MIT Department of Brain and Cognitive Sciences
9.641J, Spring 2005 - Introduction to Neural Networks
Instructor: Professor Sebastian Seung

Excitatory-inhibitory networks

Sebastian Seung

Two neural populations

- “excitatory” and “inhibitory”
- interactions
  - within populations: symmetric
  - between populations: antisymmetric

The two populations of an excitatory-inhibitory network behave as if they have opposing goals.

Minimax

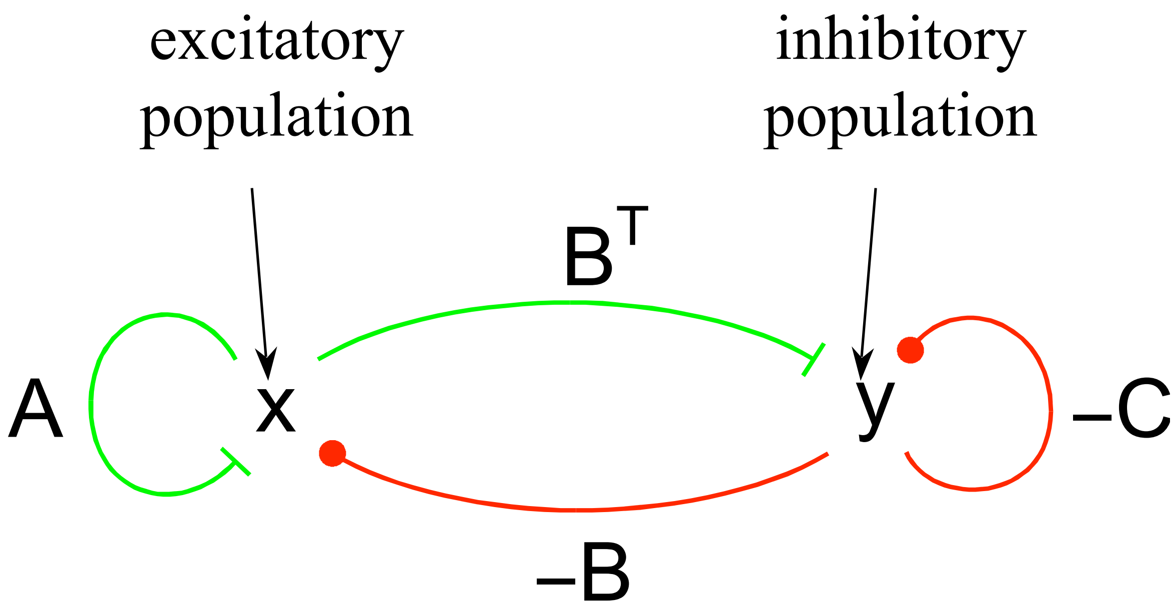
- An excitatory-inhibitory network is a method of solving a minimax problem.

$$\min_x \max_y S(x, y)$$

Multiple goals

- Analogy to game theory
  - zero-sum game
- Equilibrium
- Oscillations
- Complex non-periodic behavior

Synaptic interactions



- A and C symmetric
- excitatory-inhibitory interpretation
  - A, B, C nonnegative matrices

Matrix-vector notation

$$\tau_x \dot{x} + x = f\left(u + Ax - By\right)$$

$$\tau_y \dot{y} + y = g\left(v + B^T x - Cy\right)$$

Saddle function

- Excitatory neurons try to minimize
- Inhibitory neurons try to maximize

$$S = -u^T x - \tfrac{1}{2} x^T A x + v^T y - \tfrac{1}{2} y^T C y$$
$$+\mathbf{1}^T \overline{F}(x) + y^T B^T x - 1^T \overline{G}(y)$$

- Platt & Barr (1987)
- Mjolness & Garrett (1990)

Saddle function gradients

$$-\frac{\partial S}{\partial x} = u + Ax - By - f^{-1}(x)$$

$$= f^{-1}(\tau_x \dot{x} + x) - f^{-1}(x)$$

$$\frac{\partial S}{\partial y} = v + B^T x - Cy - g^{-1}(y)$$

$$= g^{-1}(\tau_y \dot{y} + y) - g^{-1}(y)$$

Pseudo gradient ascent-descent

$$\tau_x \dot{x} \cong -\frac{\partial S}{\partial x}$$ descent

$$\tau_y \dot{y} \cong \frac{\partial S}{\partial y}$$ ascent

- The components of these vectors have the same sign.

True gradient ascent-descent

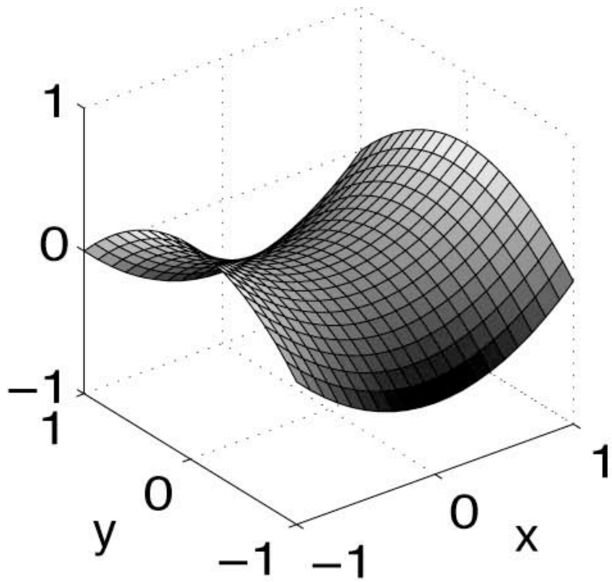
$$\dot{x} = -\frac{\partial S}{\partial x} \qquad \text{descent}$$

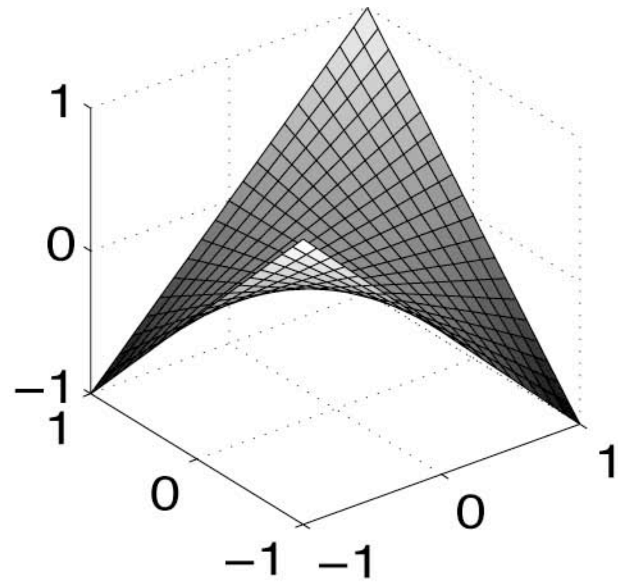
$$\dot{y} = \frac{\partial S}{\partial y} \qquad \text{ascent}$$

- When does this dynamics converge to the solution of the minimax problem?

$$\min_x \max_y S(x, y)$$

It depends





$$S = \frac{x^2}{2} - \frac{y^2}{2}$$

steady state

$$S = xy$$

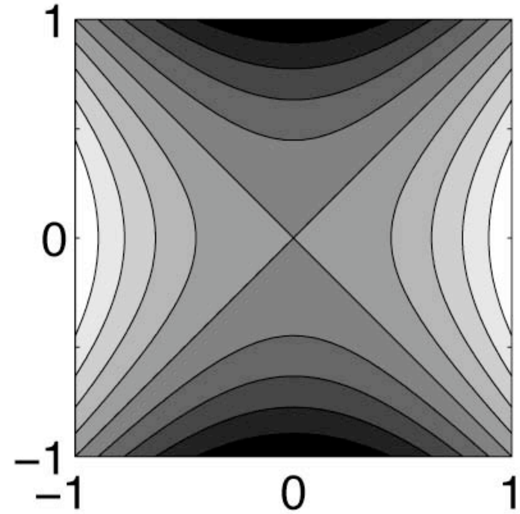
oscillations

Steady state

$$S = \frac{x^2}{2} - \frac{y^2}{2}$$

$$\dot{x} = -\frac{\partial S}{\partial x} = -x$$

$$\dot{y} = \frac{\partial S}{\partial y} = -y$$

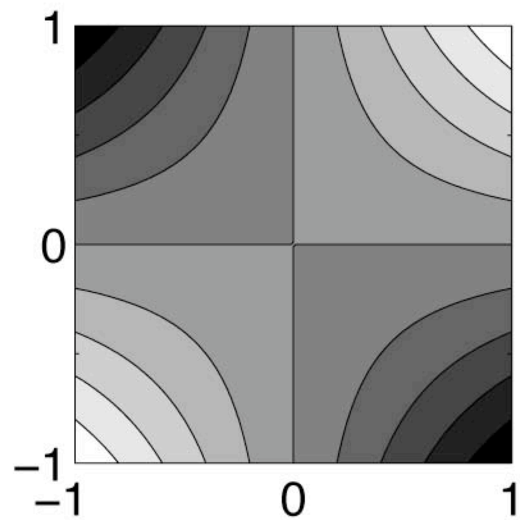


Periodic behavior

$$S = xy$$

$$\dot{x} = -\frac{\partial S}{\partial x} = -y$$

$$\dot{y} = \frac{\partial S}{\partial y} = x$$



Kinetic energy

$$T = \frac{\dot{x}^2}{2} + \frac{\dot{y}^2}{2} \qquad \dot{T} = -\dot{x}^T \frac{\partial^2 S}{\partial x^2} \dot{x} + \dot{y}^T \frac{\partial^2 S}{\partial y^2} \dot{y}$$

- lower bounded
- nonincreasing if

$$\frac{\partial^2 S}{\partial x^2} \text{ positive definite}$$

$$\frac{\partial^2 S}{\partial y^2} \text{ negative definite}$$

Proof

$$\dot{x} = -\frac{\partial S}{\partial x} \longrightarrow \ddot{x} = -\frac{\partial^2 S}{\partial x^2}\dot{x} - \frac{\partial^2 S}{\partial x \partial y}\dot{y}$$

$$\dot{y} = \frac{\partial S}{\partial y} \longrightarrow \ddot{y} = \frac{\partial^2 S}{\partial x \partial y}\dot{x} + \frac{\partial^2 S}{\partial y^2}\dot{y}$$

$$\dot{T} = \dot{x}\ddot{x} + \dot{y}\ddot{y}$$

$$= -\frac{\partial^2 S}{\partial x^2}\dot{x}^2 + \frac{\partial^2 S}{\partial y^2}\dot{y}^2$$

The saddle function could either increase or decrease

$$\frac{dS}{dt} = \dot{x}^T \frac{\partial S}{\partial x} + \dot{y}^T \frac{\partial S}{\partial y} = -\dot{x}^T \dot{x} + \dot{y}^T \dot{y}$$

Lyapunov function

$$L = T + rS$$

$$\dot{L} = -\dot{x}^T \left(\frac{\partial^2 S}{\partial x^2} + rI \right) \dot{x} + \dot{y}^T \left(\frac{\partial^2 S}{\partial y^2} + rI \right) \dot{y}$$

$$\frac{\partial^2 S}{\partial x^2} + rI \ \text{ positive definite}$$

$$\frac{\partial^2 S}{\partial y^2} + rI \ \text{ negative definite}$$

choose r to satisfy these conditions
and keep L lower bounded

Legendre transform pairs

$$F \xleftrightarrow{\text{Legendre transformation}} \overline{F}$$

$$F'(x) = f(x) \quad \overline{F}'(x) = f^{-1}(x)$$

$$\overline{F}(x) = \max_{p}\left\{px - F(p)\right\}$$

$$\Phi(p,x) = \mathbf{1}^T F(p) - p^T x + \mathbf{1}^T \overline{F}(x)$$

Generalized kinetic energy

$$\boldsymbol{\tau}_x \dot{x} + x = f\left(u + Ax - By\right)$$

$$\tfrac{1}{2}\boldsymbol{\tau}_x \dot{x}^2 \longrightarrow \boldsymbol{\tau}_x^{-1}\Phi\left(u + Ax - By, x\right)$$

$$\Phi\left(p, x\right) = \mathbf{1}^T F\left(p\right) - p^T x + \mathbf{1}^T \overline{F}(x)$$

$$\Phi\left(p, x\right) \geq 0$$

$$\Phi\left(p, x\right) = 0 \text{ for } f\left(p\right) = x$$

likewise, $\Gamma\left(q, x\right) = \mathbf{1}^T G\left(q\right) - q^T x + \mathbf{1}^T \overline{G}(x)$

Lyapunov function

$$F'(x) = f(x) \quad \overline{F}'(x) = f^{-1}(x) \quad G'(x) = g(x) \quad \overline{G}'(x) = g^{-1}(x)$$

kinetic energy

$$\Phi(p,x) = \mathbf{1}^T F(p) - p^T x + \mathbf{1}^T \overline{F}(x)$$

$$\Gamma(q,x) = \mathbf{1}^T G(q) - q^T x + \mathbf{1}^T \overline{G}(x)$$

saddle function

$$S = -u^T x - \tfrac{1}{2} x^T A x + v^T y - \tfrac{1}{2} y^T C y + \mathbf{1}^T \overline{F}(x) + y^T B^T x - 1^T \overline{G}(y)$$

Lyapunov function

$$L = \frac{1}{\boldsymbol{\tau}_x} \Phi(u + Ax - By, x) + \frac{1}{\boldsymbol{\tau}_y} \Gamma\left(v + B^T x - Cy, y\right) + rS$$

Need to verify that L is lower bounded

Sufficient conditions for stability

$$\dot{L} = \dot{x}^T A\dot{x} - \dot{y}^T C\dot{y} - \left(\tau_x^{-1} + r\right)\dot{x}^T\left[f^{-1}\left(\tau_x\dot{x} + x\right) - f^{-1}(x)\right]$$
$$+\left(r - \tau_y^{-1}\right)\dot{y}^T\left[g^{-1}\left(\tau_y\dot{y} + y\right) - g^{-1}(y)\right]$$

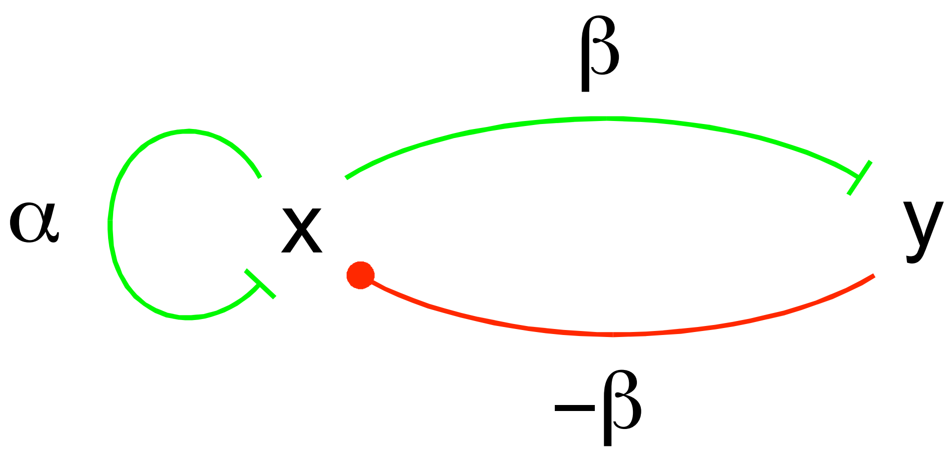
sufficient condition for $\dot{L} \le 0$

$$\max_{a,b}\frac{(a-b)^T A(a-b)}{(a-b)^T\left(f^{-1}(a) - f^{-1}(b)\right)} \le 1 + r\tau_x$$

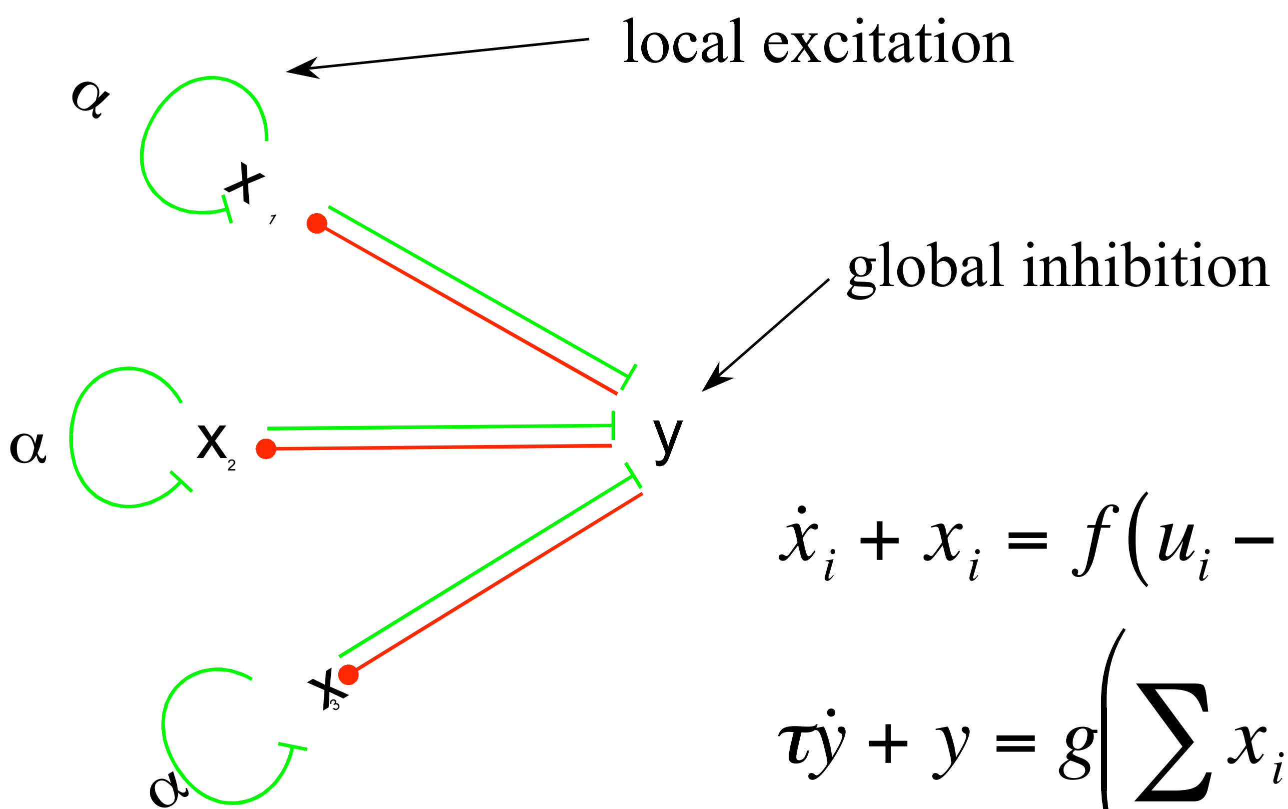
$$\min_{a,b}\frac{(a-b)^T C(a-b)}{(a-b)^T\left(g^{-1}(a) - g^{-1}(b)\right)} \ge r\tau_y - 1$$

Excitatory-inhibitory pair

- inhibitory feedback causes oscillations
- self-excitation required to sustain them



Competitive network



$$\dot{x}_i + x_i = f\left(u_i - y + \alpha x_i\right)$$

$$\tau\dot{y} + y = g\left(\sum_i x_i\right)$$

Sufficient conditions

$$T = \sum_i \left[F(u_i + \alpha x_i - y) - (u_i + \alpha x_i - y)x_i + \overline{F}(x_i) \right]$$

$$V = \sum_i \left[-u_i x_i - \frac{1}{2}\alpha x_i^2 + \overline{F}(x_i) + G\left(\sum_i x_i\right) \right]$$

$$L = T + V/\tau$$

$$\dot{L} = \sum_i \left\{ \alpha \dot{x}_i^2 - \left(\tau^{-1} + 1\right)\dot{x}_i \left[f^{-1}(\dot{x}_i + x_i) - f^{-1}(x_i) \right] \right\}$$

Conclusion

- excitatory-inhibitory network
- dynamics on a saddle
  - gradient ascent/descent
  - shape of saddle determines behavior