This was written some years ago for an Advanced Topics in Theoretical Neuroscience class, so it often says "Show that" – you can ignore that. I've filled in a few places where it used to say that. This is just FYI, for those wanting more background on how to analytically analyze the ring bump-attractor model.

The ring model.

As mentioned in class, the ring model is a workhorse of theoretical neuroscience, underlying the many "bump attractor" models, so it is worth knowing. It played a role historically in thinking about V1, though I don't think it is a good model of V1. The idea is to have a "Mexican hat" interaction – short-range excitation, long-range inhibition – with strong recurrence, to stabilize a "bump" of activity of fixed width, wherever it appears. The input selects where it appears but not its shape. The clever idea theoretically, developed by Haim Sompolinsky, was to find a version of this idea that could be solved analytically.

As applied to V1, the idea was to have the bump live on the ring of preferred orientations; the stimulus would choose the location of the bump but not its width, while the width of the bump, determined by the cortical circuitry, would yield stimulus-invariant (and thus contrast-invariant) orientation tuning. I don't think it applies to V1 because (1) as we've seen, the input already has contrast-invariant orientation tuning, once the DC is subtracted away; and (2) it's not true that orientation tuning is stimulus invariant – other stimulus attributes, such as spatial frequency, that change the tuning of the input, correspondingly change the tuning of cortical cells.

We start with a linear-threshold rate model. The neurons live on a ring of preferred stimulus variables or "orientations", θ . For V1, we should keep orientation between 0 and π , but for simplicity we'll let the ring go from 0 to 2π , this isn't hard to change. The firing rate of the neuron with preferred orientation θ is $r(\theta)$. Its inputs are inputs from other neurons, mediated by connection weights W; and external input $I(\theta)$ to the neuron at θ . The firing rate moves exponentially toward its threshold-rectified inputs:

$$\frac{d}{dt}r(\theta) = -r(\theta) + \left[\int_0^{2\pi} \frac{d\theta'}{2\pi} W(\theta - \theta')r(\theta') + I(\theta) - T\right]_+$$
(1)

Here T is the threshold; input less than T yields no output. The left side should be $\tau \frac{d}{dt}r(\theta)$ but I've set $\tau = 1$, which sets the time units.

I will sometimes write this more simply in vector notation as

$$\frac{d}{dt}\mathbf{r} = -\mathbf{r} + \left[\mathbf{W}\mathbf{r} + \mathbf{I} - \mathbf{T}\right]_{+} = -\mathbf{r} + \left[\mathbf{h} - \mathbf{T}\right]_{+}$$
(2)

where $\mathbf{h} \equiv \mathbf{Wr} + \mathbf{I}$ is the external input and \mathbf{T} is the vector all of whose elements are T. Here, the vector \mathbf{r} (or \mathbf{I}) can be thought of as having elements $\mathbf{r}_i = r(\theta_i)$, where the θ_i , i = 1, ..., N, are a discretized version of the continuus variable θ . Similarly, the matrix **W** has elements $W_{ij} = W(\theta_i - \theta_j) \frac{\Delta \theta}{2\pi}$ where $\Delta \theta$ is the interval between adjacent θ_i . A function is just a continuum limit of a vector – the discrete indices become a continuus variable – and so we will use the simpler vector notation and the more complete function notation interchangeably, although when it comes to calculation we will use functions.

Haim's clever trick is to choose cosine-tuned weights and cosine-tuned inputs:

$$W(\theta) = W_0 + 2W_1 \cos(\theta) \tag{3}$$

$$I(\theta) = I_0 + 2I_1 \cos(\theta - \theta_I) \tag{4}$$

Using cosine-tuned weights means that **Wr** only depends on the zeroeth and first harmonics of **r**: letting $r(\theta) = r_0 + 2r_1 \cos(\theta - \theta_r)$ + higher harmonics, the integral gives $W_0 r_0 + 2W_1 r_1 \cos(\theta - \theta_r)$:

$$\int_{0}^{2\pi} \frac{d\theta'}{2\pi} \cos(\theta - \theta') \cos(\theta' - \theta_r) =$$
(5)

$$\int_{0}^{2\pi} \frac{d\theta'}{2\pi} \frac{e^{i(\theta-\theta')} + e^{-i(\theta-\theta')}}{2} \frac{e^{i(\theta'-\theta_r)} + e^{-i(\theta'-\theta_r)}}{2} \tag{6}$$

$$= \frac{1}{4} \left(e^{i(\theta - \theta_r)} + e^{-i(\theta - \theta_r)} + \int_0^{2\pi} \frac{d\theta'}{2\pi} e^{i(\theta + \theta_r)} e^{-i2\theta'} + e^{-i(\theta + \theta_r)} e^{i2\theta'} \right)$$
(7)

$$= \frac{1}{4} \left(2\cos(\theta - \theta_r) - \frac{1}{2i} \left(e^{i(\theta + \theta_r)} (e^{-i4\pi} - 1) - e^{-i(\theta + \theta_r)} (e^{i4\pi} - 1) \right) \right)$$
(8)

$$=\frac{1}{2}\cos(\theta - \theta_r) \tag{9}$$

Thus,

$$\int_{0}^{2\pi} \frac{d\theta'}{2\pi} \left(W_0 + 2W_1 \cos(\theta - \theta') \right) (r_0 + 2r_1 \cos(\theta' - \theta_r)) = W_0 r_0 + 2W_1 r_1 \cos(\theta - \theta_r)$$
(10)

As a result, $\mathbf{h} = \mathbf{W}\mathbf{r} + \mathbf{I}$ depends only on the zeroeth and first harmonics of \mathbf{r} and is itself a sinusoid plus an F0:

$$h(\theta) = h_0 + 2h_1 \cos(\theta - \theta_h) \tag{11}$$

Clearly $h_0 = W_0 r_0 + I_0$. h_1 and θ_h are defined by

$$I_1 e^{-i\theta_I} + W_1 r_1 e^{-i\theta_r} = h_1 e^{-i\theta_h}$$
(12)

To see this, write the cosine in terms of complex exponentials to obtain:

$$I_1(e^{i(\theta-\theta_I)} + e^{-i(\theta-\theta_I)}) + W_1r_1(e^{i(\theta-\theta_r)} + e^{-i(\theta-\theta_r)}) = h_1(e^{i(\theta-\theta_h)} + e^{-i(\theta-\theta_h)})$$
(13)

or

$$e^{i\theta} \left(I_1 e^{-i\theta_I} + W_1 r_1 e^{-i\theta_r} - h_1 e^{-i\theta_h} \right) + e^{-i\theta} \left(I_1 e^{i\theta_I} + W_1 r_1 hi e^{i\theta_r} - h_1 e^{i\theta_h} \right) = 0$$
(14)

This is of form $e^{i\theta}X + e^{-i\theta}X^* = 0$, where X^* means complex conjugate of X. To be true for all θ , it must be the case that $X = X^* = 0$. X = 0 gives Eq. 12.

Since the dynamics of \mathbf{r} depend only on $[\mathbf{Wr} + \mathbf{I} - \mathbf{T}]_+$, they also depend only on the zeroeth and first harmonics of \mathbf{r} . Thus if we can solve for the these harmonics of \mathbf{r} , then we can derive a complete solution for \mathbf{r} ; in particular, the steady state value, \mathbf{r}_{SS} , is just a rectified sinusoid, i.e. a "bump": $\mathbf{r}_{SS} = [\mathbf{Wr}_{SS} + \mathbf{I} - \mathbf{T}]_+ = [\mathbf{h} - \mathbf{T}]_+$.

Define the Fourier transform of a function $f(\theta)$ by

$$\hat{f}_n = \int_0^{2\pi} \frac{d\theta}{2\pi} f(\theta) e^{in\theta} \equiv \mathcal{F}_n[f(\theta)]$$
(15)

$$f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}_n e^{in\theta} \equiv \mathcal{F}_n^{-1}[f(\theta)]$$
(16)

Show (as always, don't bother if it's obvious to you) that if $f(\theta) = f_0 + 2f_1 \cos(\theta - \theta_f)$, then $\hat{f}_0 = f_0$, $\hat{f}_1 = f_1 e^{-i\theta_f}$, and $\hat{f}_{-1} = \hat{f}_1^*$ (where the * indicates complex conjugate). Thus, Eq. 12 just states that $\hat{h}_1 = \hat{J}_1 \hat{r}_1 + \hat{I}_1$. More generally, **show** that $\hat{h}_n = \hat{J}_n \hat{r}_n + \hat{I}_n$, although all of the terms (except the r_n 's) are 0 for |n| > 1.

The dynamics for the n^{th} harmonic are given simply by

$$\frac{d}{dt}\hat{r}_n = -r_n + \mathcal{F}_n\left[[\mathbf{h} - \mathbf{T}]_+\right] \tag{17}$$

We solve this as follows:

• $h_0 - T > 2h_1$: the equation is linear: $[\mathbf{h} - \mathbf{T}]_+ = \mathbf{h} - \mathbf{T}$. Thus Eq. 17 becomes simply

$$\frac{d}{dt}\hat{r}_n = -\hat{r}_n + \hat{h}_n - \delta_{n0}\mathbf{T}$$
(18)

$$= = -\hat{r}_n(1 - \hat{W}_n) + \hat{I}_n - \delta_{n0}\mathbf{T}$$
(19)

So the fixed point is $\frac{\hat{I}_n - \delta_{n0} \mathbf{T}}{1 - \hat{W}_n}$.

If $\hat{W}_n > 1$, the equation for \hat{r}_n is unstable and \hat{r}_n grows exponentially away from the fixed point, which is unstable;

If $\hat{W}_n < 1$, \hat{r}_n is instantaneously exponentially decaying towards this fixed point, which is stable, with time constant $1/(1-\hat{W}_n)$.

• $h_0 - T < -2h_1$: $[\mathbf{h}]_+ = 0$ and \mathbf{r} decays to 0.

The remainder of the problem is devoted to the interesting case: $-2h_1 < h_0 - T < 2h_1$: For this case, we define the angle ψ , which is the width of the activity bump, by $\cos \psi = -\frac{h_0 - T}{2h_1}$. Thus, $h(\theta) - T = 2h_1 (\cos(\theta - \theta_h) - \cos\psi)$ and $[h(\theta) - T]_+ = 0$ for $|\theta - \theta_h| > \psi$. **Compute** $\mathcal{F}_0[[h(\theta) - T]_+]$ and $\mathcal{F}_1[[h(\theta) - T]_+]$ (\mathcal{F}_{-1} is just the complex conjugate of \mathcal{F}_1), using

$$\mathcal{F}_{n}\left[\left[h(\theta)\right]_{+}-T\right] = 2h_{1} \int_{\theta_{h}-\psi}^{\theta_{h}+\psi} \frac{d\theta}{2\pi} e^{in\theta} \left[\cos(\theta-\theta_{h})-\cos\psi\right]$$
(20)

$$= 2h_1 e^{-in\theta_h} \int_{-\psi}^{\psi} \frac{d\theta}{2\pi} e^{in\theta} \left[\cos\theta - \cos\psi\right]$$
(21)

$$\equiv 2h_1 e^{-in\theta_h} G_n(\psi) \tag{22}$$

That is, solve for the functions $G_0(\psi)$, $G_1(\psi)$. You should find

$$G_0(\psi) = \frac{1}{\pi} \left(\sin \psi - \psi \cos \psi \right)$$
(23)

$$G_1(\psi) = \frac{1}{2\pi} \left(\psi - \frac{\sin(2\psi)}{2} \right)$$
(24)

Graph G_0 and G_1 over the range $0 \le \psi \le \pi$ to get a sense of how they behave. Both are monotonically increasing, from 0 to 1 (G_0) or from 0 to 1/2 (G_1).

Note that we can rewrite the answers as follows:

$$\mathcal{F}_0[[h(\theta)]_+ - T] = 2h_1 G_0(\psi) = (h_0 - T) \frac{-G_0(\psi)}{\cos\psi} = (h_0 - T) \frac{G_0(\psi)}{\cos(\psi + \pi)}$$
(25)

$$\mathcal{F}_{1}[[h(\theta)]_{+} - T] = 2h_{1}e^{-i\theta_{h}}G_{1}(\psi) = \hat{h}_{1}[2G_{1}(\psi)]$$
(26)

Verify that for $\psi = \pi$ (no rectification) these become $h_0 - T$ and \hat{h}_1 , respectively, as they should; while for $\psi = 0$, both are 0.

We can unify our dynamical equations for all conditions, using $h_1 e^{-i\theta_h} = \hat{h}_1$, as:

$$\frac{d}{dt}r_0 = -r_0 + \frac{W_0r_0 + I_0 - T}{\cos(\psi + \pi)}G_0(\psi)$$
(27)

$$= -r_0 + e^{i\theta_h} (W_1 \hat{r}_1 + \hat{I}_1) [2G_0(\psi)], \qquad 0 < \psi < \pi$$
(28)

$$\frac{d}{dt}\hat{r}_1 = -\hat{r}_1 + (W_1\hat{r}_1 + \hat{I}_1)[2G_1(\psi)]$$
(29)

$$\psi = \cos^{-1}\left(\frac{-(h_0 - T)}{2h_1}\right), \quad |h_0| < 2h_1;$$
(30)

$$= \pi, \qquad h_0 - T > 2h_1; \tag{31}$$

$$= 0, \qquad h_0 - T < -2h_1; \tag{32}$$

(33)

Write $\hat{r}_1 = r_1 e^{-i\theta_r}$ and $\hat{I}_1 = I_1 e^{-i\theta_I}$, multiply the equation for $\frac{d}{dt}\hat{r}_1$ by $e^{i\theta_r}$, and separate real and imaginary parts to obtain separate equations for $\frac{d}{dt}r_1$ and for $\frac{d}{dt}\theta_r$. We thus have reduced the nonlinear dynamics to a set of 3 equations in the 3 real variables r_0 , r_1 , and θ_r .

You should find that $\frac{d}{dt}\theta_r \propto \sin(\theta_I - \theta_r)$, that is, θ_r evolves to the input angle θ_I . Show that, once $\theta_r = \theta_I$, then $\theta_h = \theta_r$, $h_1 = W_1r_1 + I_1$, and the equations reduce to the following two equations for the 2 real variables r_0 and r_1 :

$$\frac{d}{dt}r_0 = -r_0 + \frac{W_0r_0 + I_0 - T}{\cos(\psi + \pi)}G_0(\psi)$$
(34)

$$= -r_0 + (W_1 r_1 + I_1) [2G_0(\psi)], \qquad 0 < \psi < \pi$$
(35)

$$\frac{d}{dt}r_1 = -r_1 + (W_1r_1 + I_1) \left[2G_1(\psi)\right]$$
(36)

where

$$\psi = \cos^{-1} \left(-\frac{W_0 r_0 + I_0 - T}{2(W_1 r_1 + I_1)} \right), \qquad |W_0 r_0 + I_0 - T| < 2(W_1 r_1 + I_1); \tag{37}$$

$$= \pi, \qquad W_0 r_0 + I_0 - T > 2(W_1 r_1 + I_1); \tag{38}$$

$$= 0, \qquad W_0 r_0 + I_0 - T < -2(W_1 r_1 + I_1)$$
(39)

We now focus on understanding the steady-state solutions of these equations. Note that in any steady state with $0 < \psi < \pi$, $r_1/r_0 = G_1(\psi)/G_0(\psi)$. We first consider the case of a uniform suprathreshold input, $I_1 = 0$ with $I_0 > T$. Show that for $W_0 \neq 1$:

- 1. There is always a uniform steady-state solution with $r_1 = 0$, $\psi = \pi$ (write down the solution for r_0). This solution is unstable to perturbations of r_0 for $W_0 > 1$, stable for $W_0 < 1$.
- 2. There is also a "bump" solution specified by $2G_1(\psi)W_1 = 1$ for $W_1 > 1$, with $0 < \psi < \pi$. This represents a "spontaneous" emergence of an orientation-tuned response. For this solution, write down the solutions for r_0 and r_1 . Note that θ_r , the location of the peak of the bump, is arbitrary there is a solution for any choice of θ_r . (There are also bump solutions for $W_1 = 1$, with $\psi = \pi$; in this case, r_1 may take any value consistent with $\psi = \pi$, i.e. $r_0 > 2(r_1 + I_1)$.)
- 3. The uniform solution is unstable to perturbations of r_1 for $W_1 > 1$ in this case, any perturbation of r_1 from zero leads to the bump solution (or else to instability). (For $W_1 = 1$, the uniform solution is just the choice $r_1 = 0$ in the range of allowed bump solutions – these solutions are all marginally stable to perturbations of r_1). It is stable to perturbations of r_1 for $W_1 < 1$.
- 4. Determine the conditions for stability of the bump solution. Linearize the dynamics about the bump solution: show that if $\cos \psi = -x$, then $\frac{d}{dx}\psi = \frac{1}{\sin\psi}$, $\frac{d}{dx}G_0(\psi) = \frac{\psi}{\pi}$, and $\frac{d}{dx}G_1(\psi) = \frac{\sin\psi}{\pi}$. Use these to determine the dynamical equations to first order

in Δr_0 and Δr_1 (defined as the perturbations of r_0 and r_1 from their bump-solution values). I get that the matrix driving the linearized dynamics is

$$\frac{1}{\pi} \begin{pmatrix} W_0 \psi - \pi & 2W_1 \sin \psi \\ W_0 \sin \psi & W_1 (\psi + \sin \psi \cos \psi) - \pi \end{pmatrix}$$
(40)

Since $W_1 = \frac{1}{2G_1(\psi)}$, this can be rewritten

$$\frac{1}{\pi} \begin{pmatrix} W_0 \psi - \pi & 2\pi \frac{\sin \psi}{\psi - \sin \psi \cos \psi} \\ W_0 \sin \psi & 2\pi \frac{\sin \psi \cos \psi}{\psi - \sin \psi \cos \psi} \end{pmatrix}$$
(41)

or

$$\frac{1}{\pi} \begin{pmatrix} W_0 \psi - \pi & 4\pi \frac{\sin \psi}{2\psi - \sin 2\psi} \\ W_0 \sin \psi & 2\pi \frac{\sin 2\psi}{2\psi - \sin 2\psi} \end{pmatrix}$$
(42)

To be stable, both eigenvalues of \mathbf{M} should have negative real part, or equivalently (for a 2 × 2 real matrix), Trace $\mathbf{M} < 0$ and det $\mathbf{M} > 0$. The trace condition is

$$W_0 < \frac{\pi}{\psi} \left(1 - \frac{2\sin 2\psi}{2\psi - \sin 2\psi} \right) = \frac{\pi}{\psi} \left(1 - \frac{\sin 2\psi}{2\pi G_1(\psi)} \right)$$
(43)

For the determinant condition, I use $2\psi - \sin 2\psi > 0$ and note $\psi \sin 2\psi - 2\sin^2 \psi = -2\pi G_0(\psi) \sin \psi$ which is < 0 for $0 < \psi < \pi$, and find that it reduces to

$$W_0 < -\frac{\cos\psi}{G_0(\psi)} \tag{44}$$

This requires $\cos \psi \leq 0$ and so requires $\pi/2 \leq \psi \leq \pi$.

It's quite possible I made a calculational error somewhere, so you should certainly check these results. But assuming they're correct: examining these conditions graphically, I find that the determinant condition is always more stringent than the trace condition, and more generally get the results shown in Fig. 1.

Now consider a more general input, that is, $I_1 \ge 0$. Show that:

- 1. There is a steady state with $\psi = 0$ (and $r_0 = r_1 = 0$) if and only if $I_0 + 2I_1 < T$; this state is stable.
- 2. There is a steady state with $\psi = \pi$, $r_0 > 0$, $r_1 \ge 0$, if and only if $W_0 < 1$, $W_1 < 1$, $I_0 T > 0$, and $\frac{I_0 T}{1 W_0} \ge \frac{2I_1}{1 W_1}$; the latter condition just says $r_0 \ge 2r_1$, as required for a solution with $\psi = \pi$. This solution is stable if $r_0 > 2r_1$. (I haven't worked out the stability for $r_0 = 2r_1$.)



Figure 1: Conditions for stability. X-axis: ψ . Red line: upper bound for W_0 under Eq. 44. Blue line: upper bound for W_0 under Eq. 43. Green line: value of $W_1 = 1/2G_1(\psi)$. As can be seen, a stable bump solution arises only for $1 < W_1 < 2$, and requires $W_0 < X$ where X, described by the red curve, ranges from 0 for $W_1 = 2$ to 1 for $W_1 = 1$. Empirically, it appears that the condition for stability is approximately $W_1 + W_0 < 2$, and indeed numerically adding the red and green curves (not shown) yields a value that varies between 2.0 and 2.08.

3. Characterize the "bump" solutions with $0 < \psi < \pi$ and $I_1 \neq 0$. This part of the problem is open-ended ...

The linearized dynamics is still given by Eq. 40 (show this if you didn't show it in general previously), giving the stability conditions

$$W_0\psi + W_1(\psi + \sin\psi\cos\psi) < 2\pi \tag{45}$$

and

$$(W_0\psi - \pi)(W_1(\psi + \sin\psi\cos\psi) - \pi) - 2W_0W_1\sin^2\psi > 0$$
(46)

Starting with the expressions for the steady-state solutions for r_0 and r_1 in terms of ψ and the parameters, use these in the expression for $\cos \psi$ (Eq. 37) to find that either $\cos \psi = 0$ (which implies $I_0 - T = 0$, $I_1 > 0$) or else ψ is implicitly defined by

$$(1 - 2W_1 G_1(\psi)) - \frac{2I_1}{I_0 - T} \left(W_0 G_0(\psi) - \cos \psi \right) = 0$$
(47)

Let $\kappa = \frac{2I_1}{I_0 - T}$, representing the degree of departure of the input from spatial homogeneity. For $\kappa = 0$ we recover the previous solution, $G_1(\psi) = \frac{1}{2W_1}$. An alternative way to parameterize the input is as in Ben-Yishai and Sompolinsky, who set $2I_1 = \epsilon I$, $I_0 = (1 - \epsilon)I$ (they restrict to non-negative input, meaning $0 \le \epsilon \le 0.5$), and T = 1, so that $\frac{2I_1}{I_0 - T} = \frac{\epsilon I}{I - 1 - \epsilon I} = \frac{\epsilon}{1 - \epsilon - 1/I}$. Analyzing a somewhat different model, described below, they argued that there is a robust regime where ψ shows little dependence on I for I > 1 ("contrast-invariance" of tuning), and also shows little dependence on ϵ for $I \gg 1$ (which is equivalent to T being negligible). For the present model, clearly ψ will show little dependence on I for $I \gg 1$, but I don't think either of the above two claims will be true.

Numerically study the dependence of ψ on ϵ and I or on κ . If you're ambitious, you can, for example, do a scatterplot of stable solutions in the 3-D space of W_0 , W_1 and ψ for a range of values of κ , with different colors for different values of κ . Figures 2-3 give some snapshots of what I got for such scatterplots.

Note: Ben-Yishai and Sompolinsky (1995), who originated this model, used an inputoutput function with a maximum as well as a minimum – $\max([\mathbf{h} - \mathbf{T}]_+, \beta)$ in place of $[\mathbf{h} - \mathbf{T}]_+$. This seems to have made calculations much simpler. It allowed stable solutions for large W_1 and W_0 (large relative to $1/\beta$), allowing calculation in the large-W limit; these solutions were essentially square bumps with ψ defined by $h(\theta_h \pm \psi) = T$ and a very narrow transition from $r(\theta_h - \psi) = 0$ to $r(\theta_h - (\psi - \epsilon)) = \beta$. These bump solutions also could be much narrower than the stable bump solutions we found here for $I_1 = 0$, which are restricted



Figure 2: Different views of the locations of stable bump solutions (simultaneous solutions of Eqs. 45-47) for different values of κ (indicated as 'k' in legend).



Figure 3: More views of the locations of stable bump solutions (simultaneous solutions of Eqs. 45-47) for different values of κ (indicated as 'k' in legend).

to $\psi \geq \pi/2$. I should probably have used that version ... there's a reason why they decided to use it.

In Goldberg, Rokni, and Sompolinsky (2004), they analyze the present model (linear threshold), but with a couple of differences: (1) $W_0 = 0$; (2) The input is spatially white noise with variance σ^2 , rather than being structured (orientation-tuned). They work out the phase diagram for existence of bump solutions in the plane of σ and W_1 , They find, as here for the case $I_1 = 0$, that stable bump solutions only exist for $1 < W_1 < 2$, but the range becomes more restricted (the lower bound on W_1 increases) as the noise increases until finally the bump solutions disappear.