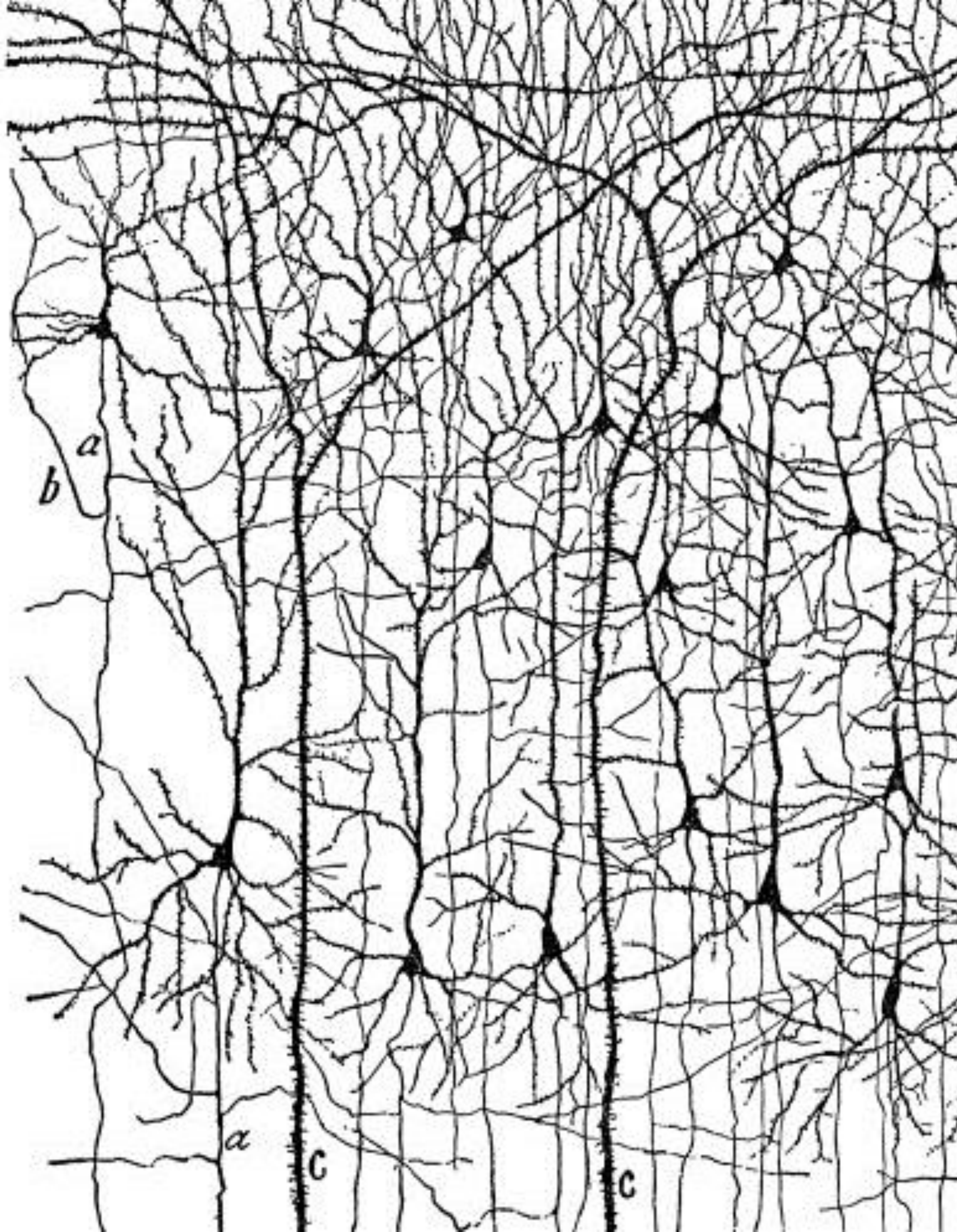
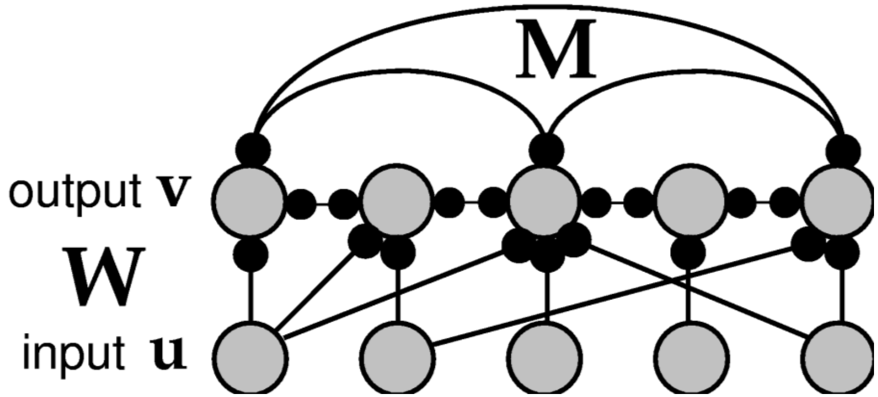
Supplementary Material for Recurrent Networks

Compiled by:
Rajesh P. N. Rao
Computational Neuroscience



Drawing by Ramón y Cajal



Solving the Differential Equation for a Linear Recurrent Network

$$\tau \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \underbrace{\mathrm{W}\mathbf{u}}_{\mathbf{h}} + \mathrm{M}\mathbf{v}$$

Image Source: Dayan & Abbott textbook

Using Eigenvectors to Solve for $\mathbf{v}(t)$

$$\tau \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \mathbf{h} + M\mathbf{v}$$

- Suppose $N \times N$ matrix M is symmetric
- M has N *orthogonal* eigenvectors $\mathbf{e}_i$ and N eigenvalues λ_i which satisfy:

$$\mathbf{M}\mathbf{e}_i = \lambda_i \mathbf{e}_i$$

- Normalize eigenvectors to have length 1
  - Divide each by its length
- Then we have a set of *orthonormal* vectors (a new "basis" or coordinate system) such that:

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \text{ for } i \neq j \text{ and } 1 \text{ otherwise}$$

Using Eigenvectors to Solve for $\mathbf{v}(t)$

- We can represent any N-dimensional vector, including our output vector $\mathbf{v}(t)$, using the orthonormal eigenvectors of M:

$$\mathbf{v}(t) = \sum_{j=1}^{N} c_j(t)\mathbf{e}_j$$

- Substituting above in the diff. equation for $\mathbf{v}$: $\tau \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \mathbf{h} + \mathrm{M}\mathbf{v}$

$$\tau \frac{d\sum_{j=1}^{N} c_j \mathbf{e}_j}{dt} = -\sum_{j=1}^{N} c_j \mathbf{e}_j + \mathbf{h} + \mathrm{M}\sum_{j=1}^{N} c_j \mathbf{e}_j$$

$$\tau \sum_{j=1}^{N} \frac{dc_j}{dt}\mathbf{e}_j = -\sum_{j=1}^{N} c_j\left(\mathbf{e}_j - \mathrm{M}\mathbf{e}_j\right) + \mathbf{h}$$

Using eigenvector equation

$$\mathrm{M}\mathbf{e}_j = \lambda_j \mathbf{e}_j$$

$$\tau \sum_{j=1}^{N} \frac{dc_j}{dt}\mathbf{e}_j = -\sum_{j=1}^{N} c_j\left(\mathbf{e}_j - \lambda_j \mathbf{e}_j\right) + \mathbf{h}$$

Using Eigenvectors to Solve for $\mathbf{v}(t)$

- Taking the dot product of both sides with any $\mathbf{e}_i$:

$$(\tau \sum_{j=1}^{N} \frac{dc_j}{dt} \mathbf{e}_j = -\sum_{j=1}^{N} c_j \left(\mathbf{e}_j - \lambda_j \mathbf{e}_j \right) + \mathbf{h}) \cdot \mathbf{e}_i$$

$$\tau \frac{dc_i}{dt} = -c_i \left(1 - \lambda_i\right) + \mathbf{h} \cdot \mathbf{e}_i$$

Using orthonormality of the $\mathbf{e}_i$

- Solve differential equation for c_i (and use it to get $\mathbf{v}(t)$!):

$$c_i(t) = \frac{\mathbf{h} \cdot \mathbf{e}_i}{1 - \lambda_i} (1 - \exp(\frac{-t(1-\lambda_i)}{\tau}) + c_i(0) \exp(\frac{-t(1-\lambda_i)}{\tau}))$$

$$\mathbf{v}(t) = \sum_{i=1}^{N} c_i(t) \mathbf{e}_i$$

Eigenvalues determine Network Stability!

$$\mathbf{v}(t) = \sum_{i=1}^{N} c_i(t)\mathbf{e}_i \qquad c_i(t) = \frac{\mathbf{h}\cdot\mathbf{e}_i}{1-\lambda_i}\left(1-\exp\left(\frac{-t(1-\lambda_i)}{\tau}\right)\right) + c_i(0)\exp\left(\frac{-t(1-\lambda_i)}{\tau}\right))$$

If any $\lambda_i > 1$, $\mathbf{v}(t)$ explodes $\Rightarrow$ network is unstable!

If all $\lambda_i < 1$, network is stable and $\mathbf{v}(t)$ converges to steady state value:

$$\mathbf{v}_{ss} = \sum_i \frac{\mathbf{h}\cdot\mathbf{e}_i}{1-\lambda_i}\mathbf{e}_i$$

Amplification of Inputs in a Recurrent Network

$$\mathbf{v}_{ss} = \sum_i \frac{\mathbf{h}\cdot\mathbf{e}_i}{1-\lambda_i}\mathbf{e}_i$$

If all $\lambda_i < 1$ and one λ_i (say λ_1) is close to 1 with others much smaller:

$$\mathbf{v}_{ss} \approx \frac{\mathbf{h}\cdot\mathbf{e}_1}{1-\lambda_1}\mathbf{e}_1$$

Amplification of input projection by a factor of $\frac{1}{1-\lambda_i}$

Memory in Linear Recurrent Networks

$$\tau \frac{d\mathbf{v}}{dt} = -\mathbf{v} + \mathbf{h} + \mathrm{M}\mathbf{v} \qquad \mathbf{v}(t) = \sum_{i=1}^{N} c_i(t)\mathbf{e}_i$$

Suppose $\lambda_1 = 1$ and all other $\lambda_i < 1$. Then, $\tau \frac{dc_1}{dt} = \mathbf{h} \cdot \mathbf{e}_1$

Solving for c_1, we get $c_1(t) = c_1(0) + \frac{1}{\tau}\int_0^t \mathbf{h}(t') \cdot \mathbf{e}_1 dt' \tau$

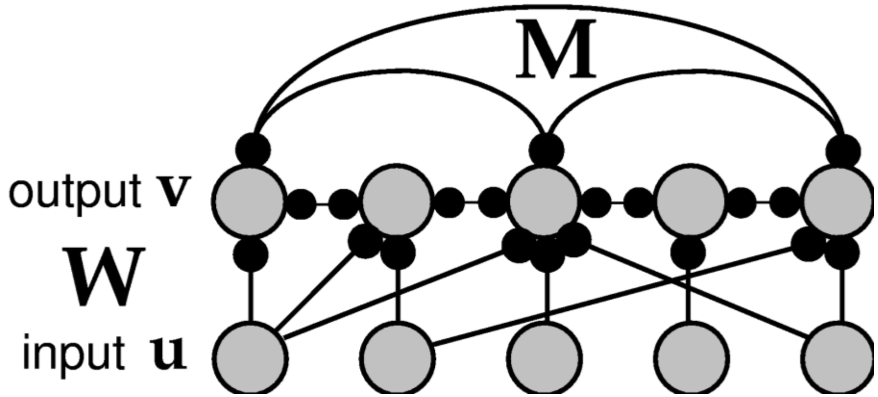
If input $\mathbf{h}$ is turned on and then off, can show that <u>even after $\mathbf{h} = 0$</u>:

$$\mathbf{v}(t) = \sum_i c_i(t)\mathbf{e}_i$$

$$\approx c_1\mathbf{e}_1 = \frac{\mathbf{e}_1}{\tau}\int_0^t \mathbf{h}(t') \cdot \mathbf{e}_1 dt' \qquad (\text{assuming } c_1(0) = 0)$$

Sustained activity without any input!

Networks keeps a memory of integral of past input



Stability of Nonlinear Recurrent Networks

$$\tau \frac{d\mathbf{v}}{dt} = -\mathbf{v} + F(\mathbf{h} + \mathbf{M}\mathbf{v})$$

Image Source: Dayan & Abbott textbook

Stability Analysis of Nonlinear Recurrent Networks

General case : $\dfrac{d\mathbf{v}}{dt} = \mathbf{f}(\mathbf{v})$

Suppose $\mathbf{v}_\infty$ is a fixed point (i.e., $\mathbf{f}(\mathbf{v}_\infty) = 0$)

We want to know: How stable is the network near $\mathbf{v}_\infty$ *?*

Near $\mathbf{v}_\infty$, $\mathbf{v}(t) = \mathbf{v}_\infty + \boldsymbol{\varepsilon}(t)$. Differentiating both sides wrt t : $\dfrac{d\mathbf{v}}{dt} = \dfrac{d\boldsymbol{\varepsilon}}{dt}$.

Taylor expansion near $\mathbf{v}_\infty$: $\mathbf{f}(\mathbf{v}(t)) \approx \mathbf{f}(\mathbf{v}_\infty) + \left.\dfrac{\partial \mathbf{f}}{\partial \mathbf{v}}\right|_{\mathbf{v}_\infty} \boldsymbol{\varepsilon}(t) = \left.\dfrac{\partial \mathbf{f}}{\partial \mathbf{v}}\right|_{\mathbf{v}_\infty} \boldsymbol{\varepsilon}(t)$

i.e. $\dfrac{d\mathbf{v}}{dt} = \mathbf{f}(\mathbf{v}(t)) = \left.\dfrac{\partial \mathbf{f}}{\partial \mathbf{v}}\right|_{\mathbf{v}_\infty} \boldsymbol{\varepsilon}(t) = J \cdot \boldsymbol{\varepsilon}(t)$

J is the stability or “Jacobian” matrix

Also, since $\dfrac{d\mathbf{v}}{dt} = \dfrac{d\boldsymbol{\varepsilon}}{dt}$, we have $\dfrac{d\boldsymbol{\varepsilon}}{dt} = J \cdot \boldsymbol{\varepsilon}(t)$

Assuming J is real with N linearly independent eigenvectors $\mathbf{e}_i$, we can write :

$$\boldsymbol{\varepsilon}(t) = \sum_i c_i(t)\mathbf{e}_i$$

(based on Mathematical Appendix A.3 in Dayan & Abbott textbook)

Stability Analysis of Nonlinear Recurrent Networks

Continued from previous page :

Substituting $\boldsymbol{\varepsilon}(t) = \sum_i c_i(t)\mathbf{e}_i$ into $\frac{d\boldsymbol{\varepsilon}}{dt} = J \cdot \boldsymbol{\varepsilon}(t)$, we find that the coefficients must satisfy :

$\frac{dc_i}{dt} = \lambda_i c_i$. The solution is : $c_i(t) = c_i(0)\exp(\lambda_i t)$

Therefore, $\boldsymbol{\varepsilon}(t) = \sum_i c_i(0)\exp(\lambda_i t)\mathbf{e}_i$ i.e., evolution of $\mathbf{v}(t)$ near $\mathbf{v}_\infty$ depends on eigenvalues of J.

Each individual term in the above sum is called a mode.

Eigenvalues of J can be complex, e.g., $\lambda_k = a_k + ib_k$, which means that :

$\exp(\lambda_k t) = \exp(a_k t)\exp(ib_k t) = \exp(a_k t)(\cos(b_k t) + i\sin(b_k t))$. This implies :

For complex λ_k ($b_k \neq 0$), the mode will oscillate with frequency b_k.

If $a_k < 0$ for all k, the oscillations will be damped exponentially to 0 and the network will be stable near the fixed point $\mathbf{v}_\infty$.

If $a_k > 0$ for any k, the oscillations for that mode will grow exponentially and the network may not be stable (unless the network's nonlinearity curbs the growth as in the following example).

Example: Non-Symmetric Nonlinear Recurrent Network

- Example: Network of Excitatory (E) and Inhibitory (I) Neurons
  - Connections can't be symmetric: Why?

10 ms → τ_E

$$\tau_E \frac{dv_E}{dt} = -v_E + [M_{EE} v_E + M_{EI} v_I - \gamma_E]^+$$

(M_{EE} = 1.25, M_{EI} = -1, γ_E = -10)

$$\tau_I \frac{dv_I}{dt} = -v_I + [M_{II} v_I + M_{IE} v_E - \gamma_I]^+$$

(M_{II} = 0, M_{IE} = 1, γ_I = 10)

Parameter we will vary to study the network → τ_I

We want to analyze stability of this network near fixed point $\mathbf{v}_\infty$ i.e., near the (v_E, v_I) which results in $\frac{dv_E}{dt} = \frac{dv_I}{dt} = 0.$

Linear Stability Analysis near Fixed Point

$$\frac{dv_E}{dt} = \frac{-v_E + [M_{EE}v_E + M_{EI}v_I - \gamma_E]^+}{\tau_E}$$

$$\frac{dv_I}{dt} = \frac{-v_I + [M_{II}v_I + M_{IE}v_E - \gamma_I]^+}{\tau_I}$$

Take derivatives of right hand side with respect to both v_E and v_I

Stability Matrix (aka the "Jacobian" Matrix):

$$J = \begin{bmatrix} \frac{(M_{EE}-1)}{\tau_E} & \frac{M_{EI}}{\tau_E} \\ \frac{M_{IE}}{\tau_I} & \frac{(M_{II}-1)}{\tau_I} \end{bmatrix}$$

Compute the Eigenvalues

- Jacobian Matrix:

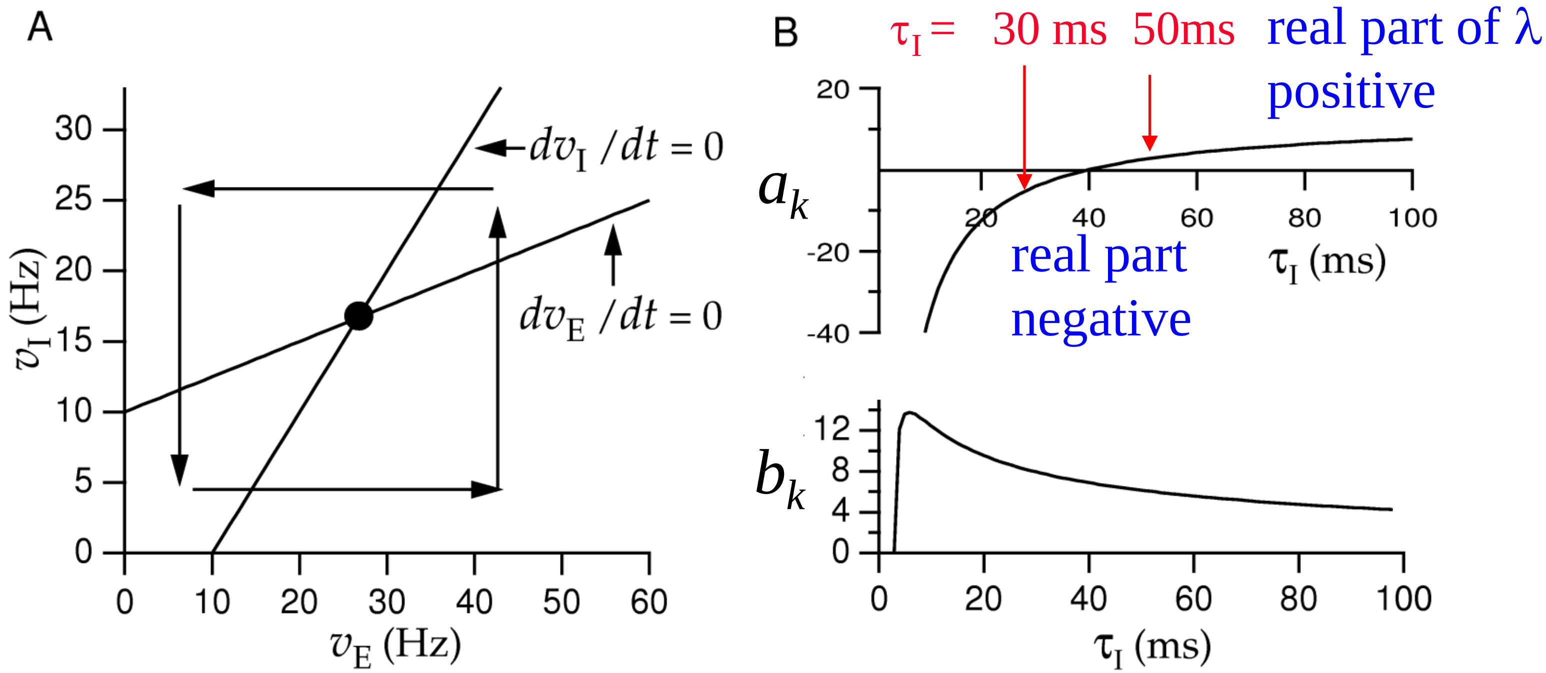
$$J = \begin{bmatrix} \dfrac{(M_{EE} - 1)}{\tau_E} & \dfrac{M_{EI}}{\tau_E} \\ \dfrac{M_{IE}}{\tau_I} & \dfrac{(M_{II} - 1)}{\tau_I} \end{bmatrix}$$

- Its two eigenvalues (obtained by solving *det*(J – λI) = 0):

$$\lambda = \frac{1}{2}\left(\frac{(\overset{1.25}{M_{EE}} - 1)}{\underset{10\text{ ms}}{\tau_E}} + \frac{(\overset{0}{M_{II}} - 1)}{\tau_I} \pm \sqrt{\left(\frac{M_{EE} - 1}{\tau_E} - \frac{M_{II} - 1}{\tau_I}\right)^2 + 4\frac{\overset{-1}{M_{EI}}\overset{1}{M_{IE}}}{\tau_E \tau_I}} \right)$$

Next page plots real and imaginary parts of λ as a function of τ_I

Phase Plane and Eigenvalue Analysis



$$10\frac{dv_E}{dt} = -v_E + [1.25v_E - v_I + 10]^+$$

$$\tau_I\frac{dv_I}{dt} = -v_I + [0 \cdot v_I + v_E - 10]^+$$

Real and imaginary parts (a_k and b_k) of λ ($= a_k + ib_k$) as a function of τ_I

Damped Oscillations in the Network

Choose $\tau_I = 30$ ms (makes real part of eigenvalues negative)

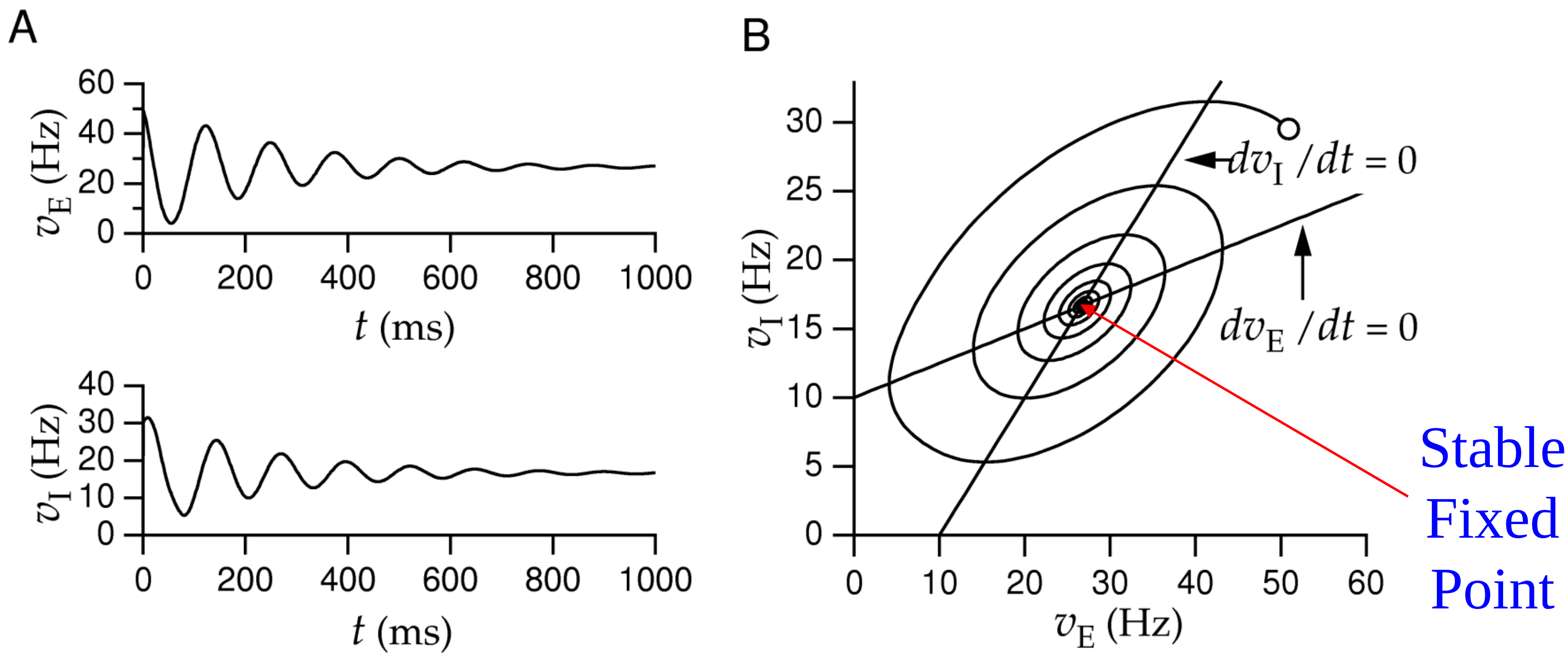
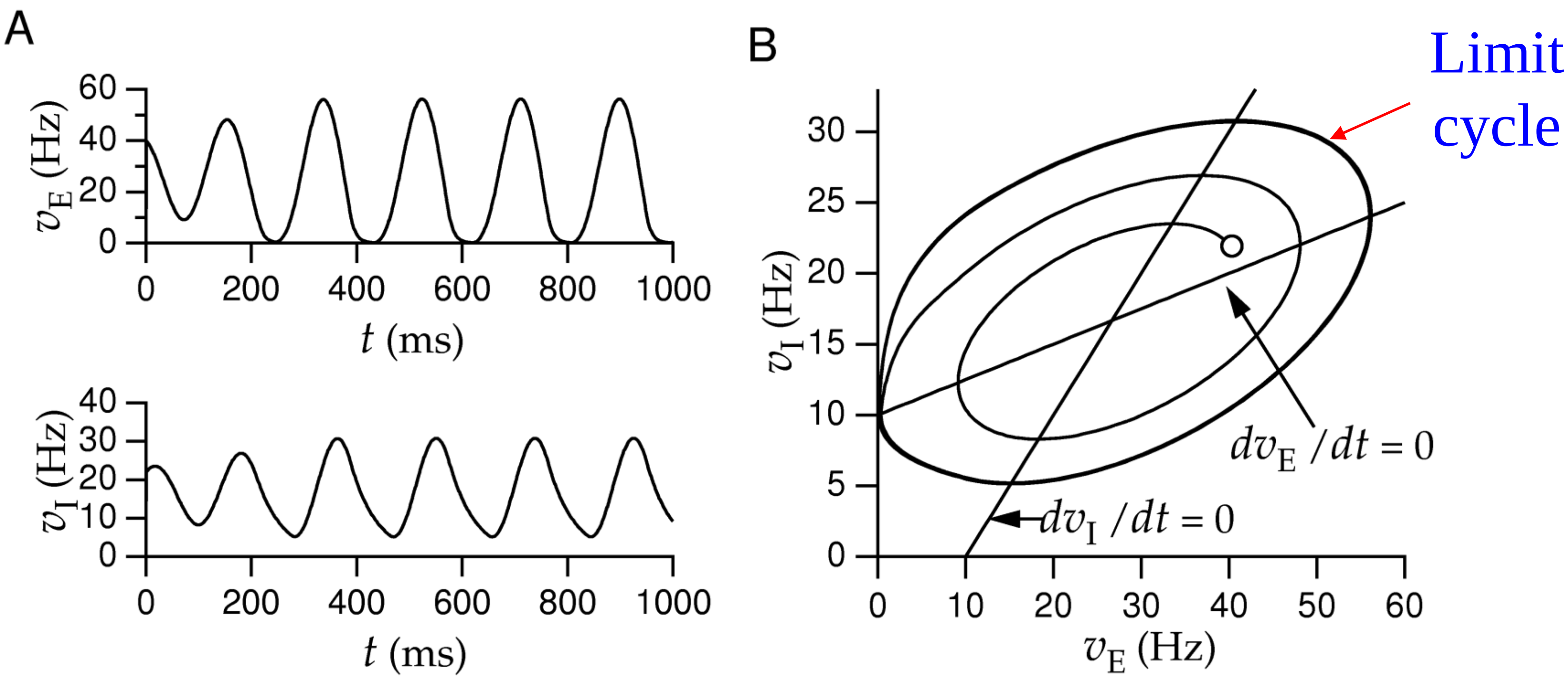


Image Source: Dayan & Abbott textbook

Unstable Behavior and Limit Cycle

Choose $\tau_I = 50$ ms (makes real part of eigenvalues positive)



Oscillations grow initially but curbed by rectification nonlinearity

Image Source: Dayan & Abbott textbook