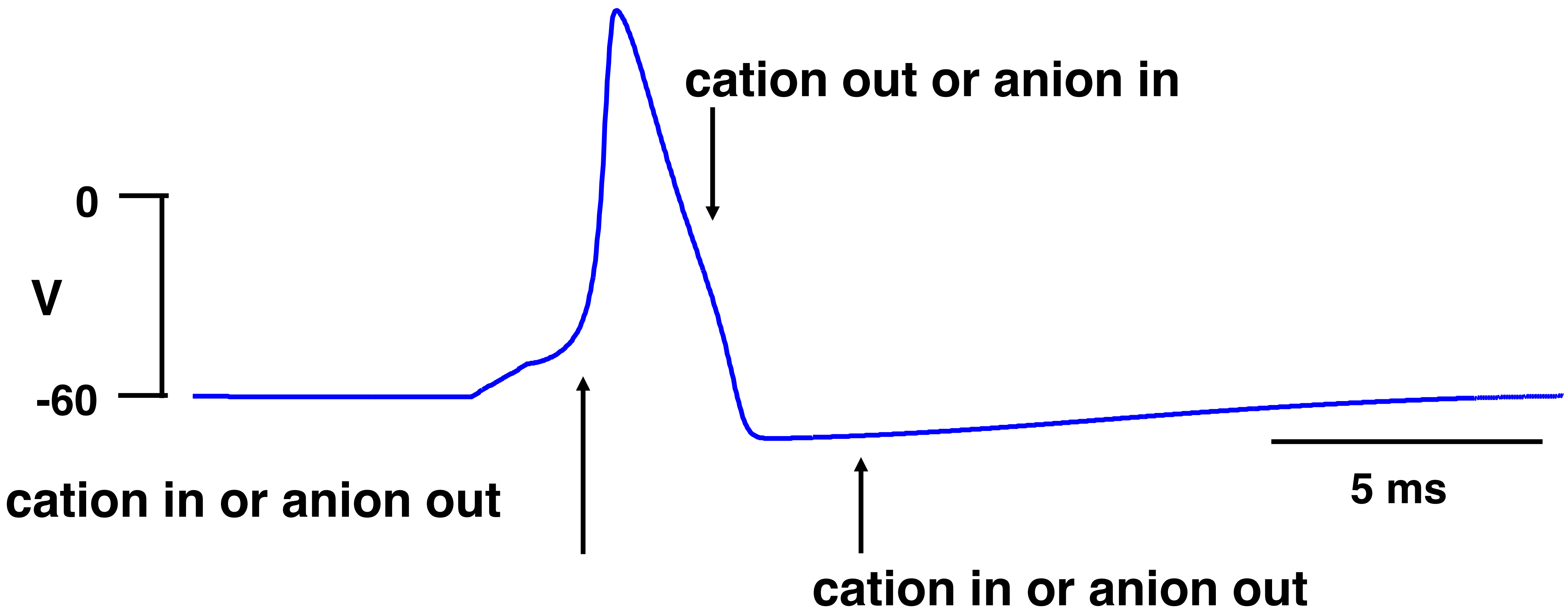
# Anatomy of a spike:



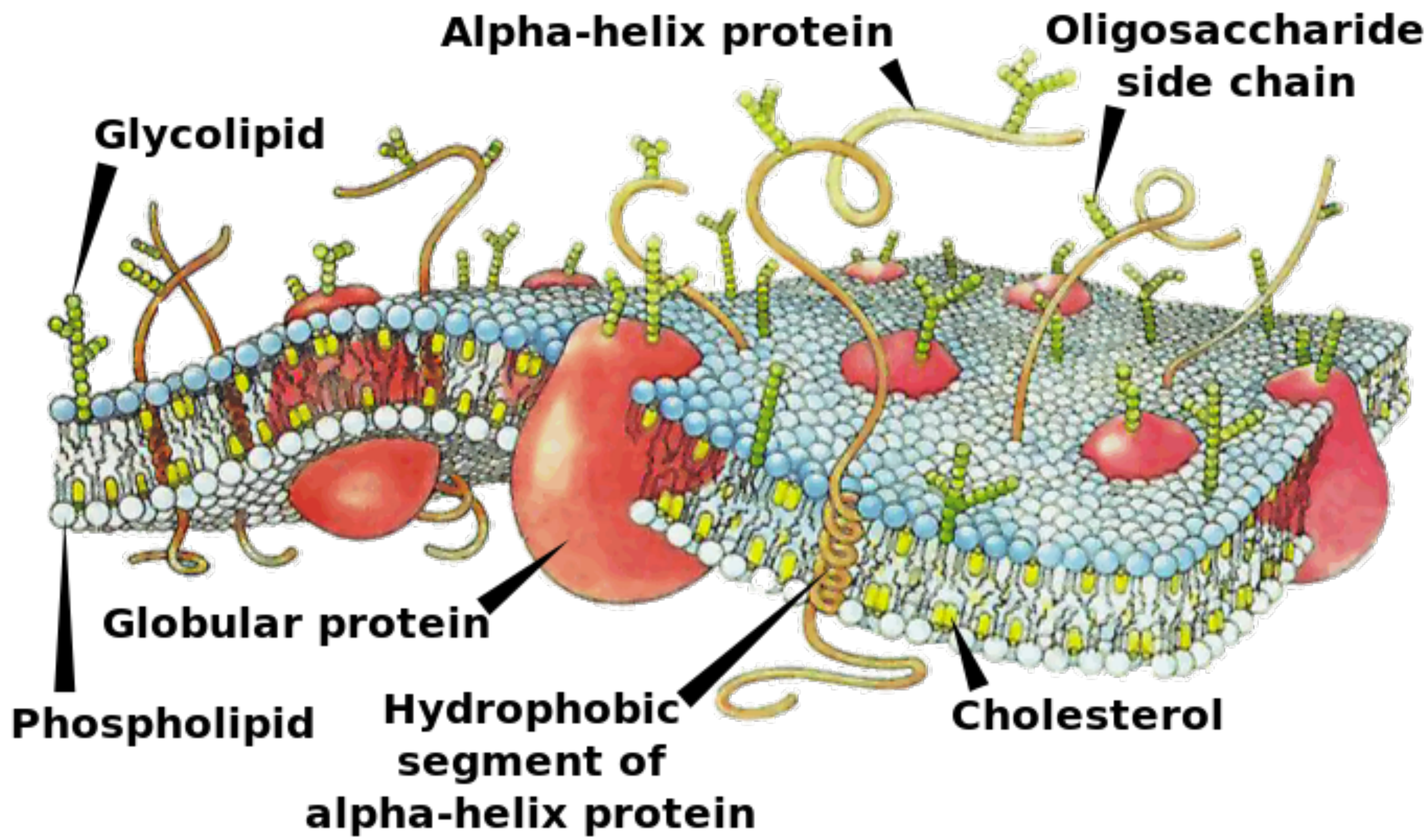
# Properties of spiking:



We seek a quantitative description of this behavior:

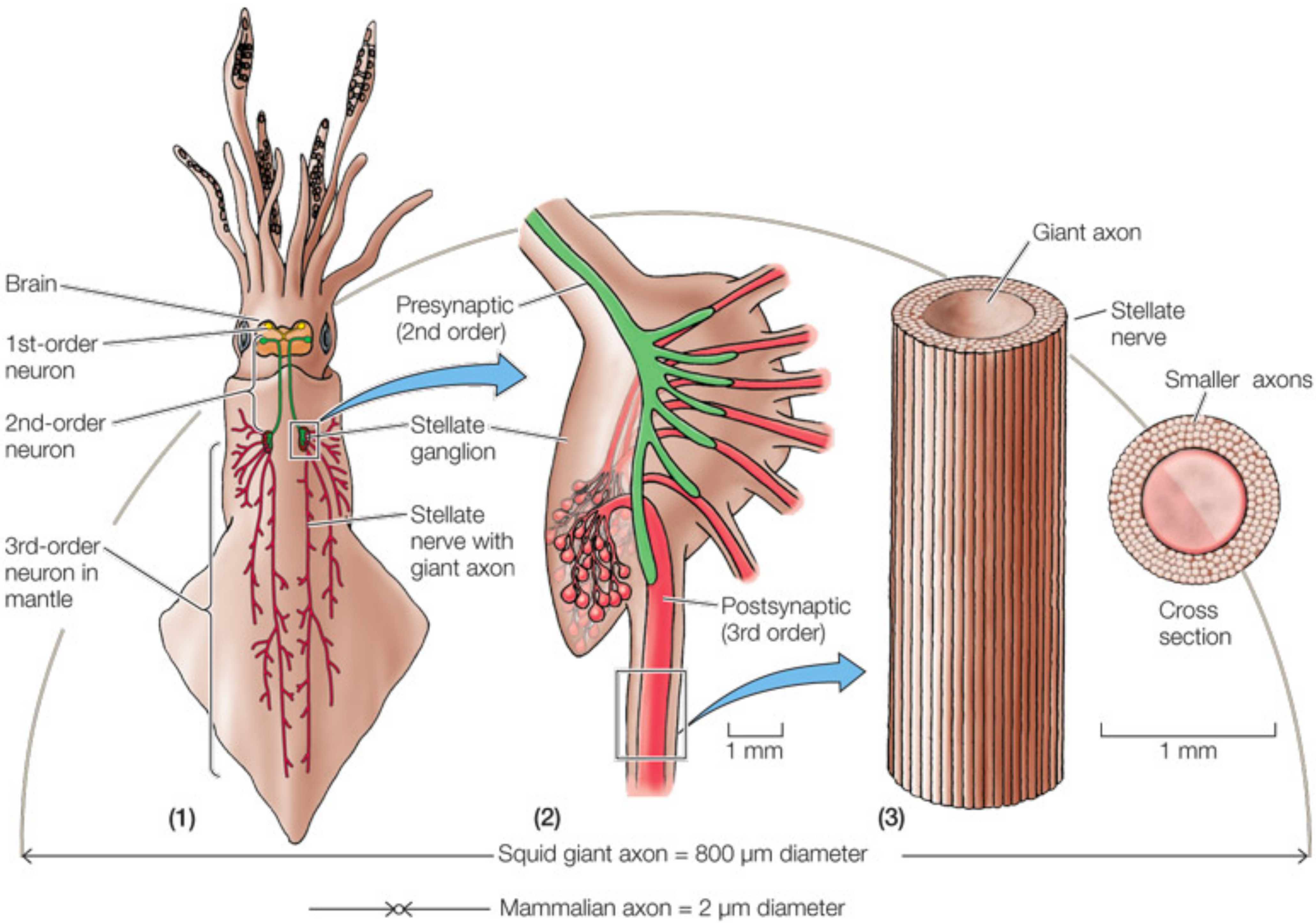


# Some basics:

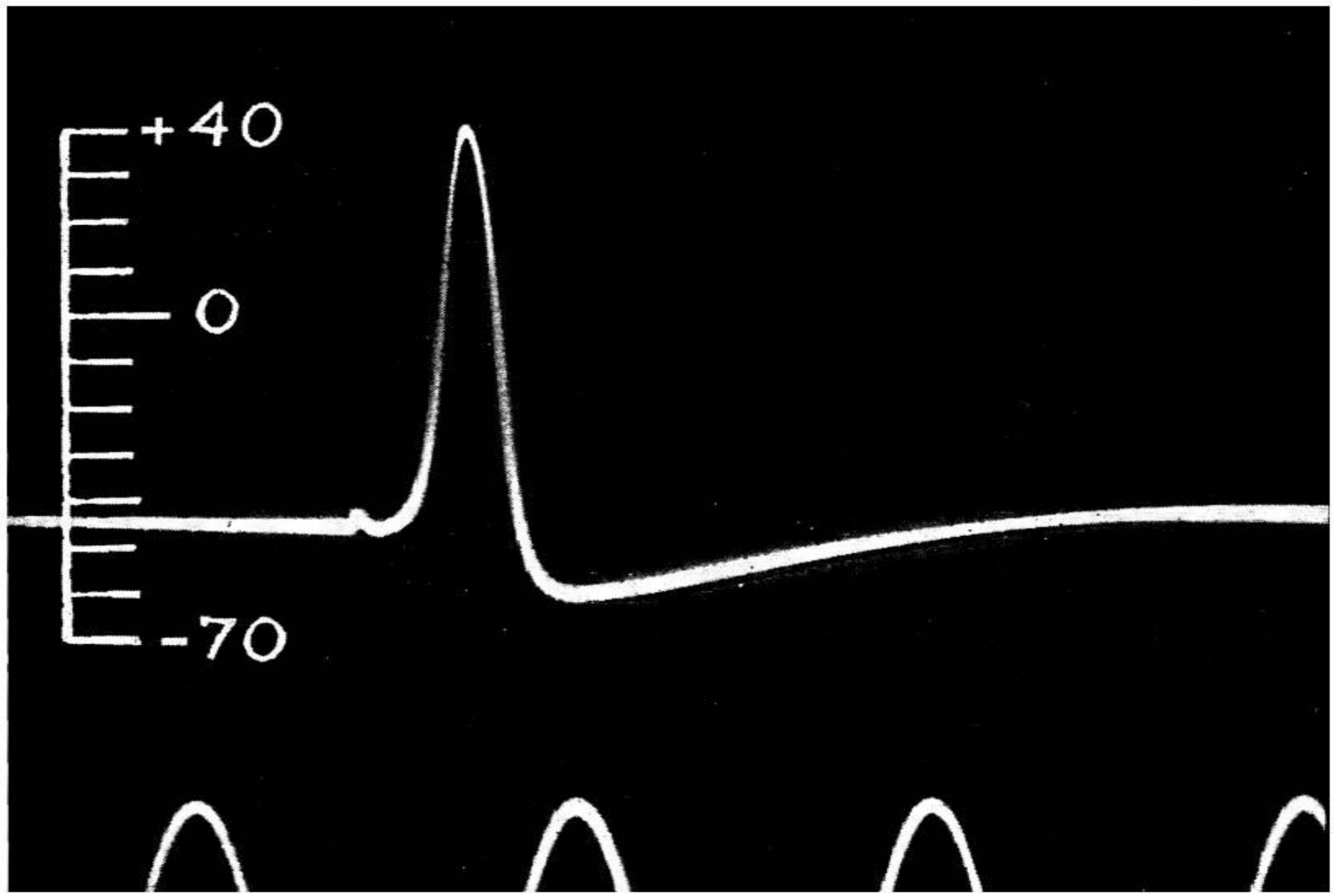



chalkboard interlude

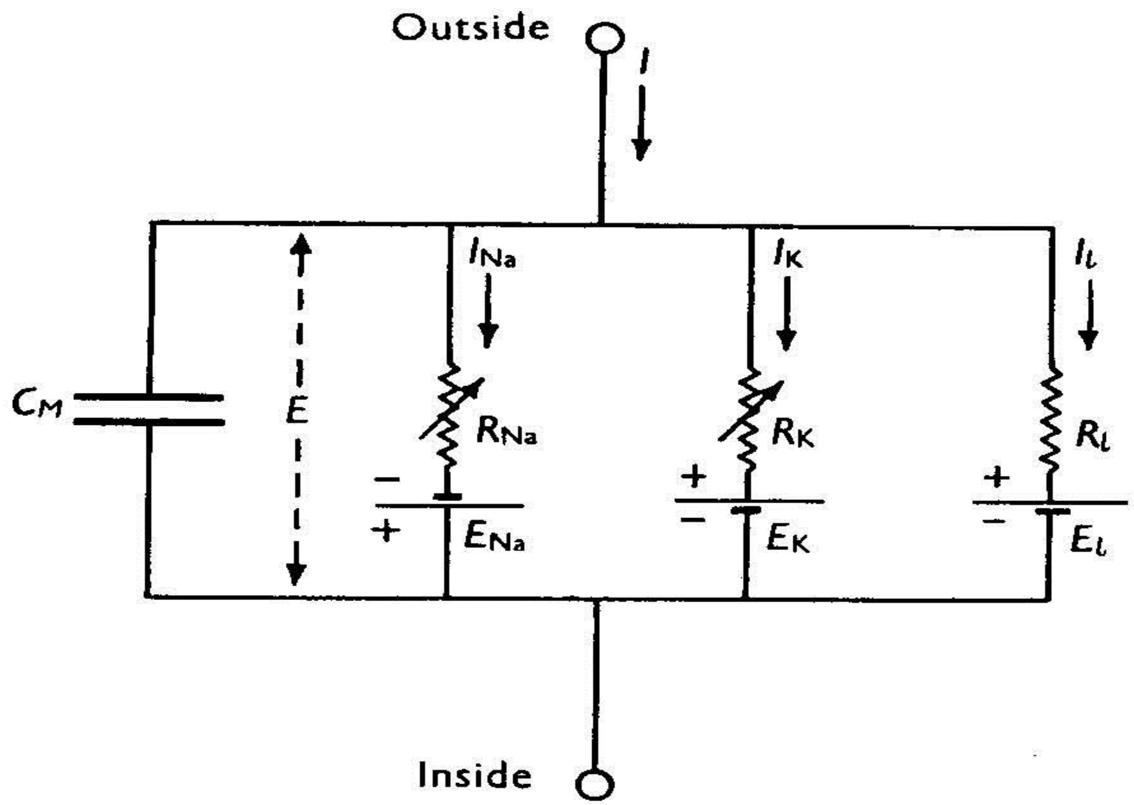
# Squid giant axon:



# The 1939 letter to Nature:



# The final model:




**Hodgkin & Huxley (1952), *J. Physiol.* 117:400.**

$$C_m \frac{dV}{dt} = -g_L(V - V_L) - \bar{g}_{Na} m^3 h (V - V_{Na}) - \bar{g}_K n^4 (V - V_K)$$

$$\frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m$$

$$\frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h$$

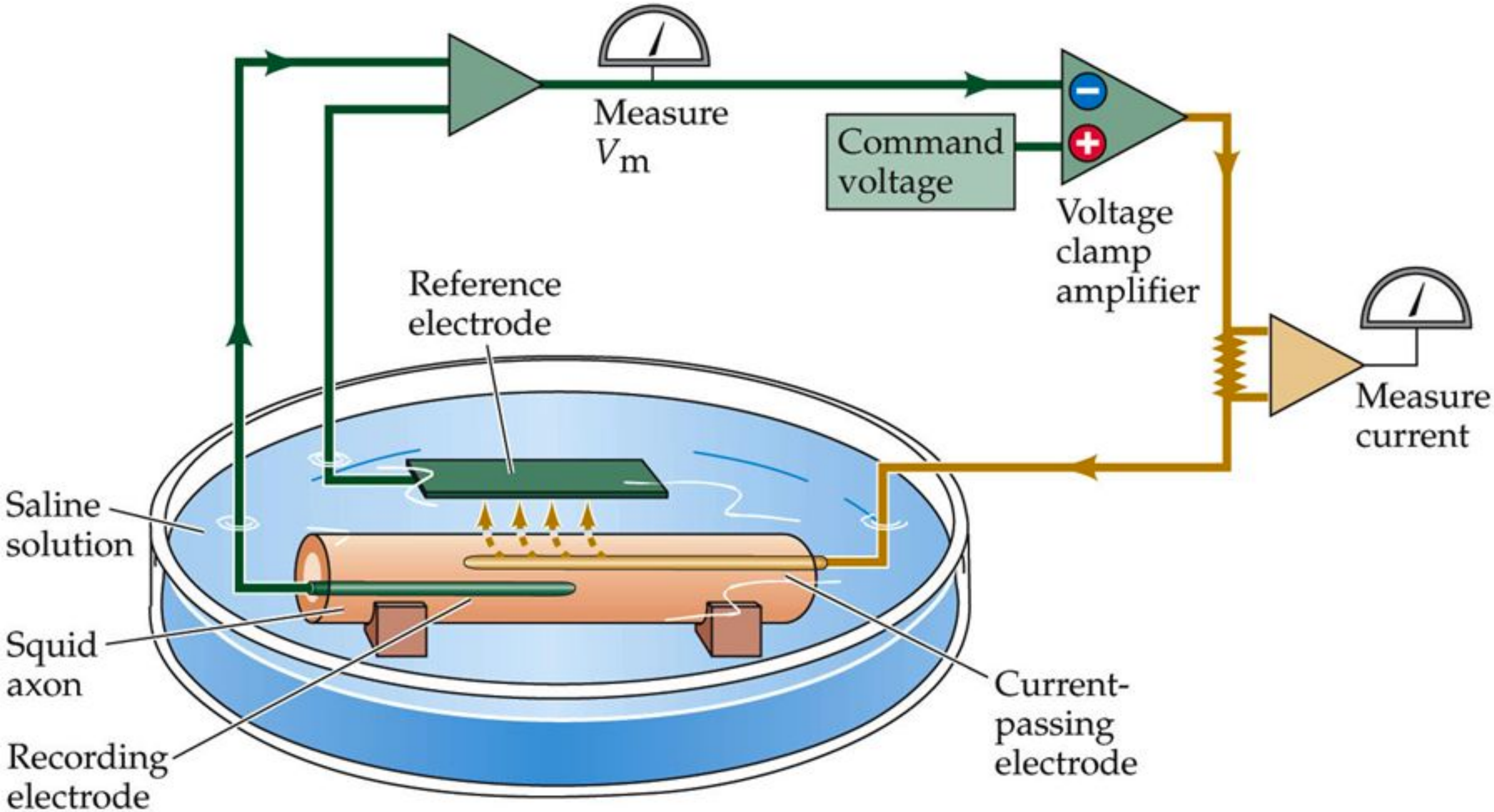
$$\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n$$

$$\alpha_m = 0.1(V_m + 35.0)/(1. - e^{(-(V_m+35.0)/10.0)})$$
$$\beta_m = 4.0\, e^{(-(V_m+60.0)/18.0)}$$

$$\alpha_h = 0.07\, e^{(-(V_m+60.0)/20.0)}$$
$$\beta_h = 1./(1 + e^{(-(V_m+30.0)/10.0)})$$

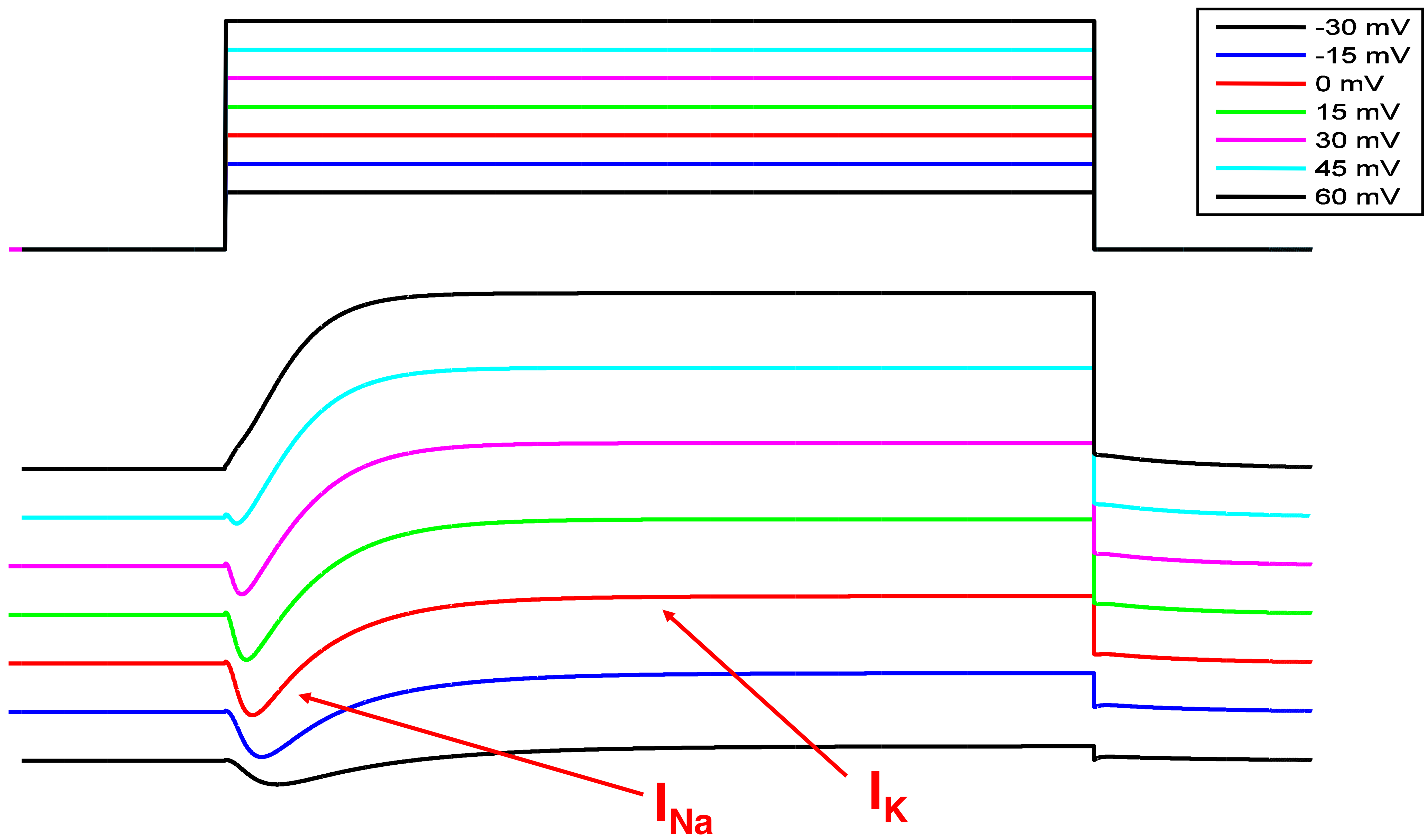
$$\alpha_n = 0.01(V_m + 50.0)/(1 - e^{(-(V_m+50.0)/10.0)})$$
$$\beta_n = 0.125\, e^{(-(V_m+60.0)/80.0)}$$

# Box 3A The Voltage Clamp Technique

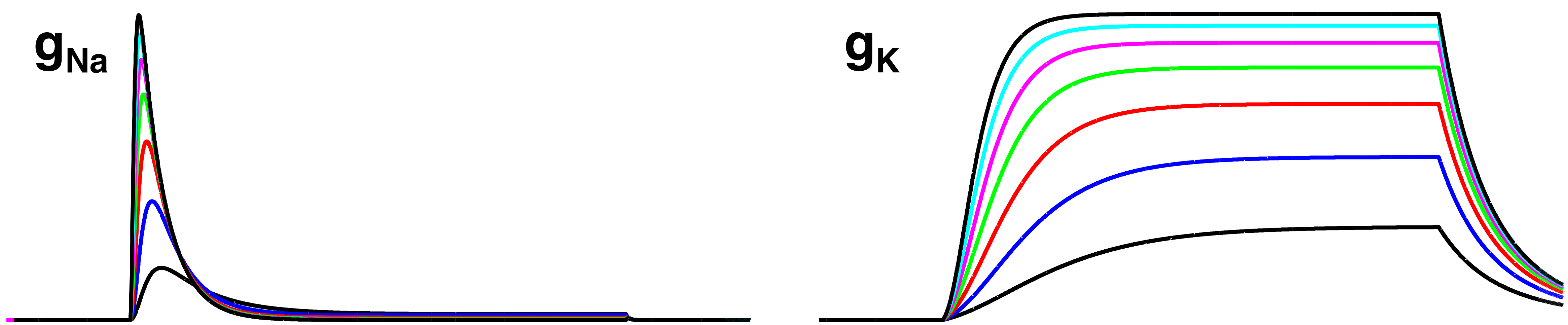



chalkboard interlude

# Make recordings, separate currents:

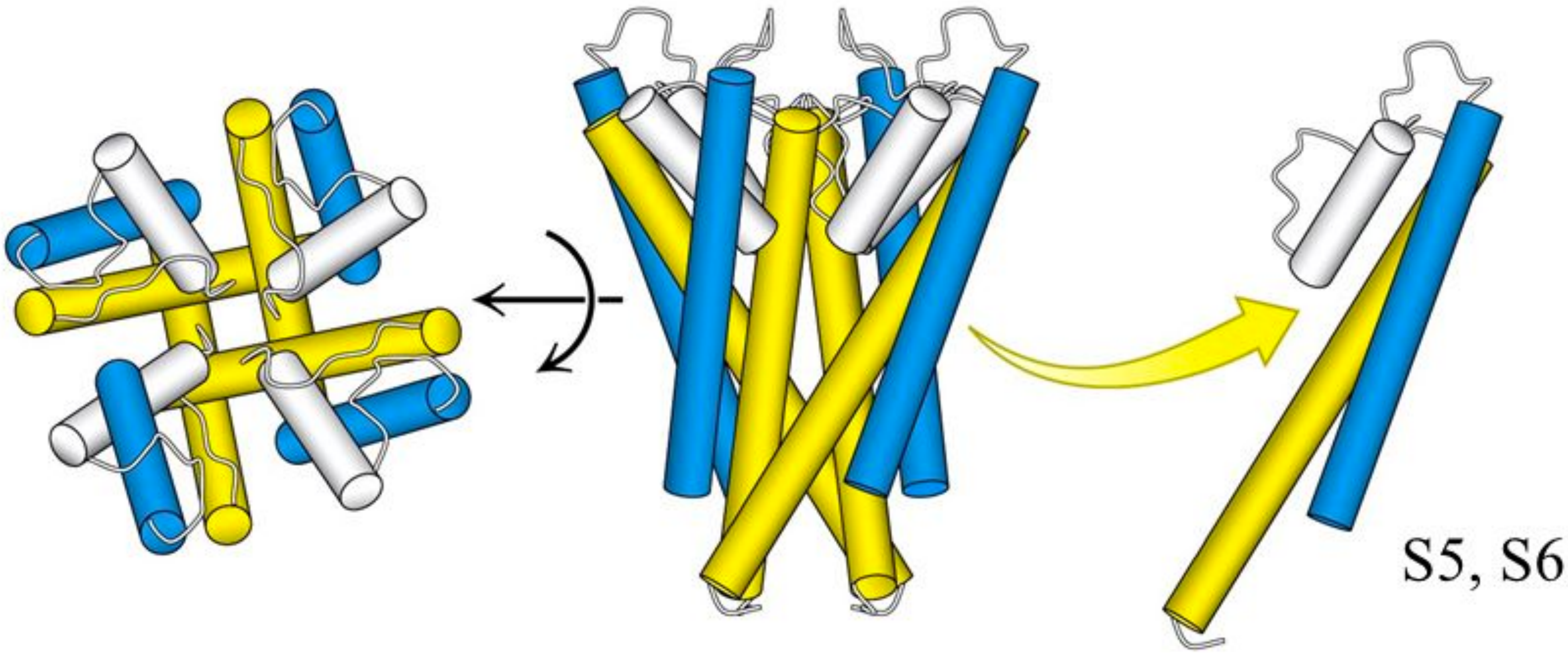


…and get conductances:




chalkboard interlude

# Structure of the potassium ion channel (tetramer)

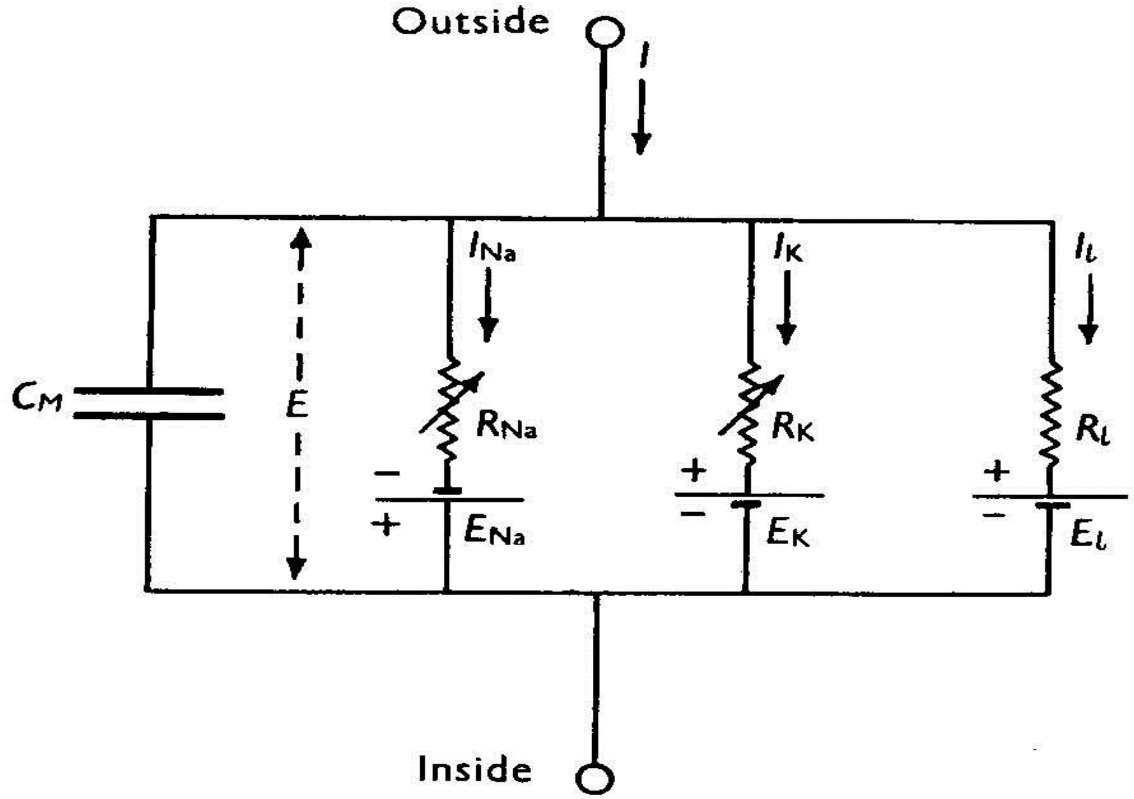



View down the pore

A single subunit

**Figure 13-17**
***Biochemistry, Sixth Edition***

# The final model:




**Hodgkin & Huxley (1952), *J. Physiol.* 117:400.**

$$C_m \frac{dV}{dt} = -g_L(V - V_L) - \overline{g}_{Na} m^3 h (V - V_{Na}) - \overline{g}_K n^4 (V - V_K)$$

$$\frac{dm}{dt} = \alpha_m(V)(1-m) - \beta_m(V)m$$

$$\frac{dh}{dt} = \alpha_h(V)(1-h) - \beta_h(V)h$$

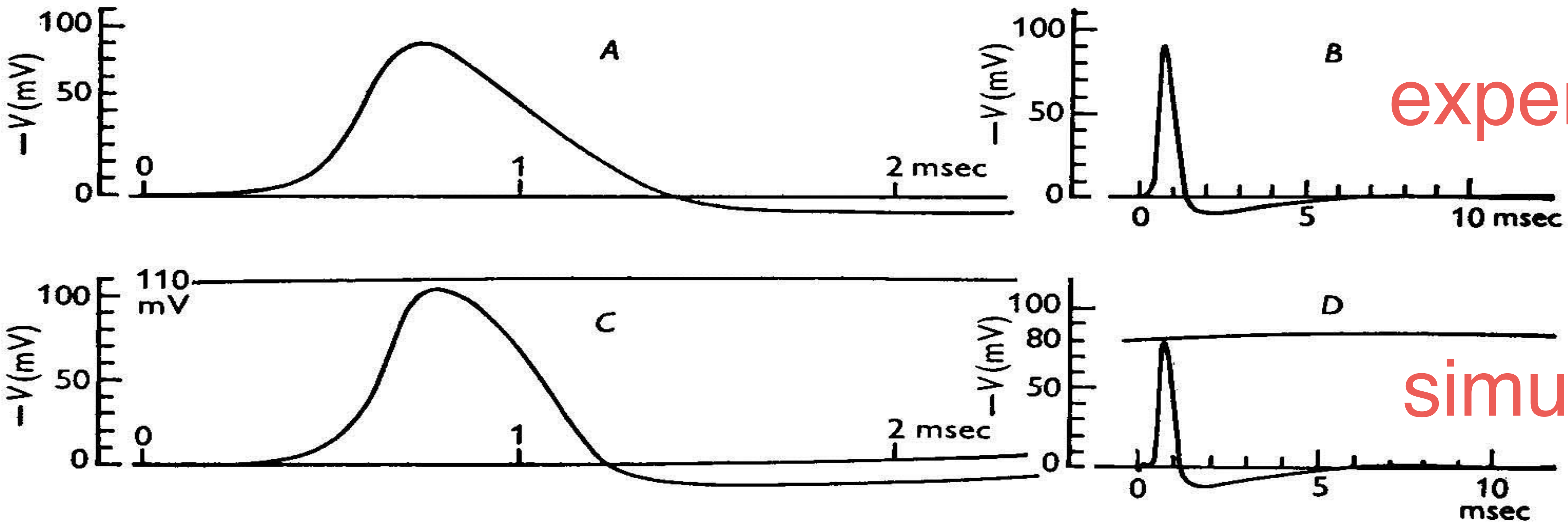
$$\frac{dn}{dt} = \alpha_n(V)(1-n) - \beta_n(V)n$$

$$\alpha_m = 0.1(V_m + 35.0)/(1. - e^{(-(V_m+35.0)/10.0)})$$
$$\beta_m = 4.0\, e^{(-(V_m+60.0)/18.0)}$$

$$\alpha_h = 0.07\, e^{(-(V_m+60.0)/20.0)}$$
$$\beta_h = 1./(1 + e^{(-(V_m+30.0)/10.0)})$$

$$\alpha_n = 0.01(V_m + 50.0)/(1 - e^{(-(V_m+50.0)/10.0)})$$
$$\beta_n = 0.125\, e^{(-(V_m+60.0)/80.0)}$$

…works well:




experimen

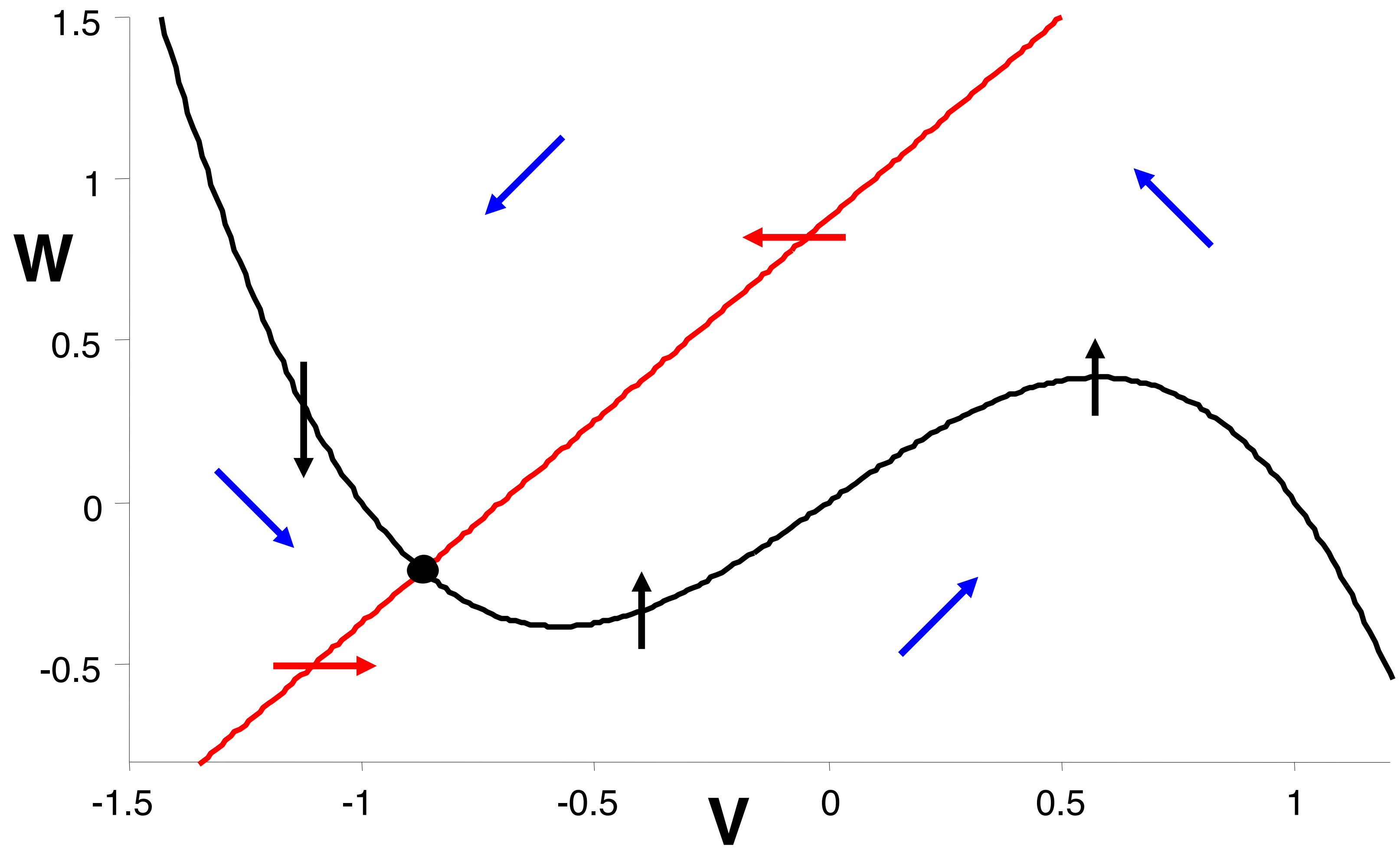
simulation

# The Fitzhugh-Nagumo model:

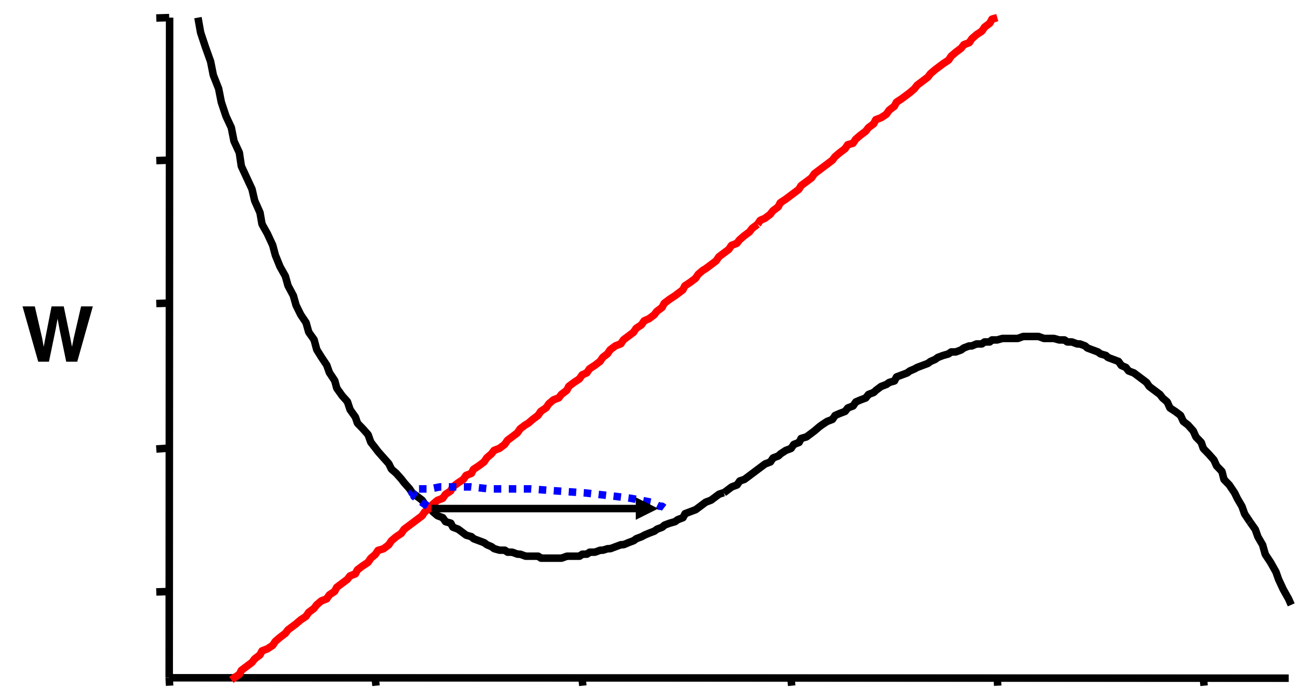
$$dV/dt = V - V^3 - W - I$$

$$dW/dt = 0.08*(V + 0.7-0.8W)$$

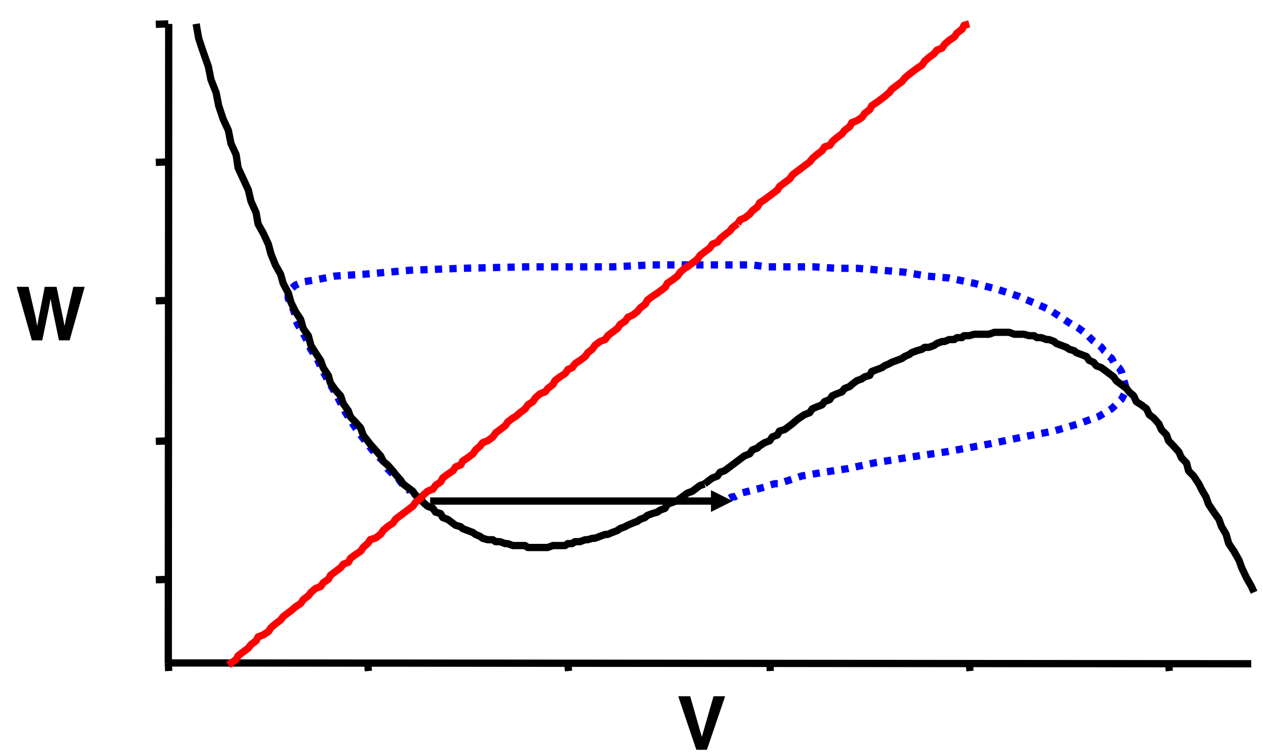
# The Fitzhugh-Nagumo model is a useful



# The Fitzhugh-Nagumo model is a useful

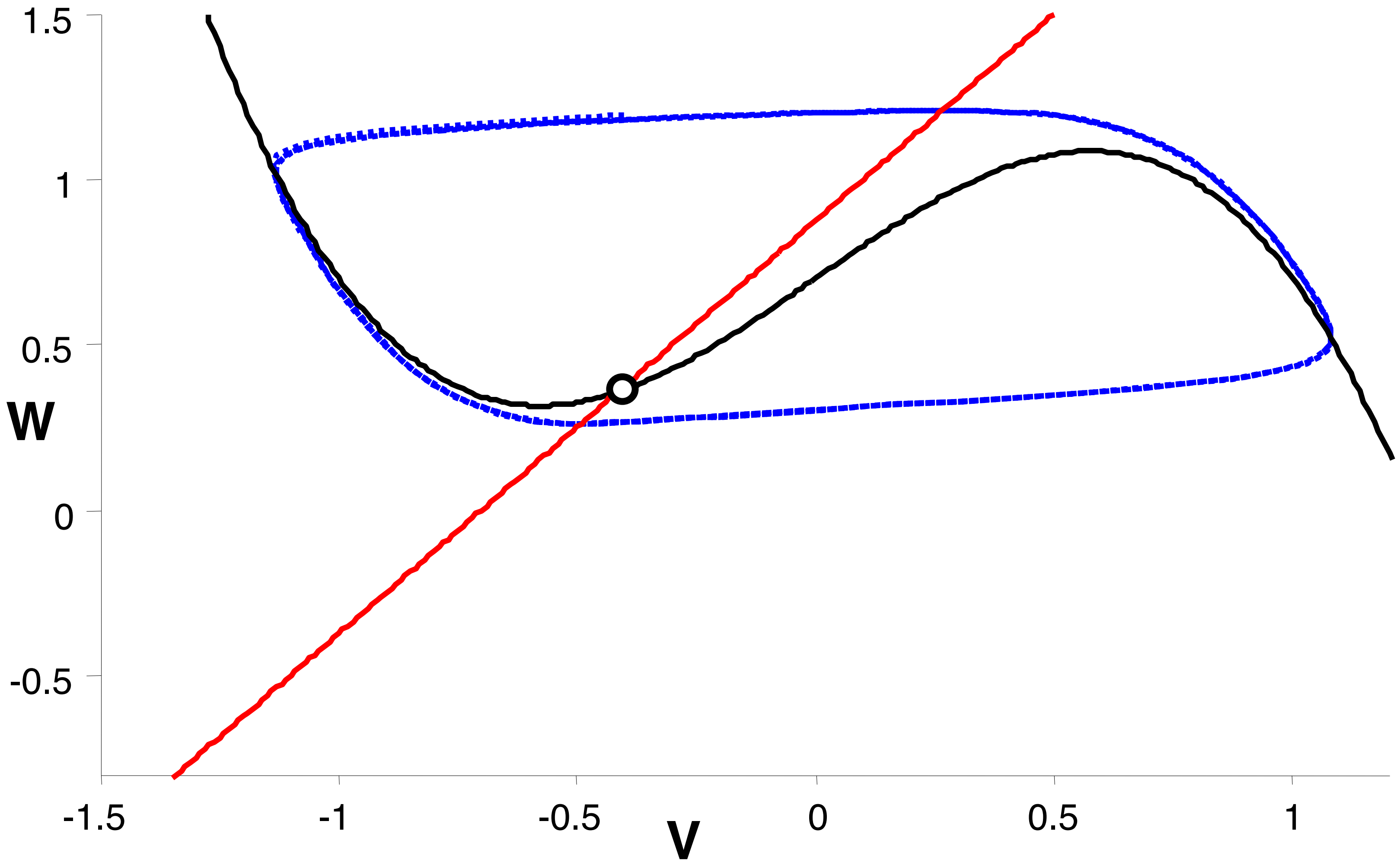



return to baseline

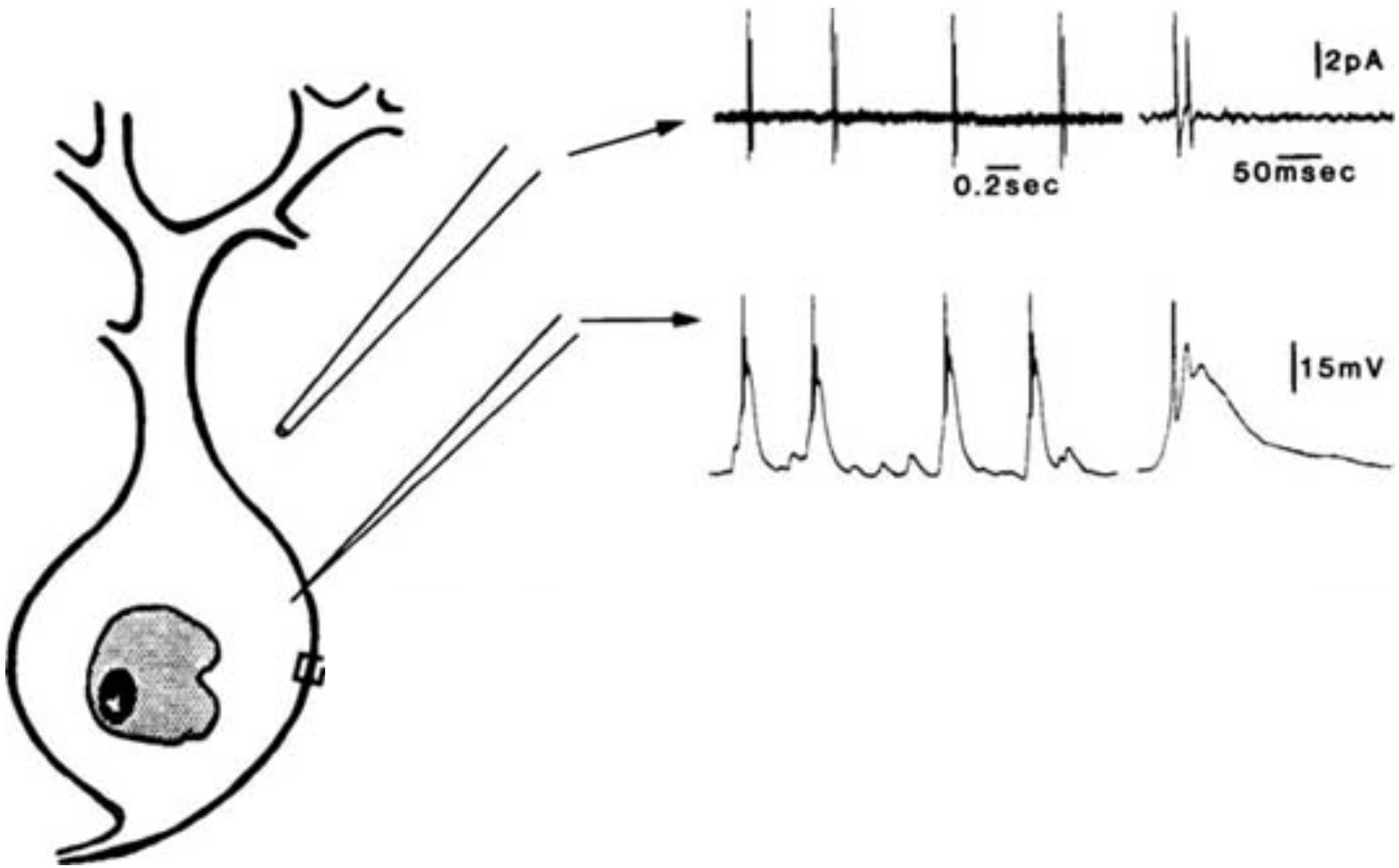



spike!

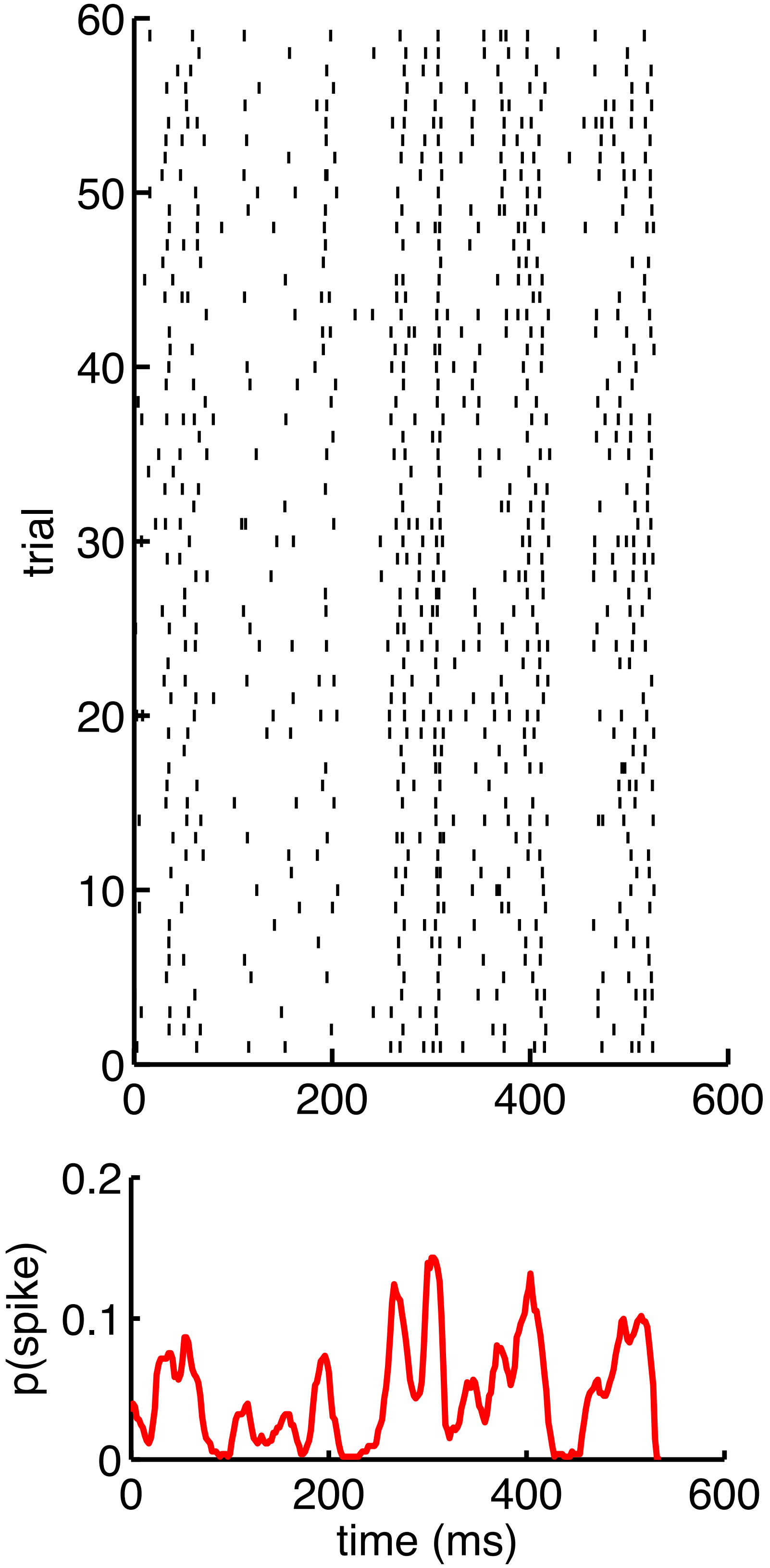
# Even explains sustained firing!



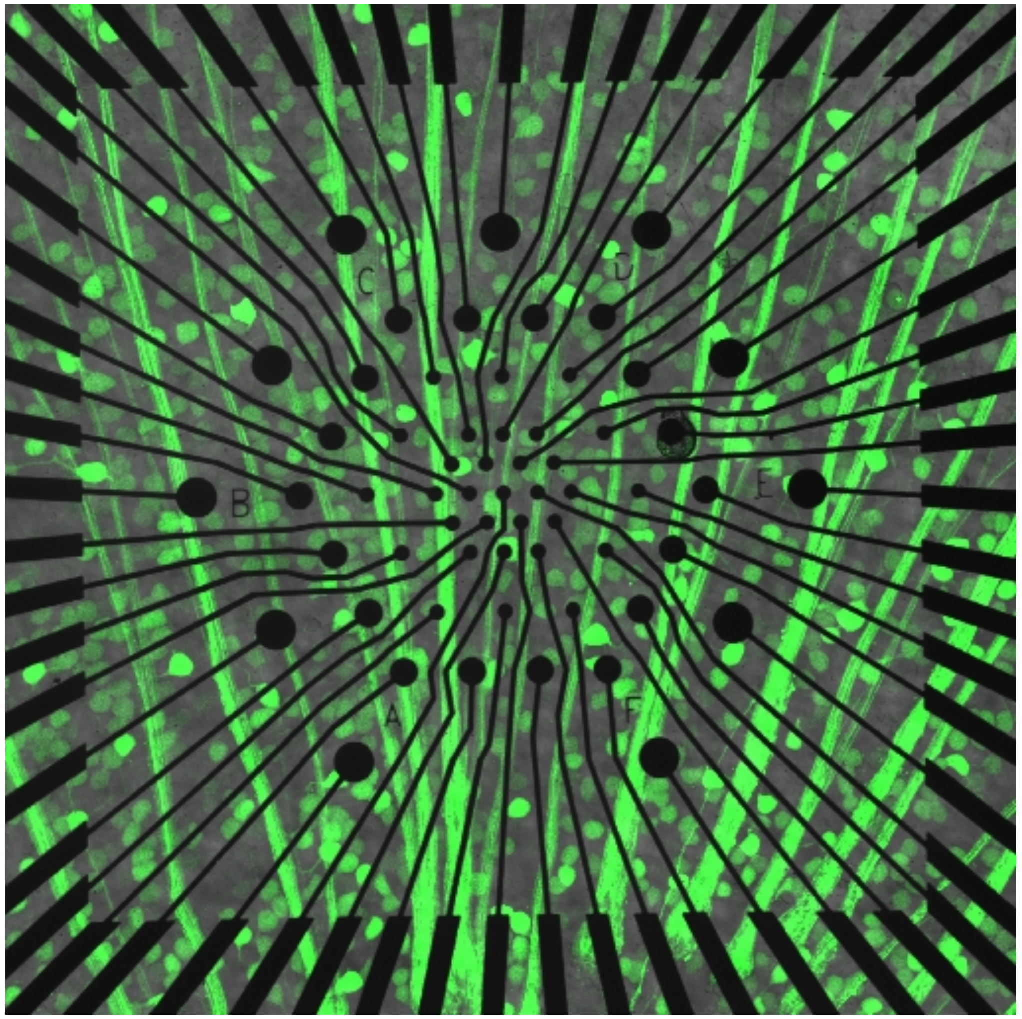
# Extracellular recording:

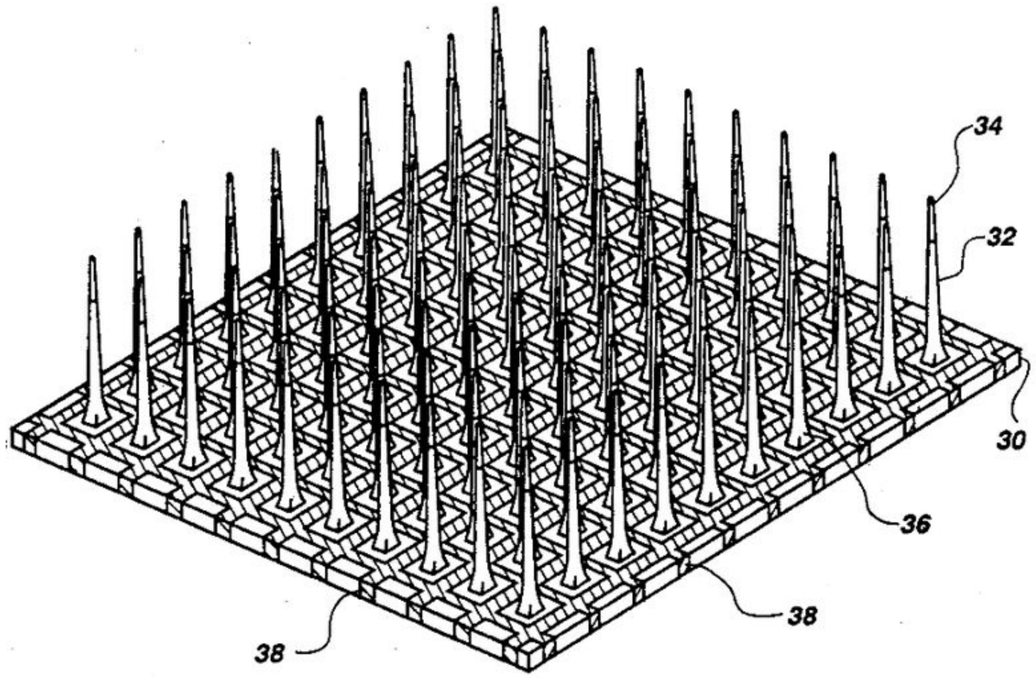


# Neuronal response is 'noisy':



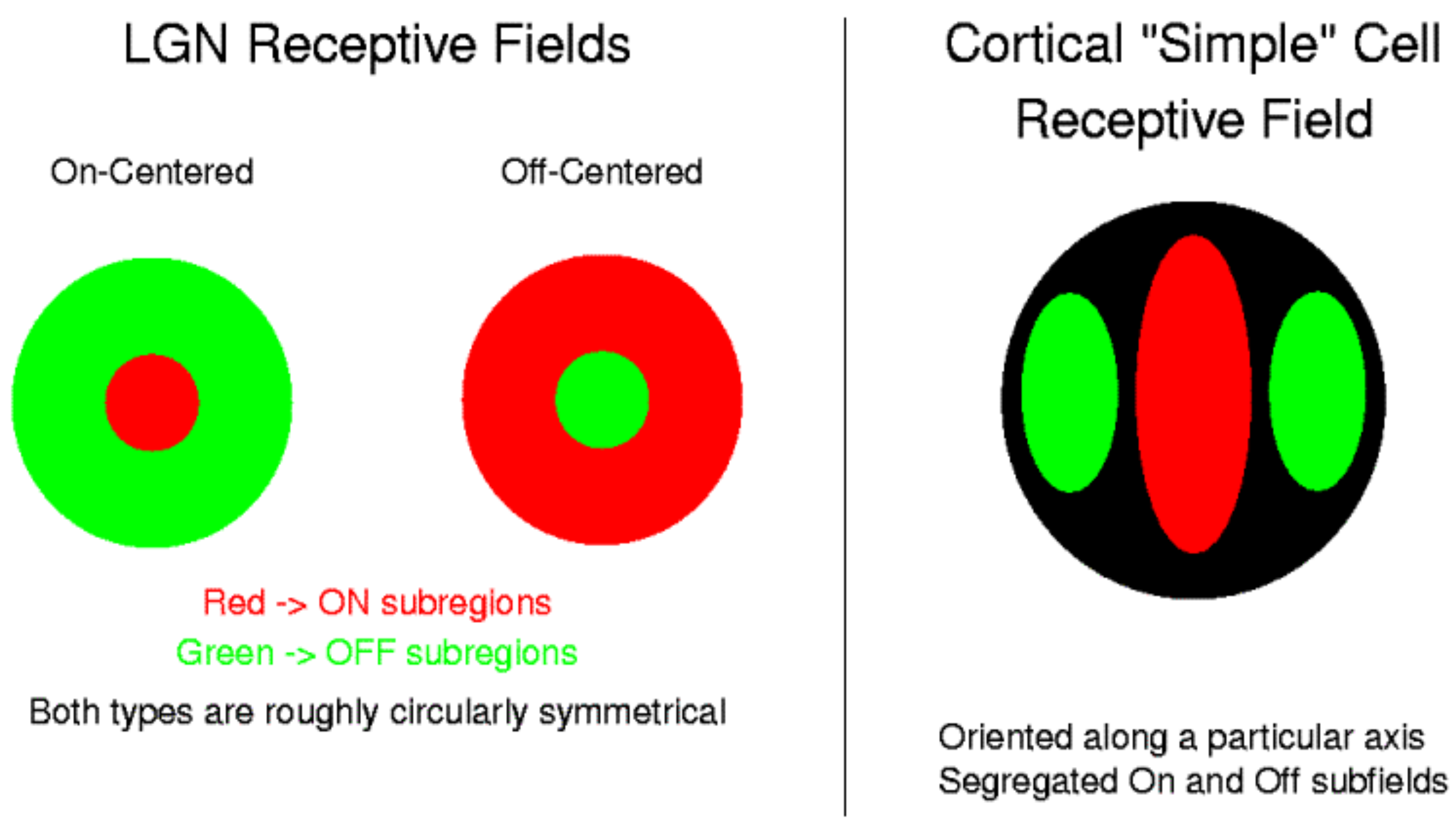
# Population recordings:



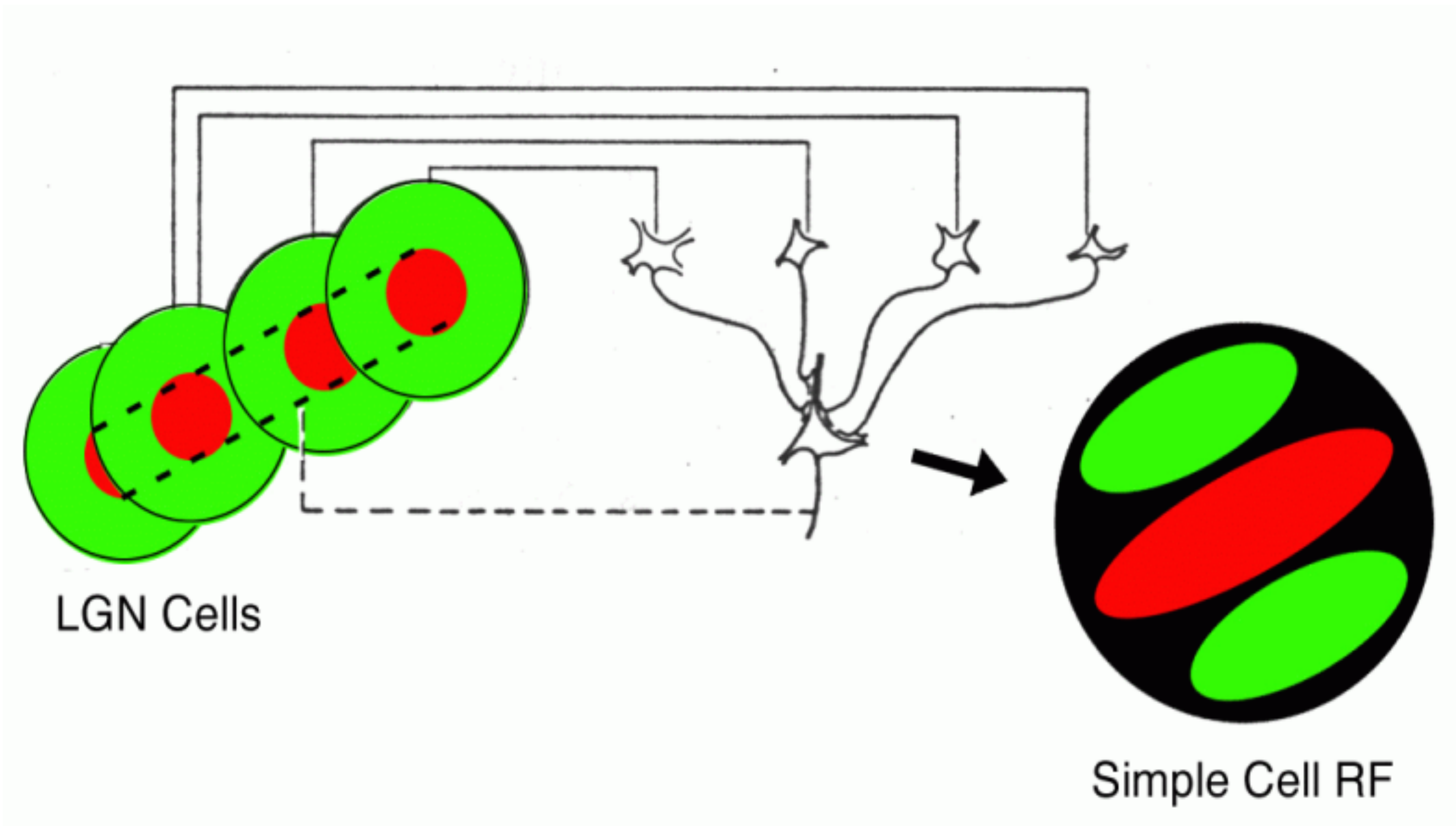



# The basic learning problem: what about the stimulus do neurons respond to?

# What about the stimulus do neurons respond to?



# What about the stimulus do neurons respond to?



Tools we use:

- statistics
- learning / inference
- optimization
- dynamical systems
- information theory

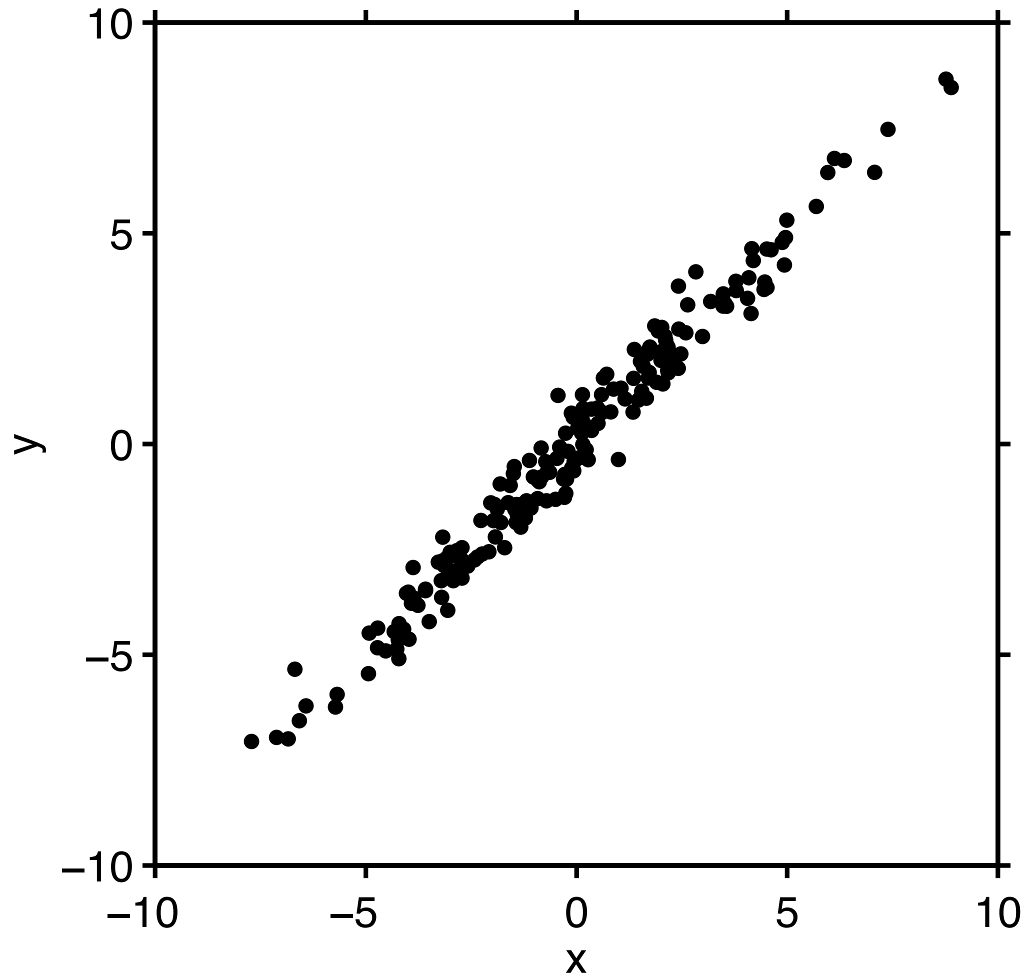
# A brief introduction to information theory:

## When is information theory useful?  When you...

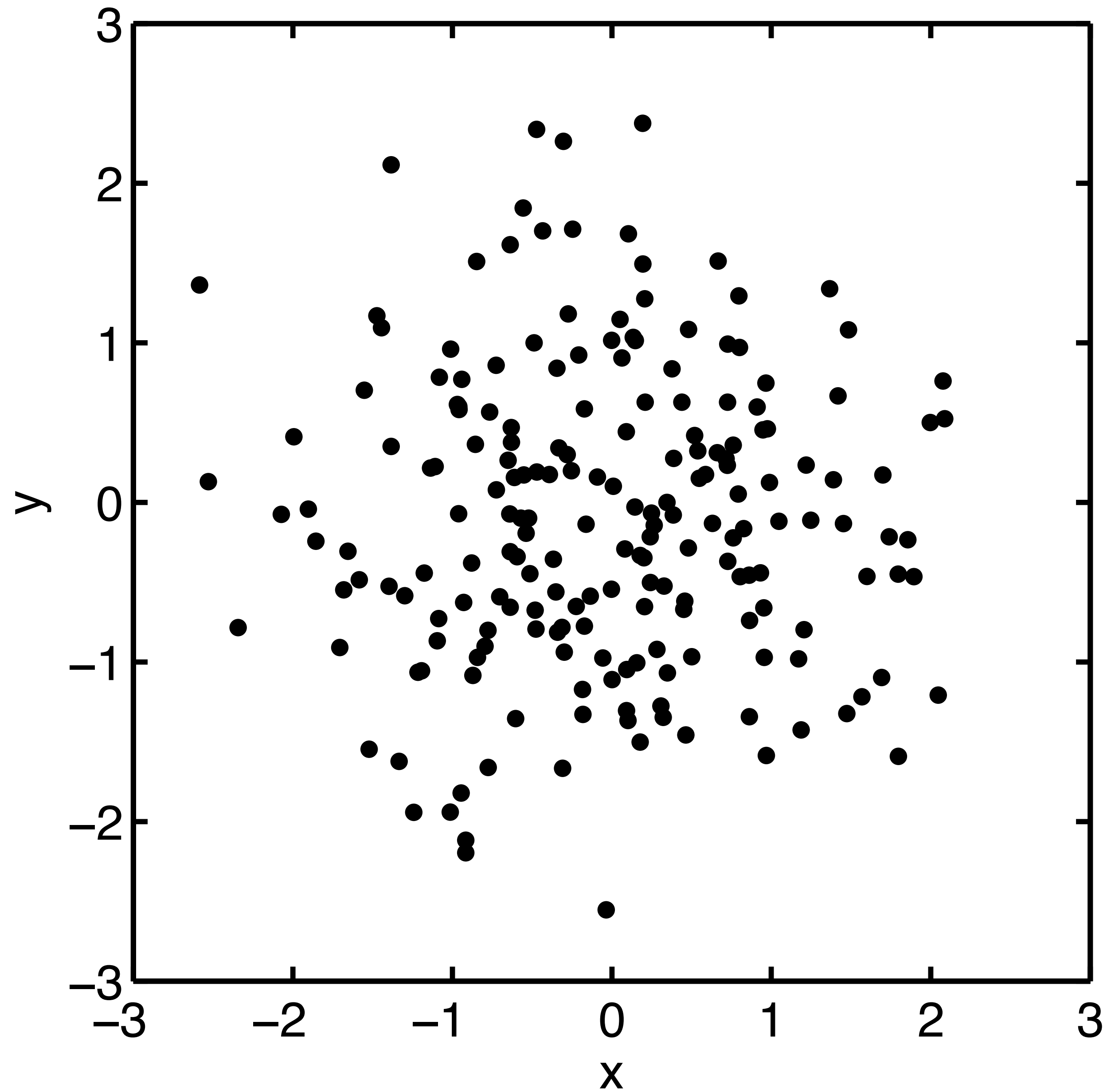
• want to go beyond linear correlation

• have enough data to sample P(x,y)

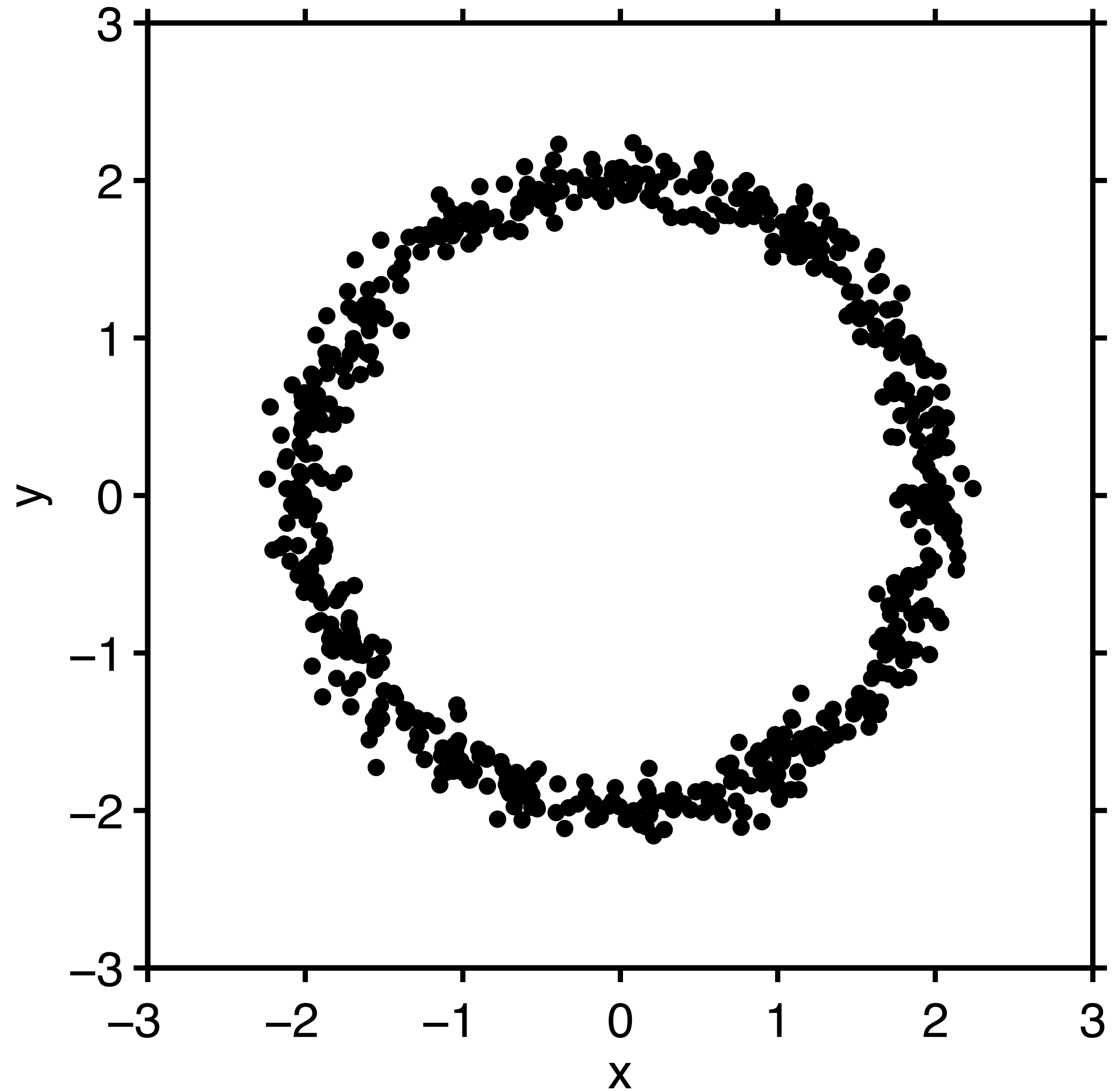
• are not sure what your 'code' is



# Zero correlation, no information:



# Zero correlation, but obvious information:



# Entropy as a measure of uncertainty:

$$\begin{aligned}\text{uncertainty} &= \log(n) \\ &= \log(1/p) \\ &= -\log(p)\end{aligned}$$

$$u_i = -\log(p_i)$$

$$\langle u_i \rangle = -\sum_i p_i \log(p_i)$$

$$S(X) = -\sum_x p(x)\log_2(p(x))$$

# Information == reduction in uncertainty

## Mutual information:

$$\begin{aligned} I(A;B) &= S(A) - S(A|B) \\ &= S(B) - S(B|A) \\ &= \sum_{a,b} P(a,b) \log_2 \left( \frac{P(a,b)}{P(a)P(b)} \right) \\ &= \sum_{a,b} P(a)P(b|a) \log_2 \left( \frac{P(b|a)}{P(b)} \right) \end{aligned}$$

# useful formulae

Product rule:

$$P(a,b) = P(a|b)P(b)$$

Sum rule:

$$P(a) = \sum_b P(a,b)$$
$$= \sum_b P(a|b)P(b)$$

Bayes' rule:

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$
$$= \frac{P(b|a)P(a)}{\sum_{a'} P(b|a')P(a')}$$

# more useful formulae

## Additivity:

$$S(A,B) = S(A) + S(B) \iff P(a,b) = P(a)P(b)$$

## Chain rule:

$$S(A,B) = S(A) + S(B|A) = S(B) + S(A|B)$$

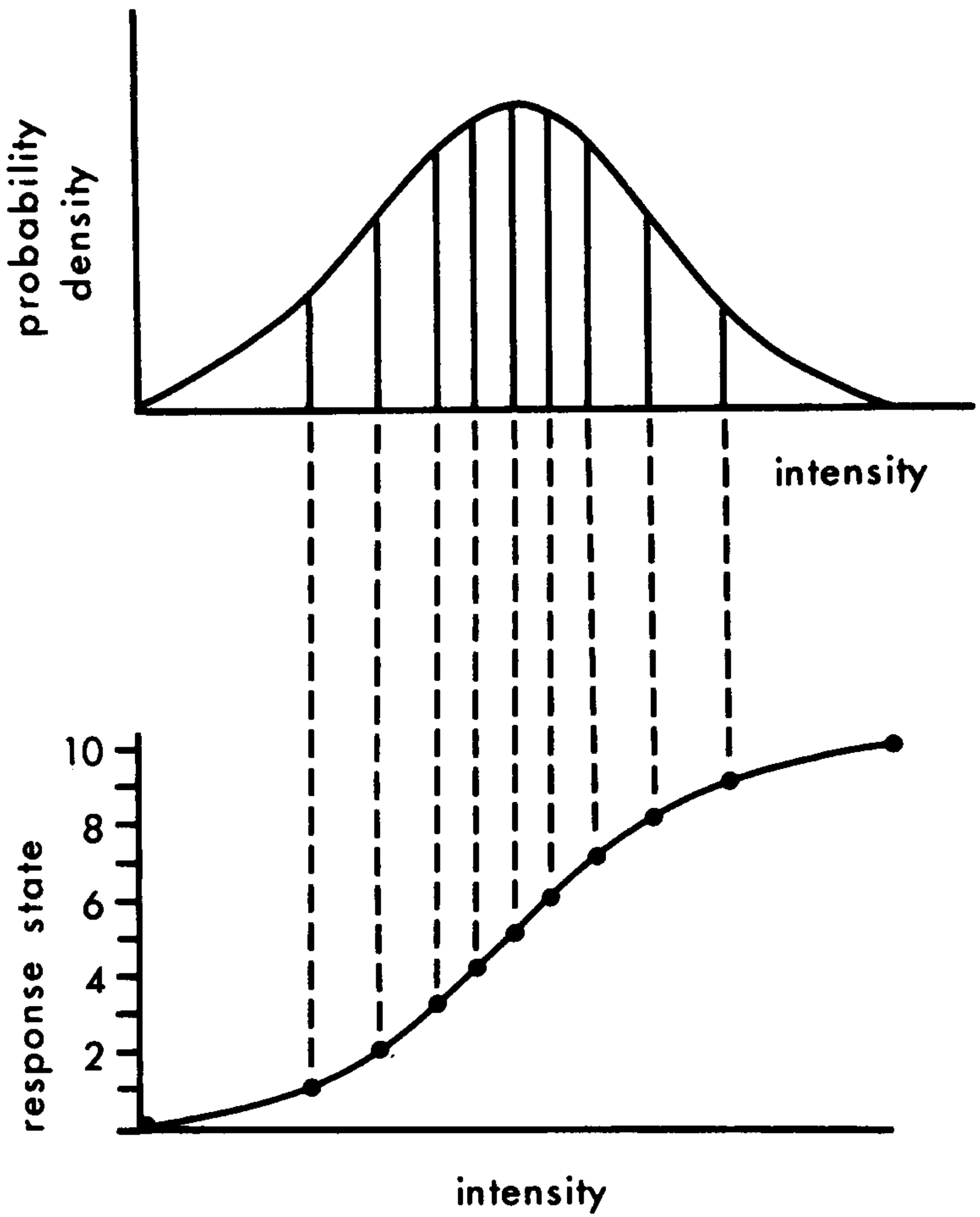
## Kullback-Liebler divergence ($D_{KL}$):

$$D_{\mathrm{KL}}(P,Q) = \sum_a P(a) \log_2 \frac{P(a)}{Q(a)}$$

## Mutual information ≥ 0:

$$I(A;B) = S(A) - S(A|B)$$

# Efficient coding in single neurons



# The efficient coding hypothesis, brief history:

- Claude Shannon (1948) *A Mathematical Theory of Communication*
- Fred Attneave (1954) *Some informational aspects of visual perception*
- Horace Barlow (1961) *Possible principles underlying the transformation of sensory messages*

**Are sensory systems optimized for information transmission?**

# Information theory example, the weighing problem:

chalkboard interlude