LETTER Communicated by Bard Ermentrout

Oscillatory Networks: Pattern Recognition Without a Superposition Catastrophe

Thomas Burwick
thomas.burwick@neuroinformatik.rub.de
Institut für Neuroinformatik, Ruhr-Universität Bochum, 44306 Bochum, Germany

Using an oscillatory network model that combines classical network models with phase dynamics, we demonstrate how the superposition catastrophe of pattern recognition may be avoided in the context of phase models. The model is designed to meet two requirements: on and off states should correspond, respectively, to high and low phase velocities, and patterns should be retrieved in coherent mode. Nonoverlapping patterns can be simultaneously active with mutually different phases. For overlapping patterns, competition can be used to reduce coherence to a subset of patterns. The model thereby solves the superposition problem.

1 Introduction

For a network with two or more active patterns, where activity is solely defined by on and off states, a subsequent stage of information processing may not be able to read out single patterns. This so-called superposition catastrophe has long been recognized as a major challenge for neural network modeling (Rosenblatt, 1961). It is related to the binding problem (see Roskies, 1999; Müller, Elliott, Herrmann, & Mecklinger, 2001).

The superposition and binding problems motivated von der Malsburg to propose grouping of neural units based on temporal correlation. This allows several nonoverlapping patterns to be active at the same time and still be separable due to different temporal properties (von der Malsburg, 1981, 1985). For example, they may synchronize with different phases for different patterns (von der Malsburg & Schneider, 1986). Subsequently, associative memory based on temporal coding has been implemented in oscillatory neural networks. Recognized patterns may then correspond to limit cycles. The first models that implemented temporal coding were based on Wilson-Cowan-like dynamics, where the oscillators were defined in terms of coupled excitatory and inhibitory units (von der Malsburg & Schneider, 1986; Baird, 1986; Freeman, Yao, & Burke, 1988; Li & Hopfield, 1989). These approaches have been extended in Wang, Buhmann, & von der Malsburg, (1990) and von der Malsburg and Buhmann (1992). Segmentation is a particular example of the superposition problem. Correspondingly, many applications are concerned with this task (e.g., Terman & Wang, 1995; Wang &

Terman, 1995, 1997) with applications such as medical image segmentation (Shareef, Wang, & Yagel, 1999).

Oscillatory networks with associative memory were also studied by using a phase description for the oscillators. Whenever such a parameterization is possible, the analysis of oscillatory systems may be simplified significantly (see Kuramoto, 1984; Winfree, 2001). In this article, we study the avoidance of the superposition catastrophe in the context of phase models. Our discussion is based on a network model, where the real-valued activity u_k of each unit k is complemented with a phase θ_k, $k = 1, \ldots, N$. The phases θ_k are supposed to parameterize a temporal structure of the signals. We consider a generalization of classical neural network dynamics:

$$\tau_u \frac{du_k}{dt} = -u_k + \frac{1}{N} \sum_{l=1}^{N} w_{kl}(\theta)\, g(u_l) + I_k, \tag{1.1a}$$

$$\tau_\theta \frac{d\theta_k}{dt} = 2\pi\ g(u_k)\ (1 + \mathcal{S}_k(u, \theta)), \tag{1.1b}$$

where

$$w_{kl}(\theta) = g_{kl}\,(\alpha + \beta \cos(\theta_k - \theta_l)), \tag{1.2}$$

$$\mathcal{S}_k(u, \theta) = \frac{\beta}{N} \sum_{l=1}^{N} g_{kl}\ \sin(\theta_l - \theta_k)\ g(u_l). \tag{1.3}$$

The activation function is $g(x) = (1 + \tanh(x))/2$, the $\tau_u, \tau_\theta > 0$ are timescales, and the I_k are external inputs. The $w_{kl}(\theta)$ are phase-dependent weights, specified by the g_{kl} and real-valued parameters $\alpha,\ \beta \geq 0$. In our examples, we use $\alpha > \beta$. In accordance with the mentioned interpretation of the phases, the model is designed so that the limit 1(0) of $g(u_k)$ describes an on(off) state, leading to high(low) frequencies of θ_k. This motivates the factor $g(u_k)$ on the right-hand side of equation 1.1b. The $\mathcal{S}_k(u, \theta)$ will correct the phase velocity and imply synchronization. For on-states with $g(u_k) \to 1$, the phase velocity will approach $(2\pi/\tau_\theta)(1 + \mathcal{S}_k)$, while for off-states, $g(u_k) \to 0$, the phase velocity will vanish. (For more detailed discussion and motivation of $w_{kl}(\theta),\ \mathcal{S}_k(u, \theta)$ see below.)

Understanding a possible relevance of equations 1.1 to biology would be of interest. An interpretation of the units in terms of single biological neurons is not intended. Possibly an interpretation in terms of populations of neurons may be found. An exact interpretation of variables in biological terms, however, is outside the scope of this letter. For the purpose of this letter, it is sufficient to see the model as a formal framework that allows the implementation of the benefits that temporal coding should hold for pattern recognition.

Specifying an oscillatory network model requires two choices to be made. First, a model for the single oscillators without coupling has to be chosen. Second, the coupling between the oscillators has to be specified. A common choice for the oscillator dynamics is the Stuart-Landau model. This assigns a complex-valued dynamics to the oscillator and may be derived from a general ordinary differential equation near the Hopf bifurcation (see Kuramoto, 1984). A natural choice for the couplings is linear in the complex-valued normal form coordinate (Kuramoto, 1975). The complete system may then be interpreted as a generalized version of the discrete complex-valued Ginzburg-Landau reaction-diffusion system. In the following, we refer to it as the generalized Ginzburg-Landau (GGL) model. Most approaches toward associative memory in the context of models with phases and amplitudes use the GGL model with vanishing shear (see Hoppensteadt & Izhikevitch, 2003, for a short review). Using pure phase models as an adiabatic approximation, associative memory has been implemented by identifying the coupling strengths with Hebbian weights (Abbott, 1990; Sompolinsky, Golomb, & Kleinfeld, 1990, 1991; Baldi & Meir, 1990; Kuramoto, Aoyagi, Nishikawa, Chawanya, & Okuda, 1992; Sompolinsky & Tsodyks, 1992; Kuzmina & Surina, 1994; Kuzmina, Manykin, & Surina, 1995). The corresponding step in the context of phase and amplitude dynamics was taken by using a complexified analog of the Cohen-Grossberg-Hopfield function (Takeda & Kishigami, 1992; Chakravaraty & Gosh, 1994; Aoyagi, 1995; Hoppensteadt & Izhikevitch, 1996). This approach imitated the method that was used for discrete dynamics (Noest, 1988a, 1988b). Reviews of complex-valued neural networks and their application to pattern recognition may also be found in Hirose (2003).

The GGL model is the simplest model for coupling Stuart-Landau oscillators, but not the only possible one. Alternative models have been studied with and without regard to associative memory. For example, Tass and Haken proposed three different models and used one to model synchronization of neural activity in the visual cortex (Tass, 1993; Tass and Haken, 1996a, 1996b). In the context of associative memory, the Stuart-Landau oscillator of the GGL model was replaced with an oscillator model that has on and off states (Aoyagi, 1995). A phase-locked loop model was introduced to allow implementations with well-developed circuit technology (Hoppensteadt & Izhikevitch, 2000). Another model was proposed that allowed a smooth transition between fixed-point mode and oscillatory mode by appropriately changing a parameter (Chakravarathy & Gosh, 1996). We will also find that the model of equations 1.1 is *not* of the GGL type.

In section 2, we review associative memory based on the GGL model with vanishing shear and identical natural frequencies. We will argue that the phase dynamics of this model is not suited to solve the superposition problem by giving different phases to different patterns and the background. In the context of weakly coupled GGL models, a solution has been proposed that relates the pooling of neural oscillators to frequency

gaps resulting from nonidentical natural frequencies (frequency modulation, FM; see Hoppensteadt & Izhikevitch, 1997, sec. 5.4.2). In section 3, we take a different approach of modifying the underlying system and starting from equations 1.1 instead. This will allow the use of phase gaps to prevent the superposition catastrophe. We give a detailed comparison of equations 1.1 and the GGL model and also relate our approach to the FM proposal. In section 4, we present simple examples. In section 5, we conclude with a summary and outlook.

2 The Generalized Ginzburg-Landau Model

We now review and discuss associative memory based on the GGL model. Notice, however, that we do not discuss the storage of complex-valued patterns. The changes may be included without difficulty by going to complex-valued Hermitean weights (Noest, 1988a, 1988b; Takeda & Kishigami, 1992; Chakravarathy & Gosh, 1994; Aoyagi, 1995). We present the models in a form that will allow a convenient comparison with equations 1.1 in section 3.

2.1 The GGL Model as Gradient System. The discrete complex Ginzburg-Landau (GGL) model may be expressed in terms of complex coordinates (Kuramoto, 1975; see also Hoppensteadt & Izhikevitch, 1997, sec. 10.1–3):

$$\text{GGL: } \frac{dz_k}{dt} = \left(\tilde{I}_k + i\omega_k\right) z_k - (\sigma_k + i\eta_k) z_k |z_k|^2 + \frac{\beta}{N} \sum_{l=1}^{N} g_{kl} z_l, \tag{2.1}$$

where $\tilde{I}_k, \omega_k, \sigma_k, \eta_k$, and the weights g_{kl} are real-valued parameters, $k = 1, \ldots, N$. Associative memory has been introduced for the parameter choices $\eta_k = 0$, corresponding to vanishing shear, and $\omega_k = \Omega$ (see Hoppensteadt & Izhikevitch, 1997, sec. 10.4, and the short review in Hoppensteadt & Izhikevitch, 2003). The latter choice allows us to set $\omega_k = 0$ by going to the comoving frame, $\theta_k \to \theta_k + \Omega t$. We may also set $\sigma_k = 1$. With $z_k = V_k \exp(i\theta_k)$, $\bar{z}_k = V_k \exp(-i\theta_k)$, where V, θ are real, equation 2.1 is then equivalent to

$$\text{GGL: } \frac{dV_k}{dt} = \tilde{I}_k V_k - V_k^3 + \frac{1}{N} \sum_{l=1}^{N} \widetilde{w}_{kl}(\theta) V_l, \tag{2.2a}$$

$$\text{GGL: } \frac{d\theta_k}{dt} = \frac{1}{V_k} \widetilde{\mathcal{S}}_k(V, \theta), \tag{2.2b}$$

where $\widetilde{w}_{kl}$ is w_{kl} with $\alpha = 0$, as given by equation 1.2, and $\widetilde{\mathcal{S}}_k$ is obtained from $\mathcal{S}_k$ by replacing $g(u_k) \to V_k$. The $g_{kl} \sin(\theta_l - \theta_k) V_l$ terms in $\widetilde{\mathcal{S}}_k(V, \theta)$ tend to synchronize (desynchronize) unit k with units l whenever $g_{kl} > 0$ ($g_{kl} < 0$). Without couplings, $g_{kl} = 0$, it is easily seen that for $\tilde{I}_k < 0$, the origin $V_k = 0$ is a stable fixed point; at $\tilde{I}_k = 0$, the unit k experiences a Hopf bifurcation, generating a stable limit cycle with radius $\sqrt{\tilde{I}_k}$ for $\tilde{I}_k > 0$.

Introducing the Lyapunov function,

$$\widetilde{\mathcal{L}}(V, \theta) = -\frac{1}{2N} \sum_{k,l} \widetilde{w}_{kl}(\theta) V_k V_l + \widetilde{\mathcal{P}}(V) \tag{2.3}$$

with

$$\widetilde{\mathcal{P}}(V) = \sum_k \left(-\frac{\tilde{I}_k}{2} V_k^2 + \frac{1}{4} V_k^4 \right) \tag{2.4}$$

allows us to express the dynamics of equations 2.2 as a gradient system (Noest, 1988a, 1988b; Takeda & Kishigami, 1992; Chakravarathy & Gosh, 1994; Aoyagi, 1995; Hoppensteadt & Izhikevitch, 1996). Assuming that the g_{kl} are symmetric, one obtains

$$\text{GGL: } \frac{dV_k}{dt} = -\frac{\partial}{\partial V_k} \widetilde{\mathcal{L}}(V, \theta), \tag{2.5a}$$

$$\text{GGL: } \frac{d\theta_k}{dt} = -\frac{1}{V_k^2} \frac{\partial}{\partial \theta_k} \widetilde{\mathcal{L}}(V, \theta). \tag{2.5b}$$

The minima of $\widetilde{\mathcal{L}}(V, \theta)$ should then correspond to storage of patterns. In section 2.2, we specify the storage of patterns and relate $\widetilde{\mathcal{L}}(V, \theta)$ to a coherence measure for the patterns.

2.2 Memory, Competition, and Coherence. We consider P patterns with components $\xi_k^p \in \{0, 1\}, k = 1, \ldots, N$. Using these patterns, we specify the weights g_{kl} of equations 1.2 and 1.3 as

$$g_{kl} = d_{kl} + \underbrace{\frac{1}{P} \sum_p \xi_k^p \xi_l^p}_{\equiv h_{kl}(\xi)} + \lambda \underbrace{\frac{-1}{P^2} \sum_{p \neq q} \xi_k^p \xi_l^q}_{\equiv c_{kl}(\xi)} . \tag{2.6}$$

The couplings d correspond to a background geometry, the excitatory part h corresponds to Hebbian memory, and the inhibitory part c establishes

competition between the patterns with $\lambda \geq 0$. Analogously, any inhibitory part of d establishes competition between units.

The complex coordinates of equation 2.1 may be used to define a coherence measure C_p for each pattern p:

$$Z_p(V,\theta) = \frac{1}{N_p}\sum_k \xi_k^p z_k = C_p \exp(i\Psi_p), \tag{2.7}$$

where Z_p, C_p, Ψ_p depend on V, θ and $|Z_p| = C_p$, $0 \leq C_p \leq 1$. These correspond to the Kuramoto coherence measure restricted to the active parts of the stored patterns (Kuramoto, 1984: see Strogatz, 2000, for a review of the Kuramoto model and references to recent results). We call Ψ_p the phase of pattern p and C_p its coherence. N_p denotes the number of nonzero components of pattern p.

Using equation 2.6 with vanishing background geometry, that is, $d = 0$, equation 2.3 may then be written as

$$\begin{aligned}\widetilde{\mathcal{L}} &= -\frac{\beta}{2N}\sum_{k,l} g_{kl}\cos(\theta_k - \theta_l)\, V_k V_l + \widetilde{\mathcal{P}} \\ &= -\frac{\beta}{2N}\mathrm{Re}\sum_{k,l} g_{kl}\exp\{i(\theta_k - \theta_l)\}\, V_k V_l + \widetilde{\mathcal{P}} \\ &= -\frac{\beta}{2}\frac{N}{P}\left(\sum_p \rho_p^2 C_p^2 - \frac{\lambda}{P}\sum_{p\neq q}\rho_p\rho_q\, C_p C_q \cos(\Psi_p - \Psi_q)\right) + \widetilde{\mathcal{P}},\end{aligned} \tag{2.8}$$

where $\rho_p = N_p/N$ denotes the density of pattern p. Since $\widetilde{\mathcal{L}}$ could be given this form, without competition, that is, $\lambda = 0$, equations 2.5 imply that the β-dependent terms of equations 2.2 introduce a tendency to maximize the coherence of all patterns. With $\lambda > 0$, this tendency is accompanied by a competition between the patterns, so that any pattern p with significant coherence C_p will try to suppress the coherence C_q of any other pattern q or will arrange for a phase difference between Ψ_p and Ψ_q so that $\widetilde{\mathcal{L}}$ is minimzed. For example, assume $P = 2$ and both patterns are nonoverlapping. The first sum in the last line of equations 2.8 will imply $C_1, C_2 \to 1$, while the second sum introduces a phase difference $(\Psi_2 - \Psi_1) \to \pi/2 \bmod \pi$ in order to minimize $\widetilde{\mathcal{L}}$. In fact, such a dynamics is exactly what we observe in example 3 in section 4.4.1.

This tendency toward coherence of a pattern and decoherence between nonoverlapping patterns bears a resemblance to the results obtained for the Wang-Terman model (Terman & Wang, 1995). The Wang-Terman model, however, is based on coupled relaxation oscillators and patterns were not stored via Hebbian connections. Instead, a specific two-dimensional

background geometry was assumed for the purpose of image segmentation. Using the two-dimensional background geometry makes the Wang-Terman model rather comparable to lattice models such as the ones studied in Sakagushi, Shinomoto, and Kuramoto (1987, 1988). The latter, however, were not applied to image segmentation, do not use inhibition, and do not have on and off states. Moreover, differences between the synchronization properties of relaxation and nonrelaxation oscillators have been described in Somers and Kopell (1993, 1995) and Izhikevitch (2000).

2.3 Phase Dynamics with Vanishing Amplitudes. The coupling in equation 2.2b has a remarkable effect on the synchronization properties (see also Hoppensteadt & Izhikevitch, 1997). It contains the terms

$$\text{GGL: } \frac{1}{V_k} \sum_{l=1}^{N} g_{kl} \sin(\theta_l - \theta_k)\ V_l \tag{2.9}$$

on the right-hand side of equation 2.2b. Due to the factor $1/V_k$, the units get more susceptible to phase differences as $V_k \to 0$. Moreover, a state V_k approaching the off-state is synchronized more strongly with on-states V_l, not with other off-states (we assume $g_{kl} > 0$). Thus, there is a strong tendendency toward global synchrony (infinitely strong for off-states, since $1/V_k \to \infty$, as $V_k \to 0$), where on- and off-states all acquire the same phase.

In section 3.2 we will assume $V_k = g(u_k)$ in order to compare equations 1.1 with the GGL model. The tendency toward global coherence, resulting from equation 2.9, is then not compatible with our interpretation of the V_k (see section 1). We intend to identify on and off states of units, respectively with high and low frequencies $d\theta_k/dt$. Therefore, the phase of a unit that is close to an off-state should not be strongly driven toward synchronization with on-state phases. When returning to equations 1a and 1b in the next section, we will therefore find that our model differs from the associative memory GGL model not only by introducing the frequency-driving activity $d\theta_k/dt = (2\pi/\tau_\theta)\, g(u_k) + \ldots$, but also by accompanying this change with an alternative phase coupling that allows off-state units to have phases that differ from on-state phases. This feature is necessary for using phase differences to avoid the superposition problem.

2.4 Weak Couplings and Temporal Coding. The central topic of this article is the superposition problem and its possible solution based on grouping of neural oscillators due to temporal properties. Therefore, it should be compared to an earlier approach that will now be reviewed in the context of weakly coupled systems.

For the case of weak couplings, the amplitude dynamics in equations 2.2, $\tilde{I}_k > 0$, $k = 1, \ldots, N$, may be adiabatically eliminated (see Kuramoto, 1984). This will leave the phase dynamics

$$\frac{d\theta_k}{dt} = \omega_k + \frac{\beta}{N} \sum_{l=1}^{N} \tilde{g}_{kl} \sin(\theta_l - \theta_k), \tag{2.10}$$

where $0 < \beta \ll 1$ and $\tilde{g}_{kl}$ is obtained from the adiabatic values $V_k \to V_k^\infty$, $V_l \to V_l^\infty$, leading to $\tilde{g}_{kl} = (V_l^\infty / V_k^\infty) g_{kl}$. Weak couplings will also imply $V_k^\infty > 0$, so that $\tilde{g}_{kl} < \infty$. In contrast, oscillator death with $V_k^\infty = 0$ (as well as self-ignition, $\tilde{I}_k < 0$) may occur for stronger couplings (see Aronson, Ermentrout, & Kopell, 1990).

In equation 2.10, we also reestablished nonidentical frequencies ω_k. In equation 2.2b, these were set to Ω and eliminated by going to the comoving frame, $\theta_k \to \theta_k + \Omega t$. Nonidentical frequencies have been used to propose neural grouping based on different frequency (see Hoppensteadt & Izhikevitch, 1997, sec. 5.4.2). This frequency modulation (FM) proposal uses a number of basic frequencies Ω_Γ, $\Gamma = 1, \ldots, G$, and assumes that

$$\omega_k \in \{\Omega_0, \Omega_1, \ldots, \Omega_G\}, \tag{2.11}$$

with $G \leq N$. Two units k and l will then belong to the same group (ensemble, pool), if their natural frequencies will be the same, $\Omega_{\Gamma(k)} = \Omega_{\Gamma(l)}$, up to terms of order β. This approach gets particularly interesting when higher-order phase couplings are included in equation 2.10. The equality condition is then generalized to a resonance condition (Hoppensteadt & Izhikevitch, 1997, sec. 9.4.3). In section 3.4, we relate the FM proposal to our approach.

3 Phase Extension of Classical Networks

In this section we return to the model of equations 1.1. The system is compared with the GGL model, pattern recognition capabilitities are discussed, and temporal coding is compared with the FM proposal mentioned at the end of the previous section.

3.1 The Model. The system of equations 1.1 reads

$$\tau_u \frac{du_k}{dt} = -u_k + \frac{1}{N} \sum_{l=1}^{N} w_{kl}(\theta)\, g(u_l) + I_k, \tag{1.1a}$$

$$\tau_\theta \frac{d\theta_k}{dt} = 2\pi\ g(u_k)\ (1 + \mathcal{S}_k(u, \theta)), \tag{1.1b}$$

where

$$w_{kl}(\theta) = g_{kl}\,(\alpha + \beta \cos(\theta_k - \theta_l)), \tag{2.2}$$

$$\mathcal{S}_k(u, \theta) = \frac{\beta}{N} \sum_{l=1}^{N} g_{kl}\, \sin(\theta_l - \theta_k)\, g(u_l). \tag{2.3}$$

The weights w_{kl} agree with the $\widetilde{w}_{kl}$ of the GGL model, except for the parameter α that was introduced to give the weights also a classical part. The term *classical* refers to phase independence (see the discussion in von der Malsburg, 1999). This difference was motivated by introducing in equations 1.1 a phase coupling that does not replace but complements classical network dynamics. In fact, for the examples of section 4, we always use $\alpha > \beta \geq 0$. The $\mathcal{S}_k(u, \theta)$ are obtained from $\widetilde{\mathcal{S}}_k(V, \theta)$ by replacing $V_l \to g(u_l)$.

With $\beta > 0$, the $\sin(\theta_l - \theta_k)$ in $\mathcal{S}_k$ introduce a tendency to synchronize (desynchronize) the units k and l, assuming $g_{kl} > 0$ ($g_{kl} < 0$), and the $\cos(\theta_k - \theta_l)$ in w_{kl} strengthens (weakens) the couplings if the units are in-phase (out-of-phase). These features agree with that of equations 2.2. Notice, however, the different strengths of phase couplings, given by $g(u_k)$ in equations 1.1 and by $1/V_k$ in equations 2.2, which lead to the different behavior of units that are close to off-states. We will return to a comparison between equations 1.1 and equations 2.2 in section 3.2.

With $\beta = 0$, the weights are constant, $w_{kl}(\theta) = \alpha g_{kl}$, and equation 1.1a decouples from the phase dynamics, thereby reducing to a classical model, while the phase dynamics reduces to uncoupled oscillations, $d\theta_k/dt = (2\pi/\tau_\theta)\, g(u_k)$. In this sense, the complete dynamics of equations 1.1a and 1.1b is an extension of the classical dynamics.

3.2 Gradient Dynamics, Frequency-Driving Activity, and Comparison with GGL Models. In the following, we identify $V_k \equiv |z_k| = g(u_k)$. Interpreting the activity $g(u_l)$ as an amplitude allows a comparison of equations 1.1 with 2.2. Notice that we do not restrict the V_k to be small; the range of the V_k is $0 < V_k < 1$, due to the definition of g. As $|u_k| \to \infty$, the V_k run into saturation.

The comparison of equations 1.1 with 2.2 will get most transparent by relating equations 1.1 to gradient dynamics, in analogy to section 2.1. Using the weights $w_{kl}(\theta)$ of equation 1.2, a natural generalization of the Cohen-Grossberg-Hopfield function (Cohen & Grossberg, 1983; Hopfield, 1984) for amplitudes and phases reads

$$\mathcal{L}(V, \theta) = -\frac{1}{2N} \sum_{k,l} w_{kl}(\theta)\, V_k V_l + \mathcal{P}(V), \tag{3.1}$$

where the potential is given by

$$\mathcal{P}(V) = \sum_k \int^{V_k} \left(g^{-1}(x) - I_k\right) dx \tag{3.2}$$

$$= \sum_k \left\{ \frac{1}{2} \left(V_k \ln V_k + (1 - V_k) \ln (1 - V_k)\right) - I_k V_k \right\}. \tag{3.3}$$

For deriving equation 3.3, the u_k may be obtained from V_k through

$$u_k = g^{-1}(V_k) = \frac{1}{2} \ln \frac{V_k}{1 - V_k}. \tag{3.4}$$

We applied the integration formula $\int \ln x \, dx = x \ln x - x$ and set the integration constant in equation 3.3 to zero.

We may now relate equations 1.1a and 1.1b to $\mathcal{L}$, just as equations 2.2 were related to $\widetilde{\mathcal{L}}$. Using $dg(u)/du = 2V_k(1 - V_k)$, equations 1.1 may be written as:

$$\tau_u \frac{dV_k}{dt} = 2V_k(1 - V_k) \left\{ -\frac{1}{2} \ln \frac{V_k}{1 - V_k} + \frac{1}{N} \sum_{l=1}^{N} w_{kl}(\theta) V_l + I_k \right\} \tag{3.5a}$$

$$= 2V_k(1 - V_k) \left(-\frac{\partial}{\partial V_k} \right) \mathcal{L}(V, \theta), \tag{3.5b}$$

$$\tau_\theta \frac{d\theta_k}{dt} = 2\pi \, V_k \left(1 + \widetilde{\mathcal{S}}_k(V, \theta)\right) \tag{3.5c}$$

$$= 2\pi V_k + 2\pi \left(-\frac{\partial}{\partial \theta_k} \right) \mathcal{L}(V, \theta). \tag{3.5d}$$

The system splits into a gradient part and the frequency-driving activity term $2\pi V_k$. For the remainder of this section, we compare the system of equations 3.5 with the GGL model of equations 2.2 and 2.5.

The amplitude dynamics of equation 3.5a is similar to equation 2.2a. The ln term in equation 3.5a is a higher-order analog of the $\tilde{I}_k V_k - V_k^3$ terms in equation 2.2a. The additional factors $V_k(1 - V_k)$ correspond to the saturation as $V_k \to 0, 1$.

The main difference is in the phase dynamics. Notice first that the GGL model could be formulated as the pure gradient system of equations 2.5a and 2.5b only by assuming vanishing shear, that is, $\eta_k = 0$ in equation 2.1. Therefore, the term $2\pi V_k$ in equation 3.5d may be compared to the shear term of the GGL model that is usually set to zero when pattern recognition is implemented (see Hoppensteadt & Izhikevitch, 1997, sec. 10.4). In

contrast, we consider the nonvanishing shear-like term to be essential for our interpretation of the oscillatory units. Setting the $2\pi\, g(u_k) + \ldots$ term on the right-hand side of equation 1.1b to zero would destroy the relation between the activity u_k and the frequency $d\theta_k/dt$ of the signals. This realizes the requirement that on and off states should correspond, respectively, to high and low frequencies of the phases.

Moreover, in order to establish a consistent picture, the effect of the $\mathcal{S}_k(u, \theta)$ should not be growing as $V_k \to 0$. In section 2.3, we mentioned that the coupling $\widetilde{\mathcal{S}}_k / V_k$ in equation 2.2b leads to global synchrony, establishing a strong tendency toward equal phases of on- and off-states. Such a coupling would be in conflict with low frequencies as $V_k \to 0$. Therefore, an obvious difference between equations 2.2 and equations 3.5 is that the latter uses the coupling $V_k \mathcal{S}_k$ instead. Thereby, synchronization between units k and l, assuming $g_{kl} > 0$, is enforced only for on-states, that is, the unit k gets less susceptible to phase differences as $V_k \to 0$. In terms of the gradient system, the difference arises due to the factor $1/V_k^2$ in equation 2.5b that is absent in equation 3.5d. Therefore, with regard to the V_k, the coupling in equations 35a to 3.5d is of higher order than the couplings of the GGL model.

3.3 Pattern Recognition Capabilities. Patterns stored according to equation 2.6 are related to minima of $\mathcal{L}$. Correspondingly, one could expect that gradient dynamics in equations 3.5a to 3.5d tends to retrieve stored patterns by approaching the minima of $\mathcal{L}$. However, due to the term $2\pi V_k$ in equation 3.5d, the system of equations 1.1 is not a pure gradient system. We find

$$\tau_u \frac{d}{dt}\mathcal{L}(V, \theta) = -\sum_k \left(2V_k(1 - V_k) \left(\frac{\partial \mathcal{L}}{\partial V_k} \right)^2 + \frac{\tau_u}{\tau_\theta} \frac{\partial \mathcal{L}}{\partial \theta_k} \left(\frac{\partial \mathcal{L}}{\partial \theta_k} - V_k \right) \right). \tag{3.6}$$

In contrast to GGL models with vanishing shear, the right-hand side includes terms that may be nonnegative. These may imply an increase in the $\mathcal{L}(V, \theta)$ values. Let us write the corresponding terms in the form

$$\sum_k \frac{\partial \mathcal{L}}{\partial \theta_k} V_k = \frac{\beta}{N} \sum_{k,l} \frac{V_k - V_l}{2} \sin(\theta_l - \theta_k)\, V_k g_{kl} V_l. \tag{3.7}$$

In the following, we will discuss the effect of these terms on the pattern recognition capability of the system.

The terms of equation 3.7 result from the first term $2\pi V_k + \ldots$ on the right-hand side of equation 1.1b. Why could this term eventually cause an increase of $\mathcal{L}(V, \theta)$? Say, $\theta_k = \theta_l + \delta\theta \pmod{2\pi}$ with $0 < \delta\theta < \pi$. On one hand, the sign of $\sin(\theta_l - \theta_k)$ is such that the θ_l tend to be fastened and the θ_k tend to be slowed down in order to reach synchrony (assuming $g_{kl} > 0$ and $\alpha > \beta \geq 0$). With $V_k - V_l < 0$, this dynamics is supported by the frequency-driving activity terms. The corresponding terms in equation 3.7 will be negative and will imply a tendency of $\mathcal{L}(V, \theta)$ to decrease. On the other hand, for $V_k - V_l > 0$, the activity term may outvote the $\sin(\theta_l - \theta_k)$ tendency: θ_k could remain faster than θ_l. This corresponds to the positive terms of equation 3.7, which may cause $\mathcal{L}(V, \theta)$ to increase. As a result, however, θ_k will approach θ_l from "the other side," that is, the situation $\theta_k = \theta_l - \delta\theta \pmod{2\pi}$ is reached. Should $V_k - V_l > 0$ still be the case, the terms of equation 3.7 will then become negative and will imply a tendency of $\mathcal{L}(V, \theta)$ to decrease again. Obviously, without additional inquiry, we may not judge whether the foregoing increase will be cancelled by the following decrease of $\mathcal{L}(V, \theta)$. Therefore, we now look more closely at the V dynamics.

Assuming an adiabatic approximation, $\tau_u \ll \tau_\theta$, we find that the only term that may become negative inside the bracket of equation 3.6 is suppressed by (τ_u/τ_θ). Moreover, due to their (fast) dynamics, the V_k will have approached their on or off states (then $\partial\mathcal{L}/\partial V_k \to 0$) when the terms of order τ_u/τ_θ get relevant. Assuming sufficiently large values of α in equation 1.2, we may arrange the set of indices k so that $V_k = 1 - \epsilon_k$, $k = 1, \ldots, M$, and $V_k = \epsilon_k$, $k = M+1, \ldots, N$, where $0 < \epsilon_k \ll 1$, for some $M \leq N$. Then

$$\sum_k \frac{\partial \mathcal{L}}{\partial \theta_k} V_k \to \frac{\beta}{N} \sum_{k,l=1}^{M} \sin(\theta_k - \theta_l)\, g_{kl} + O(\epsilon) = O(\epsilon), \tag{3.8}$$

due to the antisymmetry of $\sin(\theta_k - \theta_l)\, g_{kl}$. Combining these aspects, we find that frequency-driving activity terms only imply a $(\tau_u/\tau_\theta)\, O(\epsilon)$ correction to $d\mathcal{L}/dt$, and we may expect that these terms will not significantly affect the pattern recognition capabilities. This expectation will be confirmed by the examples in section 4.

3.4 Comparison with the Frequency Modulation Approach to Temporal coding. Before giving examples for the pattern recognition behavior, we want to relate equations 1 to the frequency modulation (FM) approach described in section 2.4. Remember the weak coupling limit in equation 2.10:

$$\frac{d\theta_k}{dt} = \underbrace{\Omega_{\Gamma(k)}} + \frac{\beta}{N} \sum_{l=1}^{N} \underbrace{\widetilde{g}_{kl}}\, \sin(\theta_l - \theta_k)\,. \tag{3.9}$$

Equation 1.1b: $\dfrac{2\pi}{\tau_\theta} g(u_k) \quad g(u_k) g_{kl} g(u_l)$

Here, we have added a comparison with equation 1.1b. We may relate the FM proposal to equations 1a and 1b by assuming two basic frequencies: $\Omega_1 = 0$ (off-states) and $\Omega_2 = 2\pi/\tau_\theta$ (on-states). The difference is that for the FM approach, the Ω_Γ have been external parameters, while in equations 1a and 1b, they are subject to the dynamics of equation 1.1a. Whether the system approaches Ω_1 or Ω_2 depends on external inputs and initial values. Notice also that in the context of the complete dynamics of equations 1.1a and 1.1b, the couplings are also subject to a dynamics. Due to the term $g(u_k)g_{kl}g(u_l)$, a synchronization is not enforced if one or both of the units k, l approach an off-state.

In the context of brain dynamics, it has been speculated that the external character of natural frequencies would actually turn dynamical in a more complete setting: "It is reasonable to speculate that the brain has mechanisms to regulate the natural frequencies Ω_Γ of its neurons so that the neurons can be entrained into different pools at different times simply by adjusting the Ω_Γ" (Hoppensteadt & Izhikevitch, 1997, sec. 5.4.2). In section 1, we mentioned that here we do not aim at a biological interpretation. Nevertheless, it is obvious that equations 1.1 provide an extension of equation 3.9, where the analog of the natural frequencies may result from a dynamical process.

Notice, however, a difference between the FM approach and the direction we take. The FM approach uses only frequency gaps for separating the neural pools, labeled by $\Omega_1, \ldots, \Omega_G$. The analogy with equations 1.1 goes only to separating on-states (Ω_1) and off-states (Ω_2). Among the on-states, we continue with a separation based on phase gaps. In particular, for overlapping patterns, competitive couplings are used to separate their phases. Using competition as the separating mechanism has been proposed already in von der Malsburg (1981).

With regard to biological interpretation, it has been suggested that temporal coding based on frequency gaps versus temporal coding based on phase gaps may be complementary (Hoppensteadt & Izhikevitch, 1999), the former being valid in the weak-coupling regime, while the latter may be more suiteable for strong couplings (Izhikevitch, 1999). Notice in this respect that the model of equations 1.1 is indeed applicable with strong couplings, in particular due to the saturation properties of the activation function.

4 Examples

In this section, we illustrate the dynamics of equations 1.1 in the context of simple examples. We continue to use the short form $V_k = g(u_k)$. In section 4.1, we comment on input choices. In section 4.2, we specify the parameters and present the patterns that we use for the examples. In section 4.3, we consider the weights without competition, that is, $\lambda = 0$ in equation 2.6. We apply external input to the network and demonstrate how

coherent pattern recognition retrieves the stored patterns. As expected, the frequency-driving activity $2\pi V_k$ separates the phases of active units from the background. We compare coherent pattern recognition to the classical dynamics. This will help in understanding the limitations of the classical approach and valuing the additional features that arise due to the phase couplings, most notably the avoidance of the superposition problem. In section 4.4, we include competition between the patterns, that is, $\lambda > 0$, and study the resulting effects.

4.1 Pattern Retrieval and Input Choices. Choosing the activation function g of equations 1.1 so that on and off states of a unit k correspond, respectively, to the limits 1 and 0 of V_k introduces a subtlety regarding the inputs I. This may be understood when writing equations 1.1 in terms of $h(x) = \tanh(x) = 2g(x) - 1$ instead. For this discussion, it is sufficient to consider $\beta = 0$, then

$$\tau_u \frac{du_k}{dt} = -u_k + \frac{\alpha}{2N} \sum_{l=1}^{N} g_{kl}\, h(u_l) + J_k\,, \tag{4.1}$$

where J is related to I through

$$I_k = J_k - \frac{\alpha}{2N} \sum_{l=1}^{N} g_{kl}\,. \tag{4.2}$$

The subtlety is related to the interpretation of vanishing inputs. Vanishing inputs should be identified not with vanishing I but with $J = 0$. The related I value is then given by the second term in equation 4.2. Only for such a value of I will $u = 0$ imply $du/dt = 0$ so that vanishing inputs correspond to a neutral or undecided state at the origin of u-space. In this section, we specify the inputs in terms of J.

4.2 A Simple Network. We now want to illustrate the dynamics of equations 1.1 by discussing simple examples. We choose $N = 6$, $P = 3$, and store the patterns

$$\xi^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \xi^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \xi^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \tag{4.3}$$

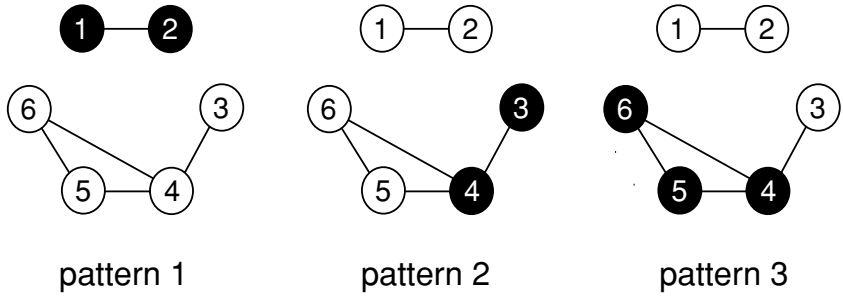

Figure 1: The stored patterns according to equation 4.3, drawn together with the resulting Hebbian connections h_{kl} of equation 4.4 and numbering of the units. The filled and empty circles correspond, respectively, to on and off states.

with $N_1 = N_2 = 2$, $N_3 = 3$. This form allows us to distinguish two cases: overlapping and nonoverlapping patterns. Pattern 1 does not overlap with patterns 2 and 3, while patterns 2 and 3 overlap at unit 4. The patterns are illustrated in Figure 1. The Hebbian weights h and the inhibitory competition terms c are given by equation 2.6:

$$h_{kl} = \frac{1}{P}\begin{pmatrix} 1\,1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0 \\ 0\,0\,1\,2\,1\,1 \\ 0\,0\,0\,1\,1\,1 \\ 0\,0\,0\,1\,1\,1 \end{pmatrix}, c_{kl} = \frac{-1}{P^2}\begin{pmatrix} 0\,0\,1\,2\,1\,1 \\ 0\,0\,1\,2\,1\,1 \\ 1\,1\,0\,1\,1\,1 \\ 2\,2\,1\,2\,1\,1 \\ 1\,1\,1\,1\,0\,0 \\ 1\,1\,1\,1\,0\,0 \end{pmatrix}. \tag{4.4}$$

We set $d = 0$ in equation 2.6 since we are not interested in background effects.

In the following, the global Hebbian parameter α is set to $\alpha = 2PN$. This is large enough to establish attractors of the system that realize storage of the patterns. On-states correspond to $V_k \to 1$, while off-states correspond to $V_k \to 0$. The timescales are set to $\tau_u = \tau_\theta/4 = \tau$. This is a mild form of the $\tau_u \ll \tau_\theta$ scenario, underlying the adiabatic approximation in section 3.3. Simulations are performed using Euler's method with time step $dt = 0.02\tau$. The initial values, random values for the θ and small random values for u_k (so that $V_k \approx 1/2$), are the same for all examples in this section. The examples are distinguished by different values for β, λ, and inputs I_k (expressed in terms of J_k; see section 4.1).

4.3 Coherent Pattern Recognition. We begin with two examples for the response to external inputs. In section 4.1 we argued that vanishing inputs

should be identified with $J_k = 0$, where J is related to I by equation 4.2. Therefore, excitatory input should be identified with $J_k > 0$ and inhibitory input with $J_k < 0$.

Both examples compare coherent pattern recognition with the classical case. The classical case corresponds to $\beta = 0$. For the phase coding scenario, we choose $\beta = 0.1\alpha$. In this section, we do not include competition—$\lambda = 0$.

4.3.1 Example 1. First, we set $J_1 = -E$, $J_6 = +E$, where $E = 10$, while all other inputs vanish. The resulting dynamics is displayed in Figure 2. We find that the inhibitory input at unit 1 suppresses pattern 1, while the excitatory input at unit 6 excites pattern 3. Moreover, due to its overlap with pattern 3, pattern 2 is also excited. This is true for the classical and the phase coding cases. It is the behavior that should be expected. If phase coding is included, $\beta = 0.1\alpha$, we find that the active units synchronize after a few multiples of the timescale.

The example confirms that we succeeded in constructing a pattern recognition mechanism that retrieves patterns in a coherent mode. The frequency-driving activity term in equation 1.1b is essential in separating the active from nonactive units. The freqencies of active units approach $2\pi/\tau_\theta$, while the frequencies of nonactive units are frozen toward zero. The implications for the superposition problem are illustrated with the next example.

4.3.2 Example 2. The superposition problem may now be demonstrated by choosing both inputs to be positive, $J_1 = J_6 = +E$. Then pattern 1 is also excited (see Figure 3). In the classical case, $\beta = 0$, the superposition of active patterns no longer allows us to distinguish the single patterns. For example, it is no longer possible to distinguish pattern 1 since all units are active. Consequently, there is a problem for information processing along the succeeding stages. This is the superposition catastrophe that plagues the classical approach (Rosenblatt, 1961; von der Malsburg, 1981).

Comparing the classical situation to the case $\beta = 0.1\alpha$ makes it obvious why the inclusion of phase coding helps to avoid the superposition problem. Now the units carry not only information about being on or off, related to high or low frequencies, but also information about the phases. We find that units 1 and 2 of pattern 1 synchronize separately from units 3 to 6 of patterns 2 and 3. As a result, in the output, pattern 1 may be distinguished from the rest due to its different phase. Whenever the succeeding stage of information processing is sensitive to this phase difference, it will be able to identify pattern 1 despite the fact that the other units are also on.

4.4 Competition Between Patterns. While the superposition problem is a severe drawback in the application of classical networks, another problem is the mixed states that correspond to a common activation of overlapping patterns. This arises when using the Hebbian weights without additional coupling of the units. It is only a problem, of course, if we aim at getting a

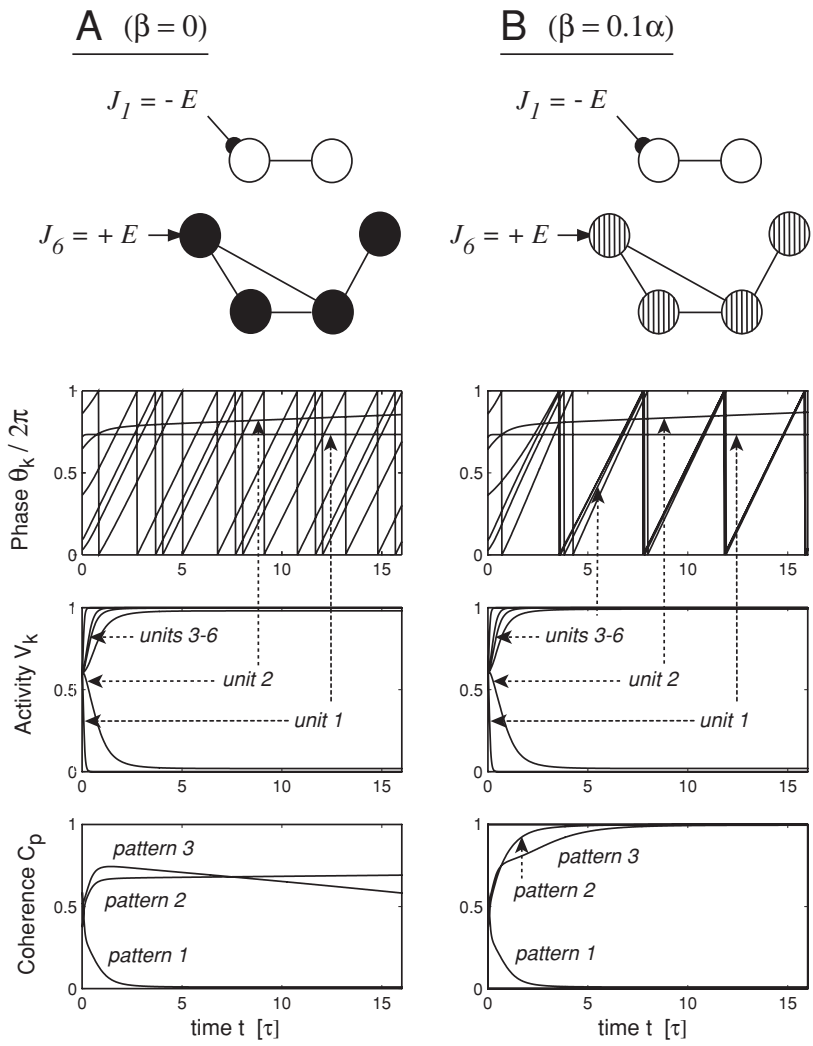

Figure 2: Example 1. The lines connecting the units indicate the Hebbian connections h_{kl} of equation 4.4 (see Figure 1). The filled and empty circles correspond, respectively, to on and off states. We compare the classical case ($\beta = 0$) with phase coding ($\beta > 0$). (A) $\beta = 0$, the classical case. The inhibitory input J_1 suppresses pattern 1. The excitatory input J_6 activates patterns 3 and 2 through unit 4. (B) $\beta = 0.1\alpha$, phase coding. Coherent activity is indicated by circles with parallel stripes. The connected active units synchronize within a few multiples of the timescale. The corresponding patterns reach maximal coherence. (Whenever mod 2π is applied to θ, the values are connected for the sake of better visualization.)

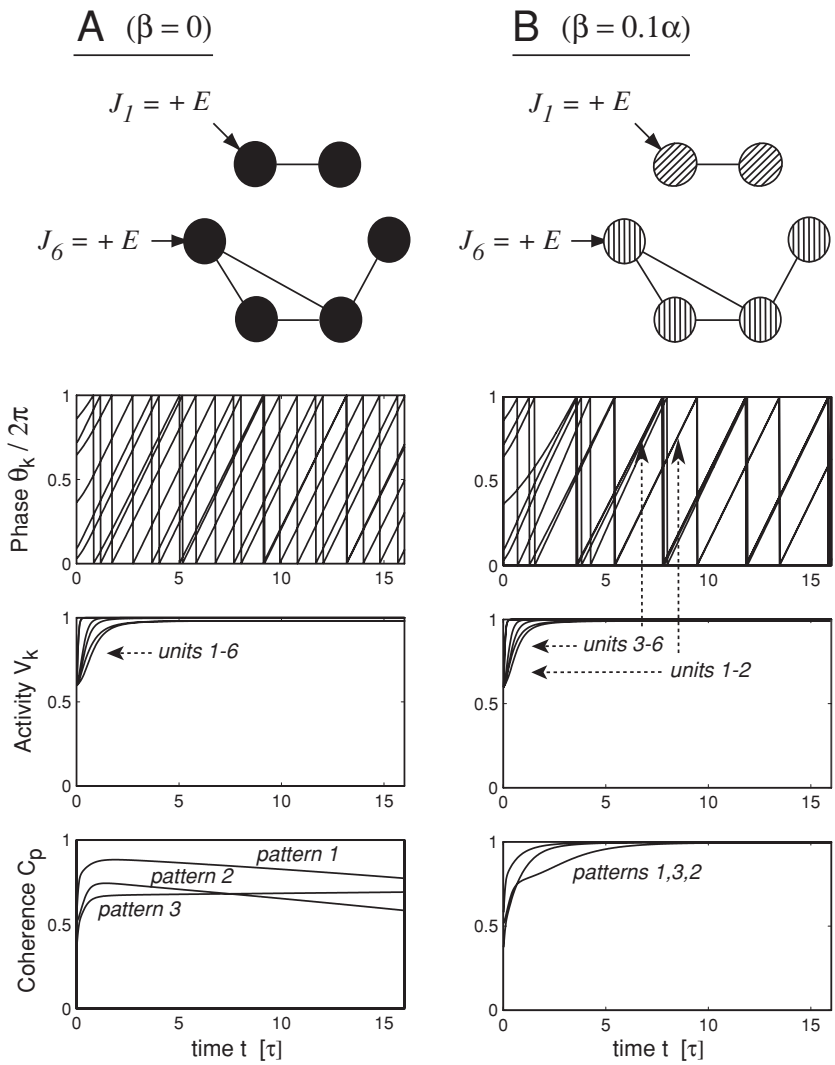

Figure 3: Example 2. Both external inputs are now excitatory. As a consequence, all units are activated. Again, we compare the classical case ($\beta = 0$) with that of phase coding ($\beta > 0$). (A) $\beta = 0$, the classical case. The superposition problem arises. The activation carries no information about single patterns. (B) $\beta = 0.1\alpha$, phase coding. Now the units of the two connected components synchronize separately. Different orientations of the stripes correspond to different phases. The superposition problem is avoided. Pattern 1 is distinguishable from the rest. Separating patterns 2 and 3 also requires competition (see example 3).

unique pattern as the output of the network. Then the situation of examples 1 and 2 is not satisfying since patterns 2 and 3 are activated simultaneously due to being connected through Hebbian connections. This is why we now include additional inhibitory weights c_{kl} that introduce competition between the patterns—that is, we use nonvanishing λ in equation 2.6.

4.4.1 Example 3. We set $\lambda = 1.2P$. No external inputs are applied. To make the influence of phase coding obvious, we now use a larger value for β by choosing $\beta = 0.8\alpha$ for the nonclassical case. The resulting dynamics is displayed in Figure 4. It may be understood as follows.

Let us first discuss the classical situation. When analyzing our random initial values for u_k, they are found to be slightly positive, so that the V_k are slightly above $1/2$ and the activities tend to run toward the on-states. This, however, invokes the competition process, and for the classical scenario, $\beta = 0$, the units of pattern 3 succeed in suppressing all the other activities, so that finally units 4 to 6 of pattern 3 approach the on-state attractor, while the other units move toward the off-state attractor. The flat lines in the phase diagram correspond to switched-off units 1 to 3. The dominance of pattern 3 is mainly due to $N_3 > N_{2,1}$.

With phase coding, the situation is different and allows an interesting observation. We find that although again pattern 3 dominates pattern 2, we now end up with a still active pattern 1. The reason for this may be understood when realizing that the two winning patterns are out of phase. In fact, the phase difference approaches π. Thus, pattern 1 escaped the competition by desynchronizing with pattern 3 so that the coupling $w(\theta)$ and thus the competition is weakened. As a result, the system ends up in a situation similar to Figure 3, but now the superposition problem between the overlapping patterns has been solved due to the competition: unit 3 and thereby pattern 2 is suppressed. Two patterns may therefore be simultaneously active and still be separable. This is another example of avoiding the superposition problem.

Notice that an analogous behavior occurs when global inhibition is used—that is, competition between units instead of competition between patterns. A possible example is obtained from $\lambda = 0$ and $d_{kl} = -1$ for all k, l in equation 2.6. Again, in the classical case, pattern 3 is winning, while in the presence of phase coding, patterns 1 and 3 are active with different phases, as illustrated for the case of Figure 4.

5 Summary and Outlook

In this letter, we used the oscillatory network model of equations 1.1 to study a solution to the superposition catastrophe in the context of phase models. The system of equations 1.1 was obtained by extending classical neural network models to include phase dynamics. It was designed to meet two basic requirements: on and off states should correspond, respectively,

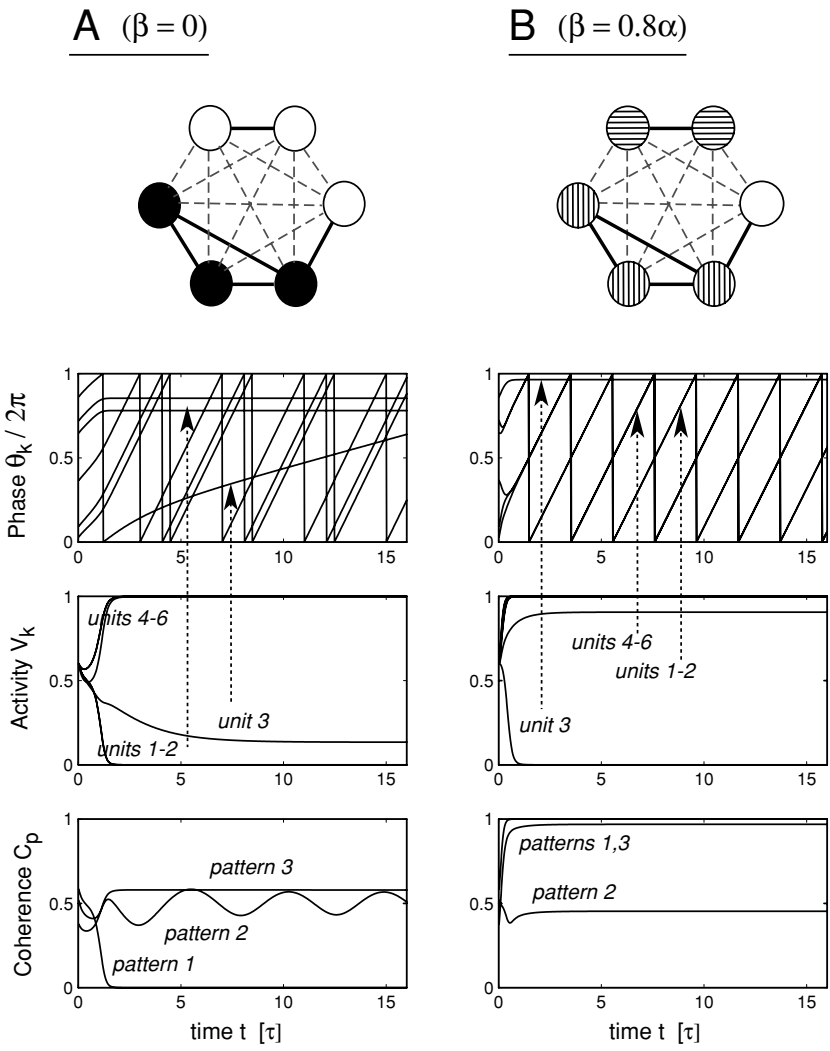

Figure 4: Example 3. The network now also includes inhibitory weights λc_{kl}, $\lambda = 1.2P$. These introduce competition between patterns. The network is now completely connected, as illustrated by the additional broken lines. No external input is applied. (A) $\beta = 0$, the classical case. Pattern 3 is the winning pattern (it is larger than the others). (B) $\beta = 0.8\alpha$, phase coding. Again pattern 3 wins over pattern 2. Pattern 1, however, is now able to survive. Phase coding allows pattern 1 to escape the competition by desynchronizing with pattern 3. This is illustrated by circles with stripes of orthogonal orientation.

to high and low frequencies of the phases, and patterns should be retrieved in a coherent (synchronized) mode.

Identifying the activity $g(u_k)$ of equations 1.1 with an oscillation amplitude V_k allows us to compare the model with equations 2.2, a generalized version of the discrete complex Ginzburg-Landau (GGL) model with vanishing shear and identical natural frequencies. This model describes Stuart-Landau oscillators (i.e., oscillators close to a Hopf-bifurcation) coupled to lowest order, a model that is frequently used for implementing associative memory (see Hoppensteadt & Izhikevitch, 2003). With $V_k = g(u_k)$, the GGL model does not obey the first of the two requirements. That phase dynamics leads to different results as oscillatory units k approach their off-state, $V_k \to 0$. We also compared our approach with an alternative proposal for grouping oscillatory units based on temporal properties (see Hoppensteadt & Izhikevitch, 1997, sec. 5.4.2). This frequency modulation (FM) proposal is founded on choosing different natural frequencies for oscillatory units, resulting in the grouping of units with nearly identical natural frequencies. With regard to a separation of on- and off-states, equations 1.1 may be seen as a dynamical extension of this approach. However, with regard to a grouping among on-states, the FM approach would be based on frequency gaps, while the model of equations 1.1 uses phase gaps. Correspondingly, we mentioned that (regarding biological systems) a complementarity has been suggested (Izhikevitch, 1999), according to which temporal coding with frequency gaps arises for weakly coupled systems, while temporal coding based on phase gaps may be more suitable for strong couplings.

In the context of equations 1.1, the solution of the superposition problem is straightforward. It is in accordance with original proposals of temporal coding (von der Malsburg, 1981, 1985; see also von der Malsburg, 2003). The frequency-driving acitivity term $2\pi g(u_k)$ in equation 1.1b separates the active units from the background so that active units, $V_k \to 1$, approach the frequency $(2\pi/\tau_\theta)(1 + \mathcal{S}_k)$, while nonactive units, $V_k \to 0$, are frozen to zero frequency. Among the active units, the phases provide an additional labeling of the activities. In case of coherent pattern recognition, due to synchronizing phase couplings, each component may be identified with a single phase. Different patterns may then have the same frequency and may still be distinguished by the phase differences between the patterns. In order to avoid mixed states of overlapping patterns, a competition between the patterns may be introduced. This was achieved by introducing appropriate inhibitory weights in addition to the excitatory Hebbian weights. The coherence of the patterns may then be reduced to the winning subset.

In this letter, we did not specify mechanisms by which successive stages of information processing can read out the phases of different components. Future work should approach this issue so that a more complete picture arises. Moreover, taking the limit $N \to \infty$ is essential for obtaining phase

transitions between ordered and disordered states (see Kuramoto, 1984; Strogatz, 2000). Studying the proposed and related models in this limit should be of particular interest. Higher-order phase couplings, and possibly related phenomena such as clustering, should also be of interest (see Tass, 1999). In the context of the FM proposal, these led to resonance conditions for the natural frequencies.

An early version of the Wang-Terman model was studied for patterns that were overlapping at one unit, resembling our example in section 4 (Wang et al., 1990). Presentation of overlapping patterns as input led to common activation of these patterns: the units at the overlap participated in each of the overlapping patterns (see Wang et al., Figure 3). This differs from our example 3, where competition led to a winner among the overlapping patterns. If a common activation of overlapping patterns should be desirable, different approaches might work for the phase models. For example, the phase model could be extended to the coupling of higher-order modes. Overlapping units could then participate in a higher-frequency mode that may simultanously synchronize with the lower-frequency nonoverlapping parts of the patterns via resonances. This would still allow the patterns to desynchronize in the nonoverlapping parts. Modifying the phase model to establish such a feature, however, is beyond the scope of this article.

We presented simple examples to illuminate the dynamical content of equations 1.1. Evidently, to gain importance as a pattern recognition method, it should be demonstrated that the method stands the test of the more advanced applications. In case of locally coupled relaxation oscillators, this step was taken by Wang and Terman, who studied segmentation of real images. They observed that a straightforward application of relaxation oscillator networks led to an unsuitable segmentation of the real image, leading to many tiny fragments, a problem they called fragmentation. This problem was solved by introducing a lateral potential for each oscillator (Wang & Terman, 1997; see also Shareef et al., 1999). It will be of interest to study how the phase model should be applied to real images. In case fragmentation arises, it would be particularly interesting to study whether some hierarchical processing could help to combine the fragments, possibly in analogy to hierarchical processing supported by higher-level cortical areas. Given the relative simplicity of the phase models in comparison with coupled relaxation oscillator models, one may hope that necessary steps to include high-level processing could be more easily analyzed and implemented.

Acknowledgment

It is a pleasure to thank Christoph von der Malsburg for inspiring and helpful discussions.

References

Abbott, L. F. (1990). A network of oscillators, *J. Phys. A: Math. Gen.*, *23*, 3835–3859.

Aoyagi, T. (1995). Networks of neural oscillators for retrieving phase information, *Physical Review Letters*, *74*(20), 4075–4078.

Aronson, D., Ermentrout, G., & Kopell, N. (1990). Amplitude response of coupled oscillators. *Physica D, 41*, 403–449.

Baird, B. (1986). Nonlinear dynamics of pattern formation and pattern recognition in the rabbit olfactory bulb. *Physica D*, *22*, 242–252.

Baird, P., & Meir, R. (1990). Computing with arrays of coupled oscillators: An application to preattentive texture discrimination. *Neural Computation*, *2*, 458–471.

Chakravarthy, S. V., & Gosh, J. (1994). A Neural Network–based associative memory for storing complex-valued patterns. In *Proc. IEEE Int. Conf. Syst. Man Cybern.* (pp. 2213–2218). Piscataway, NJ: IEEE.

Chakravarthy, S. V., & Gosh, J. (1996). A complex-valued associative memory for storing patterns as oscillatory states. *Biological Cybernetics*, *75*, 229–238.

Cohen, M. A., & Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transaction on Systems, Man, and Cybernetics, SMC-13*, 815–826.

Freeman, W. J., Yao, Y., & Burke, B. (1988). Central pattern generating and recognizing in olfactory bulb: A correlation learning rule. *Neural Networks*, *1*, 277–288.

Hirose, A. (Ed.). (2003). *Complex-valued neural networks.* Singapore: World Scientific.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences of the U.S.A.*, *81*, 3088–3092.

Hoppensteadt, F. C., & Izhikevitch, M. E., (1996). Synaptic organization and dynamical properties of weakly connected neural oscillators: II. Learning phase information. *Biological Cybernetics*, *75*, 129–135.

Hoppensteadt, F. C., & Izhikevitch, E. M., (1997). *Weakly connected neural networks*. Berlin: Springer-Verlag.

Hoppensteadt, F. C., & Izhikevitch, E. M. (1999). Thalamo-Cortical interactions modeled by weakly connected oscillators: Could the brain use FM radio principles? *BioSystems*, *48*, 85–94.

Hoppensteadt, F. C., & Izhikevitch, E. M. (2000). Pattern recognition via synchronization in phase-locked loop neural networks. *IEEE Transactions on Neural Networks*, *11*, 734–738.

Hoppensteadt, F. C., & Izhikevitch, E. M. (2003). Canonical neural models, In M. Arbib (Ed.), *Brain theory and neural networks* (2nd ed., pp. 181–186). Cambridge, MA: MIT Press.

Izhikevitch, E. M. (1999). Weakly connected quasi-periodic oscillators, FM interactions, and multiplexing in the brain. *BioSystems*, *48*, 85–94.

Izhikevitch, E. M. (2000). Phase equations for relaxation oscillators. *SIAM J. Appl. Math.*, *60*, 1789–1805.

Kuramoto, Y. (1975). Self-entrainement of a population of coupled non-linear oscillators. In H. Araki (ed.), *International Symposium on Mathematical Problems in Theoretical Physics* (pp. 420–422). Berlin: Springer-Verlag.

Kuramoto, Y. (1984). *Chemical oscillations, waves, and turbulence*. Berlin: Springer-Verlag.

Kuramoto, Y., Aoyagi, T., Nishikawa, I., Chawanya, T., & Okuda, K. (1992). Neural network model carrying phase information with application to collective dynamics. *Progr. Theor. Phys.*, *87*, 1119–1126.

Kuzmina, M., Manykin, E., & Surina, I. (1995). Associative memory oscillatory networks with Hebbian and pseudo-inverse matrices of connections. In *Proceedings of the Third European Congress on Intelligent Techniques and Soft Computing (EUFIT'95), Aachen, Germany, August 28–31* (pp. 392–395). Bellingham, WA: SPIE International Society for Optical Engineering.

Kuzmina, M., & Surina, I. (1994). Macrodynamical approach for oscillatory networks. In *Proceedings of SPIE: Optical Neural Networks* (Vol. 2430, pp. 229–235). Aachen: ELITE Foundation.

Li, Z., & Hopfield, J. J. (1989). Modeling the olfactory bulb and its neural oscillatory processings. *Biological Cybernetics*, *61*, 379–392.

Müller, H. J., Elliott, M. A., Herrmann, C. S., & Mecklinger, A. (Eds.). (2001). *Visual Cognition*, 8 [Special issue].

Noest, A. J. (1988a). Associative memory in sparse phasor neural networks. *Europhysics Letters*, *6*, 469–474.

Noest, A. J. (1988b). Discrete-state phasor neural networks. *Physical Review A*, *38*, 2196–2199.

Rosenblatt, F. (1961). *Principles of neurodynamics: Perception and the theory of brain mechanism*. Washington, DC: Spartan Books.

Roskies, A. L. (Ed.). (1999). *Neuron, 24* [Special topic].

Sakagushi, H., Shinomoto, S., & Kuramoto, Y. (1987). Local and global self-entrainments in oscillator lattices. *Progr. Theor. Phys.*, *77*, 1005–1010.

Sakagushi, H., Shinomoto, S., & Kuramoto, Y. (1988). Mutual entrainment in oscillator lattices with nonvariational type interaction. *Progr. Theor. Phys.*, *79*, 1069–1079.

Shareef, N., Wang, D., & Yagel, R. (1999). Segmentation of medical images using LEGION. *IEEE Transactions on Medical Imaging*, *18*, 74–91.

Somers, D., & Kopell, N. (1993). Rapid synchronization through fast threshold modulation. *Biological Cybernetics*, *68*, 393–407.

Somers, D., & Kopell, N. (1995). Waves and synchrony in networks of oscillators of relaxation and non-relaxation type. *Physica D*, *88*, 1–14.

Sompolinsky, H., Golomb, D., & Kleinfeld, D. (1990). Global processing of visual stimuli in a neural network of coupled oscillators. *Proc. Natl. Acad. Sci. USA*, *87*, 7200–7204.

Sompolinsky, H., Golomb, D., & Kleinfeld, D. (1991). Cooperative dynamics in visual processing. *Physical Review A, 43*, 6990–7011.

Sompolinsky, H., & Tsodyks, M. (1992). Processing of sensory information by a network of coupled oscillators. *International Journal of Neural Systems*, *3* (Suppl.), 51–56.

Strogatz, S. H. (2000). From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Physica D*, *143*, 1–20.

Takeda, M., & Kishigami, T. (1992). Complex neural fields with a Hopfield-like energy function and an analogy to optical fields generated in phase-conjugate resonators. *J. Opt. Soc. Am. A*, *9*, 2182–2192.

Tass, P. A. (1993). *Synchronisierte Oszillationen im visuellen Cortex—ein synergetisches Modell*. Unpublished doctoral dissertation, Institut für Theoretische Physik und Synergetik der Universitt Stuttgart.

Tass, P. A. (1999). *Phase resetting in medicine and biology*. Berlin: Springer-Verlag.

Tass, P., & Haken, H. (1996a). Synchronization in networks of limit cycle oscillators. *Z. Phys. B*, *100*, 303–320.

Tass, P., & Haken, H. (1996b). Synchronization oscillations in the visual cortex—a synergetic model. *Biological Cybernetics*, *74*, 31–39.

Terman, D., & Wang, D. (1995). Global competition and local cooperation in a network of neural oscillators. *Physica D*, *81*, 148–176.

von der Malsburg, C. (1981). *The correlation theory of brain function* (Internal Rep. 81-2). Max-Planck Institute for Biophysical Chemistry.

von der Malsburg, C. (1985). Nervous structures with dynamical links. *Ber. Bunsenges. Phys. Chem., 89*, 703–710.

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, *24*, 95–104.

von der Malsburg, C. (2003). Dynamic link architecture. In M. Arbib (Ed.), *Brain theory and neural networks* (2nd ed., pp. 365–368). Cambridge, MA: MIT Press.

von der Malsburg, C., & Buhmann, J. (1992). Sensory segmentation with coupled neural oscillators. *Biological Cybernetics*, *67*, 233–242.

von der Malsburg, C., & Schneider, W. (1986). A neural cocktail-party processor. *Biological Cybernetics*, *54*, 29–40.

Wang, D., Buhmann, J., & von der Malsburg, C. (1990). Pattern segmentation in associative memory. *Neural Computation*, *2*, 94–106.

Wang, D., & Terman, D. (1995). Locally excitatory globally inhibitory oscillator networks. *IEEE Transaction on Neural Networks*, *6*, 283-286.

Wang, D., & Terman, D. (1997). Image segmentation based on oscillatory correlation. *Neural Computation*, *9*, 805–836.

Winfree, A. T. (2001). *The geometry of biological time* (2nd ed.). Berlin: Springer-Verlag.

Received November 10, 2004; accepted June 30, 2005.