

Mathematical Conditions for Brain Stability

John W. Ryon

Computer and Information Science Department
New Jersey Institute of Technology
Newark, NJ 07102
ryon@cis.njit.edu

We investigate the stability of a dynamic brain model using conventional techniques of linearization of the state space equations of motion for two nearby solutions followed by solution via expansion in the eigenvectors of the Jacobian matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. The fundamental stability condition for brains is then $|\lambda_k| \equiv 1$. By equating terms in two versions of the characteristic polynomial of the Jacobian we obtain a set of equations relating principle minors of the Jacobian to the fundamental symmetric polynomials formed from the eigenvalues. These equations may be considered as constraints on the Jacobian since the eigenvalues are constrained by the stability conditions. In turn, these equations may be regarded as constraints on brain structure since the Jacobian embodies the dynamical structure of the brain model.

Introduction

In this paper we investigate the conditions a brain must satisfy to be stable when considered as a dynamic system. We know that normal brains are stable in the sense that they do not spontaneously evolve toward coma or brain death on the one hand, or toward mania or seizure on the other. Those conditions are severely abnormal. The task is to translate this physiological statement into a form that applies to the dynamical model presented below.

To accomplish this we take coma and brain death to correspond to brain states in which the neuron firing rates are either zero, constant, very weak, or repetitive. In dynamical systems such states are called fixed points, steady states, and limit cycles. We also take mania and seizure to correspond to brain states in which the firing rates approach their maximum possible values. In dynamical systems such trajectories are called unstable.

In the theory of nonlinear dynamical systems, instability occurs when the dynamic quantities that define a system state evolve progressively toward infinite values, while fixed points, steady states, and limit cycles are regarded as attributes of stable systems. In this paper we regard steady states and cycles as *too stable*. Thus a *stable brain* is one that succeeds in continuously evolving in the intermediate region between states of little or no activity, and states of excessive activity. Stable brains are able to maintain indefinitely a level of activity that is neither too low nor too high, but instead is "just right."

To investigate this we follow the conventional approach in which we compare the evolution of two systems (brains) that initially are in nearly identical states. We do this by computing the difference between the two systems as time advances. If the equations force the two systems to converge to an identical state, then the systems are evolving toward a steady state or limit cycle. If the two system states are forced to diverge without limit, then the system is unstable. As we have noted, both situations are undesirable for brains. Lastly, if the equations constrain the two system states to remain close together, neither converging nor diverging, then the systems are dynamically stable. Such dynamically stable systems are consistent with normal brain functioning.

Brain Model

Here we give a *brief* presentation of the mathematical model used to represent brains. A more detailed discussion may be found elsewhere [Ryon, 1997]. See Kandel, *et al* for a thorough discussion of the neural science material.

We assume that the firing rate of neurons is the physical quantity of interest. We denote the firing rate of neuron i at time t by $r_i(t)$ and define it to be the average rate at which action potentials are initiated in the axon hillock of neuron i at time t . There is a time delay between the time an action potential fires in neuron i and the time it affects the firing of a successor neuron j . This time delay, denoted by t_{ij} , consists of the time required for an action potential to travel down the axon of neuron i , cross the synapse to neuron j , and diffuse along the dendrite tree and around the cell body to the hillock (trigger zone) of neuron j . Thus the firing rate of a neuron i is determined by the firing rates of the neurons that synapse upon it *at earlier times*.

We take the following *rate equations* as the model of neuron activity in brains:

$$r_i(t) = f_i(r_1(t-t_{1i}), r_2(t-t_{2i}), \dots, r_n(t-t_{ni}))$$

Here n is the total number of neurons in the brain, and f_i is the function that represents the transformation of the combined effects of the neurons making synapses on neuron i into the firing rate of neuron i .

If a neuron k makes no synapse upon neuron i then f_i does not depend upon r_k which may be expressed as $\partial f_i / \partial r_k = 0$.

A *solution* to these rate equations consists of functions $r_i(t)$ defined for all times in some non empty interval that satisfy the equations in the same time interval. We may also refer to solutions as *trajectories*.

Observe the generality of the these rate equations. There is no explicit appearance of synapse strengths or sigmoid curves. Thus our results will possess similar generality. Of course, one may always specialize the rate equations to account for particular choices of neuron models.

Rate Equations for Nearby Solutions

Let $r_i(t)$ be a solution of the rate equations and let $s_i(t)$ be a nearby solution at time t . Then, following the conventional approach presented in many places [see Drazin 1994, Jackson 1995, Ott 1993, Saaty 1981, etc.], we write

$$s_i(t) = r_i(t) + \mathbf{e}_i(t)$$

where the $\mathbf{e}_i(t)$ are the small differences between the two solutions which we take to be of first order in smallness. On substituting this into the rate equations and expanding in a Taylor series to first order we obtain

$$\begin{aligned} r_i(t) + \mathbf{e}_i(t) &= f_i(r_1(t-t_{1i}) + \mathbf{e}_1(t-t_{1i}), \dots, r_n(t-t_{ni}) + \mathbf{e}_n(t-t_{ni})) \\ &= f_i(r_1(t-t_{1i}), \dots, r_n(t-t_{ni})) + \sum_{k=1}^n \frac{\partial f_i}{\partial r_k} \mathbf{e}_k(t-t_{ki}) \end{aligned}$$

Since $r_i(t)$ is a solution of the rate equations we have

$$\mathbf{e}_i(t) = \sum_{k=1}^n \frac{\partial f_i}{\partial r_k} \mathbf{e}_k(t-t_{ki}) = \sum_{k=1}^n f_{ik} \mathbf{e}_k(t-t_{ki})$$

where $f_{ik} = \partial f_i / \partial r_k$ is the Jacobian matrix of the system. Note that it is evaluated at the retarded times $t - t_{ki}$.

Thus the first order differences satisfy a *linear* set of equations. To further simplify these equations so we may employ standard techniques we take advantage of the fact that the time delays t_{ki} are themselves small quantities of the first order. Thus we replace $\mathbf{e}_k(t-t_{ki})$ with $\mathbf{e}_k(t-\tau)$ where τ is the average value of the t_{ki} . Thus the linearized equations become

$$\mathbf{e}_i(t) = \sum_{k=1}^n f_{ik} \mathbf{e}_k(t-\tau)$$

We rewrite this as follows

$$\mathbf{e}_i(t+\tau) = \sum_{k=1}^n f_{ik} \mathbf{e}_k(t)$$

Stability Conditions

The linearized equations will represent a stable brain system if the differences change little with time. More precisely, the system will be stable if the differences remain within given upper and lower bounds as time advances.

Consider the eigenvalue problem

$$\sum_{k=1}^n f_{ik} u_k = \mathbf{I} u_i$$

Under conditions to be discussed in the next section this problem will have solutions u_k^p with the eigenvalues λ_p where $p = 1, \dots, n$. The eigenvalues are either real numbers or occur in complex conjugate pairs. The eigenvectors, u_k^p , form a complete set in \mathbf{R}^n , the n-dimensional space in which ε_k is defined. Thus it is possible to expand ε_k as follows

$$\mathbf{e}_k = \sum_{p=1}^n a_p u_k^p$$

Therefore

$$\mathbf{e}_i(t + \mathbf{t}) = \sum_{k=1}^n f_{ik} \mathbf{e}_k(t) = \sum_{p=1}^n a_p \sum_{k=1}^n f_{ik} u_k^p = \sum_{p=1}^n a_p \mathbf{I}_p u_i^p$$

and thus

$$\mathbf{e}_i(t + m\mathbf{t}) = \sum_{p=1}^n a_p (\mathbf{I}_p)^m u_i^p$$

From this last form it is clear that the magnitude of $\varepsilon_k(t)$ will remain within given bounds if and only if the following *stability conditions* hold.

$$|\lambda_p| \cong 1$$

Solving the Eigenvalue Problem

The eigenvalue problem $\sum_{k=1}^n f_{ik} u_k = \mathbf{I} u_i$ will be solvable if and only if the following condition holds

$$\det(\mathbf{F} - \lambda \mathbf{I}) = |\mathbf{F} - \lambda \mathbf{I}| = 0$$

where $\mathbf{F} = [f_{ik}]$ and \mathbf{I} is the n-dimensional identity matrix.

In treatments of determinants and matrices such as that of Aitken it is shown that the determinant $|\mathbf{F} - \lambda \mathbf{I}|$ may be expanded in the following *characteristic polynomial*.

$$|\mathbf{F} - \lambda \mathbf{I}| = |\mathbf{F}| - sp_{n-1}(\mathbf{F})\lambda + sp_{n-2}(\mathbf{F})\lambda^2 - \dots + (-1)^{n-1} sp_1(\mathbf{F})\lambda^{n-1} + (-1)^n \lambda^n$$

where $sp_k(\mathbf{F})$ stands for the sum of the principle minors of $|\mathbf{F}|$ of order k . A *minor* of $|\mathbf{F}|$ is a determinant of $|\mathbf{F}|$ obtained by suppressing any $n-k$ rows and $n-k$ columns of $|\mathbf{F}|$. A *principle minor* of $|\mathbf{F}|$ is a minor of $|\mathbf{F}|$ in which its elements are symmetrically arranged with respect to the main diagonal of \mathbf{F} . For example, the main diagonal elements of \mathbf{F} , f_{11} , f_{22} , ... f_{nn} , are each a principle minor of \mathbf{F} of order 1. The minors of a real matrix, including its principle minors, are real numbers.

Accordingly, $sp_1(\mathbf{F})$ is given by

$$sp_1(\mathbf{F}) = f_{11} + f_{22} + \dots + f_{nn} = tr(\mathbf{F})$$

which is the trace of \mathbf{F} as indicated.

In brains, neurons rarely, if ever, synapse upon themselves, so we have

$$f_{ii} = \frac{\sum f_i}{\sum r_i} = 0$$

Therefore

$$sp_1(\mathbf{F}) = tr(\mathbf{F}) = 0$$

We see that solving the eigenvalue problem is equivalent to finding the roots of the characteristic polynomial given above. Because the polynomial has real coefficients, its roots are either real or occur in complex conjugate pairs.

Stability Constraints

The characteristic polynomial introduced in the previous section has n roots, $\lambda_1, \lambda_2, \dots, \lambda_n$. Thus the characteristic polynomial may be written

$$\begin{aligned} |\mathbf{F} - \mathbf{I}\mathbf{I}| &= (\mathbf{I}_1 - \mathbf{I})(\mathbf{I}_2 - \mathbf{I}) \dots (\mathbf{I}_n - \mathbf{I}) \\ &= \mathbf{I}_1 \mathbf{I}_2 \dots \mathbf{I}_n - \sum_i \frac{\mathbf{I}_1 \mathbf{I}_2 \dots \mathbf{I}_n}{\mathbf{I}_i} \mathbf{I} + \sum_{i < j} \frac{\mathbf{I}_1 \mathbf{I}_2 \dots \mathbf{I}_n}{\mathbf{I}_i \mathbf{I}_j} \mathbf{I}^2 - \dots \\ &\quad + (-1)^{n-2} \sum_{i < j} \mathbf{I}_i \mathbf{I}_j \mathbf{I}^{n-2} + (-1)^{n-1} + (-1)^{n-1} \sum_i \mathbf{I}_i \mathbf{I}^{n-1} + (-1)^n \mathbf{I}^n \end{aligned}$$

By equating corresponding terms in the two forms of the characteristic polynomial we obtain the following:

$$\begin{aligned}
|\mathbf{F}| &= I_1 I_2 \cdots I_n \\
sp_{n-1}(\mathbf{F}) &= I_1 I_2 \cdots I_n \sum_i \frac{1}{I_i} \\
sp_{n-2}(\mathbf{F}) &= I_1 I_2 \cdots I_n \sum_{i < j} \frac{1}{I_i I_j} \\
&\dots \\
sp_2(\mathbf{F}) &= \sum_{i < j} I_i I_j \\
sp_1(\mathbf{F}) &= \sum_i I_i
\end{aligned}$$

or more compactly

$$sp_m(\mathbf{F}) = \sum_{k_1 < k_2 < \cdots < k_m} I_{k_1} I_{k_2} \cdots I_{k_m}$$

These equations in conjunction with the previous stability conditions

$$|\lambda_p| \cong 1$$

constitute a set of constraints that the matrix \mathbf{F} must satisfy for the dynamical system determined by \mathbf{F} to be stable. We call these the *stability constraints*.

Discussion

At first glance the stability constraints appear to be extremely severe. For example, the equation $|\mathbf{F}| = I_1 I_2 \cdots I_n$ and the condition $|\lambda_p| \cong 1$ seem to say that $|\det(\mathbf{F})| \cong 1$. When we remember that the dimensionality of \mathbf{F} is of the order of 10^{11} it would appear that constraining its determinant to be of magnitude unity would be extremely difficult to achieve. However, this view is misleading as we shall show.

Note that the fact that neurons do not synapse upon themselves implies the following

$$sp_1(\mathbf{F}) = tr(\mathbf{F}) = \sum_i I_i = 0$$

Let the eigenvalues λ_j be defined as follows

$$I_i = \pm 1 + \eta_i$$

where the η_i are small. Then we have

$$\sum_i I_i = \sum_i (\pm 1 + \mathbf{h}_i) = \sum_i (\pm 1) + \sum_i \mathbf{h}_i = 0$$

For the moment, assume that the number n of neurons is even. Assume further that half the λ_i are defined with $+1$, the other half with -1 . Then the sum $\sum(\pm 1)$ vanishes and we are left with

$$\sum_i \mathbf{h}_i = 0$$

Thus we have

$$\begin{aligned} |\mathbf{F}| &= \prod_i I_i = \prod_i (\pm 1 + \mathbf{h}_i) = (+1 + \mathbf{h}_1) \cdots (+1 + \mathbf{h}_m) (-1 + \mathbf{h}_{m+1}) \cdots (-1 + \mathbf{h}_{2m}) \\ &= (-1)^m + (\mathbf{h}_1 + \cdots + \mathbf{h}_m) (-1)^m + (\mathbf{h}_{m+1} + \cdots + \mathbf{h}_{2m}) (-1)^{m-1} + \mathcal{O}(\mathbf{h}^2) \\ &= (-1)^m (1 + \mathbf{h}_1 + \cdots + \mathbf{h}_m - \mathbf{h}_{m+1} - \cdots - \mathbf{h}_{2m}) + \mathcal{O}(\mathbf{h}^2) \end{aligned}$$

where $m = n/2$. Define $\mathbf{d}_i = \mathbf{h}_i - \mathbf{h}_{m+i}$ and drop terms of order higher than one.

$$(-1)^m |\mathbf{F}| = 1 + \mathbf{d}_1 + \cdots + \mathbf{d}_m = 1 + m \left(\frac{1}{m} \sum_i \mathbf{d}_i \right) = 1 + m \bar{\mathbf{d}}$$

where $\bar{\mathbf{d}}$ is the average of the deltas. Recall that m is large, on the order of 10^{10} , so we finally have

$$|\det(\mathbf{F})| \cong 1 + 10^{10} \bar{\mathbf{d}}$$

It is now clear that, unless $\bar{\mathbf{d}}$ is very small, $|\mathbf{F}|$ is not overly constrained in magnitude.

Nevertheless, the stability constraint equations can be useful in at least two ways. First, given a theoretic model in the form of specific rate equations, one may analyze the circumstances under which the constraints would be satisfied. Second, given a computer or simulation model, one may calculate the Jacobian and verify that the system is stable.

Specializing the Brain Model

In this section we specialize our brain model so that, in the following section, we may give an example of applying the stability criteria developed above. In specializing our model, we ensure it conforms to known properties of neurons.

The contribution neuron j makes to the membrane potential at the hillock of neuron i at time t is given by

$$v_{ij}(t) = v_{ij}(r_j(t - t_{ji})) = w_{ij}r_j(t - t_{ji})$$

We take the potential to be a linear function of the firing rate of neuron j at the earlier time t_{ji} . This means only that doubling the firing rate of neuron j at time $t - t_{ji}$ has the effect of doubling its contribution to the hillock membrane potential of the post synaptic neuron i at the later time t .

The total membrane potential at the hillock of neuron i at time t is then the sum of the contributions from all its pre synaptic neurons.

$$v_i(t) = \sum_{j=1}^n v_{ij}(t) = \sum_{j=1}^n w_{ij}r_j(t - t_{ji})$$

The firing rate of neuron i at time t is determined by its hillock membrane potential

$$r_i(t) = \mathbf{s}(v_i(t)) = a_i \mathbf{s}_0 \left(x_1 + (v_i(t) - v_1) \frac{\Delta x}{\Delta v_i} \right)$$

where a_i is a constant, $\Delta v_i = v_{i,2} - v_{i,1}$ is the range of possible values of the hillock membrane potential of neuron i , $\Delta x = x_2 - x_1$ is the *effective range* of the argument x of the sigmoid function $\sigma_0(x)$, and the sigmoid function is the frequently used

$$\mathbf{s}_0(x) = \frac{e^x}{1 + e^x}$$

The effective range of the sigmoid argument x simply allows us to map the membrane potential range Δv_i onto any desired range of x values, Δx . In particular it allows us to model the threshold value of the membrane potential that prevents the neuron firing for potential values that fall below the threshold .

Example

Here we apply the stability criteria to the specialized model of the previous section.

We note that negative values of the hillock membrane potential $v_i(t)$ completely inhibit the firing of neuron i . Consequently, we map the positive range of $v_i(t)$ onto the effective range Δx . We define Δx so that $\sigma_0(x)$ effectively vanishes outside the range. Thus

$$\Delta x = 3.0 - (-3.0) = 6.0$$

$$\Delta v_i = v_{i,\max} - 0 = v_{i,\max}$$

The firing rate of neuron i becomes

$$r_i(t) = a_i \mathbf{s}_0 \left(-3.0 + v_i(t) \frac{6}{v_{i,\max}} \right)$$

Now the elements of the Jacobian matrix may be computed as follows:

$$f_{ij} = \frac{\mathcal{J}r_i(t)}{\mathcal{J}r_j} = a_i \frac{d\mathbf{s}_0(x)}{dx} \frac{\mathcal{J}v_i(t)}{\mathcal{J}r_j} \frac{6}{v_{i,\max}} = \frac{6a_i}{v_{i,\max}} \mathbf{s}'_0(x(v_i(t))) w_{ij} = b_i \mathbf{s}'_{0,i} w_{ij}$$

where $b_i = 6a_i/v_{i,\max}$ and $\mathbf{s}'_{0,i} = \mathbf{s}'_0(x(v_i(t)))$

Therefore the determinant $|\mathbf{F}|$ becomes

$$|\mathbf{F}| = \begin{vmatrix} b_1 \mathbf{s}'_{01} w_{11} & b_1 \mathbf{s}'_{01} w_{12} & \cdots & b_1 \mathbf{s}'_{01} w_{1n} \\ b_2 \mathbf{s}'_{02} w_{21} & b_2 \mathbf{s}'_{02} w_{22} & \cdots & b_2 \mathbf{s}'_{02} w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n \mathbf{s}'_{0n} w_{n1} & b_n \mathbf{s}'_{0n} w_{n2} & \cdots & b_n \mathbf{s}'_{0n} w_{nn} \end{vmatrix} = \prod_{k=1}^n b_k \prod_{k=1}^n \mathbf{s}'_{0k} \begin{vmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{vmatrix}$$

$$= \prod_{k=1}^n b_k \prod_{k=1}^n \mathbf{s}'_{0k} |\mathbf{W}|$$

where $|\mathbf{W}|$ is the determinant of the matrix of synapse strengths w_{ij} . The coefficients b_k and synapse strengths w_{ij} are constant over the short term. Thus any time dependence of $|\mathbf{F}|$ comes from the factors \mathbf{s}'_{0k} .

Computing b_i

Consider the coefficients $b_i = 6a_i/v_{i,\max}$. The constant a_i represents the maximum possible firing rate of neuron i . We can estimate $v_{i,\max}$ in terms of a_i as follows

$$v_{i,\max} = \sum_{j=1}^n w_{ij}^+ r_{j,\max} = \sum_{j=1}^n w_{ij}^+ a_j$$

where w_{ij}^+ is defined as follows:

$$w_{ij}^+ = \begin{cases} w_{ij} & (w_{ij} > 0) \\ 0 & (w_{ij} \leq 0) \end{cases}$$

Thus $v_{i,\max}$ includes contributions from all the excitatory neurons synapsing on i at their maximum rates but none from any of the inhibitory neurons synapsing on i . The coefficient b_i becomes

$$b_i = \frac{6a_i}{\sum_j w_{ij}^+ a_j}$$

Assume Constant a_i

We now assume that all neurons have approximately the same maximum firing rate given by the inverse of the refractory period τ . Thus $a_i = 1/\tau$ and we obtain

$$v_{i,\max} = \frac{1}{\tau} \sum_{j=1}^n w_{ij}^+ = \frac{n}{\tau} w_{i,\text{avg}}^+$$

where $w_{i,\text{avg}}^+$ is the average strength of the excitatory synapses on neuron i . The coefficient b_i becomes

$$b_i = \frac{6a_i}{v_{\max}} = \frac{6}{w_{i,\text{avg}}^+}$$

Therefore the determinant $|\mathbf{F}|$ becomes

$$|\mathbf{F}| = \prod_{k=1}^n \frac{6}{w_{k,\text{avg}}^+} \prod_{k=1}^n \mathbf{s}'_{0k} |\mathbf{W}| = \prod_{k=1}^n (6\mathbf{s}'_{0k}) \overline{\mathbf{W}}$$

where $\overline{\mathbf{W}}$ is the matrix with elements $w_{ij}/w_{i,\text{avg}}^+$.

Estimating $|\mathbf{F}|$

The derivative of the sigmoid function σ_0 is given by

$$\mathbf{s}'_0 = \mathbf{s}_0(1 - \mathbf{s}_0)$$

Its maximum value occurs at $x = 0$ and equals $1/4$. Thus the term $6\mathbf{s}'_{0k}$ in the expression above for $|\mathbf{F}|$ is on the order of unity. Its exact value of course depends upon the current state of the system of neurons.

The elements of $\overline{\mathbf{W}}$ are also on the order or unity, or less since negative contributions have been excluded from $w_{i,\text{avg}}^+$.

Thus we see that the terms in the expression for $|\mathbf{F}|$ are all on the order of unity. This does not guarantee that the eigenvalues of \mathbf{F} will satisfy the stability conditions $|\lambda_p| \cong 1$, but clearly there is no evident inconsistency. Our previous discussion demonstrated that the

stability conditions could hold even though $|\mathbf{F}| = I_1 I_2 \cdots I_n$ might have a value far from unity.

References

- Aitken, A. C.**, *Determinants and Matrices*, (Oliver and Boyd (Interscience), New York NY, 1959)
- Cullen, C.G.**, *Matrices and Linear Transformations*, (Dover Publications, N.Y., 1990)
- Drazin, P.G.**, *Nonlinear Systems*, (Cambridge University Press, New York, 1994)
- Jackson, E. A.**, *Perspectives of Nonlinear Dynamics*, Vol. 1 & 2, (Cambridge University Press, New York, 1995)
- Kandel, E.R., Schwartz, J.H., and Jessell, T.M.**, *Essentials of Neural Science and Behavior*, (Appleton & Lange, Stamford Ct., 1995)
- Levy, H., and Lessman, F.**, *Finite Difference Equations*, (Dover Publications, N.Y., 1992)
- Mostow, G.D., Sampson, J.H., and Meyer, J.-P.**, *Fundamental Structures of Algebra*, (McGraw-Hill Book Co., N.Y., 1963)
- Ott, Edward**, *Chaos in Dynamical Systems*, (Cambridge University Press, New York, 1993)
- Ryon, J. W.**, "Neuron Networks as General Dynamic Systems", (NJIT, 1997)
- Schneider, H. and Barker, G.P.**, *Matrices and Linear Algebra*, (Dover Publications, N.Y., 1989)
- Saaty, Thomas L.**, *Modern Nonlinear Equations*, (Dover Publications, N.Y., 1981)