Biol. Cybern. 80, 393–409 (1999)

Biological Cybernetics

Neural 'bubble' dynamics in two dimensions: foundations

J.G. Taylor

Institut für Medizin, Forschungszentrum-Juelich, Germany and Centre for Neural Networks, King's College, London, UK

Received: 26 March 1997 / Accepted in revised form: 16 December 1998

Abstract. An extension to two dimensions of recent results in continuum neural field theory (CNFT) in one dimension is presented here. Focus is placed on the treatment of receptive fields and of learning on afferent synapses to obtain topographic maps.

1 Introduction

There was considerable progress in the 1970s and 1980s in the development of a model of neural dynamics in terms of a one-dimensional continuum neural field theory (CNFT) [1–5]. This was based on a considerable body of earlier and related work in the late 1950s and 1960s in which the nature of activity of sheets of neurons was investigated both experimentally (in cortical slabs) [6], by simulation and through related mathematical analysis [7, 8]. For example, there is the powerful analysis of bifurcation of spatially periodic solutions from zero neural activity in the seminal work of [9]; this has led to interesting speculations about the manner in which instability in the cortical sheet can be the origin of hallucinations under drugs [10]. At the same time the existence of 'bubbles of neural activity', which are compact regions of autonomous neural firing, were shown to exist in a number of these earlier works [1–3].

The results from this program of work are of great interest in showing the possibility of the creation of autonomous activity in the cortex and of the formation of topographic maps due to suitable inputs. However, the mathematical analysis was only performed for the case of one-dimensional neural fields. More recently interest in this model, now considered as a two-dimensional neural field theory, has been resurgent, and in particular, it has been applied to modelling saccadic control by the superior colliculus [11], the formation of orientation detectors in the early visual cortex [12, 13], the representation of spatial orientation [14], and changes in somatosensory cortical receptive fields with training [15]. However, these studies and others about the formation of various forms of autonomous activity [10, 16] have involved simulation and not mathematical analysis to allow a possibly deeper understanding of the principles involved.

It is the purpose of this paper to develop an understanding of the mathematical nature of the activation and learning processes of two- dimensional CNFT. This will then allow for application to various problems of sensory processing, culminating in associative cortical regions, as well as provide a mathematical framework for the increasingly numerous applications referred to earlier. The problems of the transformation of visual inputs into various codes of shape, colour, motion, and texture are then possible to consider mathematically in the terms of two-dimensional CNFT applied to modules suitably adapted to these codes. Some examples of the manner in which this approach sheds light on various types of effects, in particular observed in stabilised images and apparent motion, will be discussed in the second part of this paper. This first part is devoted to setting out the mathematical framework and proving some of the basic theorems both extending the one-dimensional results of earlier work [1–9] and providing new results on some aspects of the temporal dynamics of bubbles.

There are two parts to the development of a two-dimensional CNFT. One involves solely the dynamics of the neurons, assuming that the synaptic weights are fixed. The other is concerned with learning on the afferent inputs, and is based on the earlier analysis of the on-going neural activity to allow the learning to take place in terms of the correct dynamical responses of the neurons to their total inputs, both afferent and lateral.

The analysis of the dynamics of the neurons has been performed mainly for time-independent activity. This corresponds to assuming that only activity in an asymptotic state needs to be considered. However, there are considerable data on the temporal development of activity in various cortical areas, especially from single cell recordings [17]; there is also increasing data arising from non-invasive instruments, such as magnetoencephalography (MEG) and electroencephalograph (EEG)

Correspondence to: J.G. Taylor

recordings, as well as from invasive cortical analysis as obtained by the use of optical dyes [18]. Thus, it is necessary to consider non-stationary dynamical aspects of CNFT in order to relate to such experimental details. That will be discussed in Sects. 3 and 4, after an initial construction of the framework of two-dimensional CNFT and a consideration of autonomous activity. In section 3 the temporal features of CNFT will be analysed in order to understand the time course of activity in the neural sheet M. It will be related to some experimental data on neural activity traces observed in the cortex [19, 20].

It is crucial to understand learning dynamics in order to build up the codes used for processing early sensory data so that they may then be efficiently processed at a later stage. This aspect of CNFT will be discussed in Sect. 4, where the extension of the notions of receptive field and region of influence will be given for the two-dimensional case. This will be used to develop a theory of learning, culminating in discussion of topographic maps for a certain class of inputs. The manner in which the receptive field structure changes with input class is considered in a further section.

This paper continues, in its second part, with applications of these ideas to a set of visual phenomena, and ends with a conclusion.

Before launching into the detailed mathematics of CNFT in two dimensions, I would like to state the more general fundamental assumptions of the theory. In particular:

(a) The activity of the neurons is approximated by their mean firing rates; the neurons do not spike. This may remove very important temporal information contained in the timing of the emission and receipt of spikes that occurs in real neurons, so that results, such as those in [21], will not be available. Thus, coherence of neuronal activity cannot be discussed in such a framework. The analysis extending the present discussion to that case will be presented elsewhere, using the results of the present analysis regarded as a temporally averaged limit for single neuron activity.
There are several justifications for this simplification:
- the mathematical problems met in the analysis of two- dimensional CNFT are non-trivial;
- interesting results with relevance to living neural systems have already been shown to arise from it, as the earlier references indicate [10–15]. Thus, certain of the underlying principles of information processing in the cortex can be gleaned by the use of mean firing rate neurons in continuum neural models;
- the majority of the cells in the cortex fire repetitive spike trains [22], and for these standard arguments allow the response to be reduced to that for which the mean firing rate is a sigmoidal function of the summed activity arriving at the cell at a given time;
- there is a considerable body of ongoing work on mean-firing rate neurons (reviewed, for example, in [23]) which indicates the usefulness of this approach in developing understanding of the nature of pattern formation in excitable media.

(b) The neural system is approximated by a continuous neural sheet. There is considerable support for this approximation, although it is one which will have to be extended by consideration of intermixed populations (to take account of different sorts of cell types) and of boundary conditions (for finiteness of the sheet), as well as to look at how well large but discrete populations of cells behave in the limit as their number tend to infinity.

(c) The refractory period and time delays arising from finite transmission time are neglected. However, these are effects that can be included in a later, more extensive analysis.

In any case the mathematical problems are of sufficient difficulty that they need to be resolved without the added complications of spiking, discreteness and time delays noted above.

2 Two-dimensional continuum neural field theory

The basic equations for the CNFT are well known but will be repeated here for completeness and to set up our conventions. The case of two dimensions will be considered in detail, but much of the analysis will be extendable to higher dimensions. It was not felt appropriate to dwell on higher dimensions than two since that may be hoped to give a good approximation to a given cortical layer, such as layer 2/3 or layer 4. In any case many of the results obtained here can be extended to the higher-dimensional case.

2.1 The two-dimensional equation

The neuronal positions will be labelled by the vector $\mathbf{x}$ as a two-component quantity which will be assumed to be taken from the two-dimensional manifold M. We can think of M as R^2. It would be possible to consider more detailed two-dimensional manifolds, such as the two-dimensional torus or sphere (possibly with handles), but questions of compactness or homology will not be at the centre of attention here, unless otherwise explicitly stated.

The membrane potential of a neuron at the point $\mathbf{x}$ and time t will be denoted by $u(\mathbf{x}, t)$. It will be assumed that there is lateral connectivity on M defined by the lateral connection weight function $w(\mathbf{x} - \mathbf{x}')$ between the two neurons at the relevant points $\mathbf{x}$ and $\mathbf{x}'$. The connection weight will usually be of Mexican hat form as a function of the Euclidean distance $|\mathbf{x} - \mathbf{x}'|$. There is also an afferent connection weight function $s(\mathbf{x}, \mathbf{y})$ from the thalamic position $\mathbf{y}$ to the cortical point $\mathbf{x}$. The response function of a neuron will be taken to be determined by its mean firing rate, which is given as some function f of the membrane potential u of the relevant cell.

The membrane potential $u(\mathbf{x}, t)$ will satisfy the CNFT equation [1–3]

$$\tau \partial u(\mathbf{x}, t)/\partial t = -u(\mathbf{x}, t) + \int d\mathbf{x}'\, w(\mathbf{x} - \mathbf{x}') f[u(\mathbf{x}', t)] + \int d\mathbf{y}\, s(\mathbf{x}, \mathbf{y}) I(\mathbf{y}, t) + h \quad (1)$$

where $I(\mathbf{y}, t)$ is the input to the thalamic position $\mathbf{y}$ at time t, h is the neuron threshold, and the integration over the lateral connection weight is over the manifold M of neurons.

2.2 The one-dimensional case

There are well-known autonomous solutions to (1) in the case when M is one-dimensional [3]. It is the purpose of the rest of this section to extend the results of that paper to the two-dimensional case. Let us first restate some of the results obtained for the one-dimensional case in [3] (where similar results were also obtained in [1] and [2]).

In that case (1), for a static solution and with no input, becomes:

$$u(x) = \int w(x - x')\, f[u(x')]\, dx' + h \quad (2)$$

It is simplest to consider the case of a sharp threshold response function $f = \theta$ (the step function) so that (2) becomes

$$u(x) = \int w(x - x')\, \theta[u(x')]\, dx' + h \quad (3)$$

A 'bubble' is defined to have a positive membrane potential over an interval, independent of input. Let us consider the bubble from $x = 0$ to $x = a$:

$$u(x) > 0, \quad 0 < x < a; \quad u(0) = u(a) = 0 \quad (4)$$

and otherwise $u < 0$. Then from (3), u is obtained explicitly as

$$u(x) = \int_0^a w(x - x')\, dx' + h = W(x) - W(x - a) \quad (5)$$

where the function W is defined by

$$W(x) = \int_0^x w(x')\, dx' \quad (6)$$

Necessary conditions for the bubble to exist are that the membrane potential vanishes at the ends of the interval $[0, a]$, so

$$u(0) = u(a) = 0 = W(a) + h \quad (7)$$

It is then possible to show that

$$u(x) > 0 \text{ for } 0 < x < a \text{ if } h < 0; \quad u(x) < 0 \text{ otherwise.}$$

Stability of the resulting solution then requires [3]

$$dW(a)/da < 0, \quad \text{or} \quad w(a) < 0 \quad (8)$$

Thus, the one-dimensional bubble exists under the conditions (7) and (8).

There are a number of further important results derived in [3] concerning the nature of bubble solutions and their extension to input dependence which will be briefly summarised here:

(a) The parameter ranges for h and for the parameters in W can be determined so as to allow for autonomous solutions of various types ($\oslash$ or the trivial one, ∞ or the constant non-zero one, an a-solution as the bubble of finite length described above, and a spatially periodic solution);
(b) Complete determination of those patterns which are stable and those which are unstable, from amongst the stationary solutions described above;
(c) Response to input stimulus patterns: a bubble of finite length moves to a position of maximum input;
(d) Two bubbles interact, with attraction if close (from the Mexican hat connection weight function), with repulsion if more distant, and with no effect on each other if very distant;
(e) Spatially homogeneous temporal oscillations can occur (between a layer of excitatory and one of inhibitory cells);
(f) Travelling waves can persist.

2.3 The two-dimensional extension

It is now necessary to determine which of these features extend to the two-dimensional situation. In that case the lateral connection strength between two neurons is assumed to be of Mexican hat form in the Cartesian distance between the two. An example of this is the difference of two Gaussians (DOG):

$$w(\mathbf{x}, \mathbf{x}') = w(|\mathbf{x} - \mathbf{x}'|) = A \exp[-d^2/2a^2] - B \exp[-d^2/2b^2] \quad (9)$$

where $d = d(\mathbf{x}, \mathbf{x}') = |\mathbf{x} - \mathbf{x}'|$ is the Euclidean distance between $\mathbf{x}$ and $\mathbf{x}'$ and A, B, a, and b are all taken as positive constants independent of the position on the two-dimensional surface M. To have a Mexican hat form for the lateral connections, it is required that $A > B$, $a < b$. More generally, the function w will be taken solely as a function of $|\mathbf{x} - \mathbf{x}'|$, with a Mexican hat shape in its single variable. The general conditions for central excitation and lateral fall-off connection strengths defined by the lateral connection strength function w is that:

$$w(0) > 0,\ w'(0) < 0 \quad (10)$$

If there is a bubble of finite extent, then by symmetry it must be circular. Let it have radius R; it will be called an $R[2]$-solution ($R[n]$-solution will be the corresponding spherical one in n-dimensional space). The origin will be taken to be the centre of the disc D: $\{\mathbf{x}: |\mathbf{x}| < R\}$. Then the condition on the stationary membrane potential $u(\mathbf{x})$ for the existence of an $R[2]$-solution equal to the disc D is, from (1),

$$u(\mathbf{x}) = \int_D w(\mathbf{x} - \mathbf{x}')\, d\mathbf{x}' + h \tag{11}$$

with the necessary condition:

$$\begin{aligned} u(\mathbf{x}) = 0,\ \mathbf{x} \in \partial D\,(\text{the boundary of } D),\\ u(\mathbf{x}) > 0 \,\text{inside}\, D,\\ u(\mathbf{x}) < 0 \,\text{outside}\, D \end{aligned} \tag{12}$$

It is possible to evaluate certain properties of the integral

$$W(|\mathbf{x}|, R) = \int_D w(x - x')\, d\mathrm{x}' \tag{13}$$

By manipulation of (13) it can be shown that

$$W'(0) = 0, \quad W''(0) = \pi a w'(a) < 0, \quad W'(\infty) = 0 \tag{14}$$

where the dashed suffix on W' denotes differentiation with respect to its first variable $|\mathbf{x}|$. Moreover, it is clear from the fact that the integration region of W, in terms of its variable $|\mathbf{x} - \mathbf{x}'|$, progressively covers the region of negative values of w as $|\mathbf{x}|$ increases, that the condition

$$W(|\mathbf{x}|, R) + h = 0 \tag{15}$$

only has one solution for $|\mathbf{x}|$, say the value R', provided that $h < 0$. Moreover, $u(\mathbf{x})$ defined by (11) will be positive for $|\mathbf{x}| < R'$ and negative outside that region. This detailed form of argument is an attempt to give the simplest extension of that of Amari [3]. However, it fails to show that R and R' are identical; if they are not, then the bubble size assumed, that is the value R, is different than that ensuing from the above analysis of (11), that being the value R'. This contradicts the basic assumption of solution (11).

It is necessary to proceed slightly differently, by considering (15) only to be valid for the vector $\mathbf{x}$ on the boundary of D. In that case the function to be considered has the form $W(|\mathbf{x}|, R)$ only for such restricted positions of $\mathbf{x}$. Let us denote this restricted function as $G(x)$, with

$$G(R) = W(R, R) = \int_D w(\mathbf{x} - \mathbf{x}')\, d\mathbf{x}' \tag{16}$$

where now $\mathbf{x}$ is only allowed to be on the boundary of D, which has radius $|\mathbf{x}| = R$. When w is chosen only to be a function of the squared length of its vector variable, it is possible to evaluate the derivatives of G with respect to R and to prove (see Appendix A) that

$$\begin{aligned} &dG(0)/dR|_{R=0} = 0\\ &d^2G/dR^2|_{R=0} = 2\pi w(0) > 0 \end{aligned} \tag{17}$$

on further use of (10).

It is now possible to apply the methods of [3] to deduce the same results as in the one-dimensional case for some of the questions (a) to (f) raised in Sect. 2.2. Before discussing the full two-dimensional extensions of those results, we should note that there are semi-infinite solutions obtained by neglecting the dependence of $u(\mathbf{x})$ in (1) on one or other of its variables. Thus, if there is independence of u on its second variable, the lateral interaction becomes 'marginalised' to reduce to a function of the remaining variable, and so is replaced by

$$v(x) = \int w(x, y)\, dy \tag{18}$$

where $\mathbf{x} = (x, y)$. All of the results of [3] apply to this reduced one-dimensional case, with conditions now being on the marginalised lateral interaction v of (18). We will not consider these cases any further here.

Returning to the full two-dimensional situation, we define

$$W_\infty = \lim_{R\to\infty} W(x, R) = \lim_{R\to\infty} G(R) \tag{19}$$

where the middle term in (19) is independent of the value x in the limit. In terms of the above limit, it is possible to show that:

(a) **Theorem 1.** In the absence of input:

1. There exists a $\oslash$ solution (which is zero everywhere) iff $h < 0$.
2. There exists a ∞-solution (which extends to infinity in both directions) iff $W_\infty > -h$.
3. There exists a R[2]-solution (of finite size) iff $h < 0$ and $R > 0$ satisfies

$$G(R) + h = 0 \tag{20}$$

The proof of theorem 1 is given in Appendix B.

It is possible to extend the classification of the solutions for varying levels of the stimulus h. Let $G_\mathrm{m} = \max_x G(x)$.

Theorem 2. The nature of the various solutions for different parameter ranges is as in Fig. 1.

The proof of theorem 2 is given in Appendix B.

(b) To determine which of these solutions is stable, it is necessary to extend the one-dimensional discussion of [3] to two (or higher) dimensions. From (12) the boundary of D, defined by the radius $R(t)$ at time t, satisfies the constraint

$$u(R(t), t) = 0 \tag{21}$$

On differentiation of (21) with respect to t and use of (1), we obtain

$$dR/dt = -[G(R(t)) + h]/\tau v \tag{22}$$

where v is the gradient of u normal to ∂D and is negative. Then the equilibrium case of (20) results on setting the right-hand side of (22) to zero. The stability of this solution is determined, following the argument in [3], by the sign of dG/dR:

$$dG/dR < 0 \Leftrightarrow \text{stability} \tag{23}$$

This leads to the stability classification of the solutions as given in theorem 2.

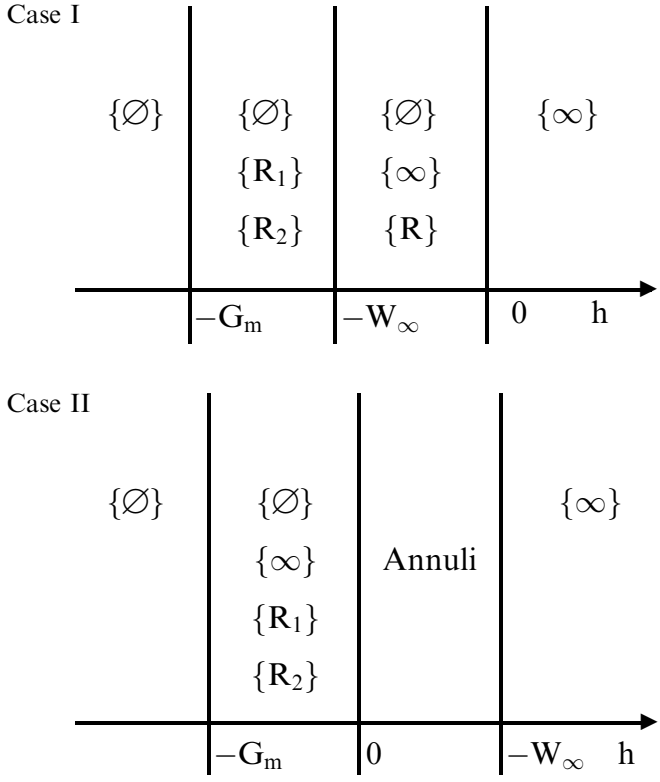

Fig. 1. The variation of the range of the stationary 'bubble' solutions to the two-dimensional continuum neural field theory (CNFT) equations as the parameters h (running along the x-axis) of the threshold of the neurons varies. The $\oslash$ solution denotes the trivial one which is equal to h (so negative) everywhere; ∞ denotes the solution which is a positive constant everywhere, while $\{R\}$ denotes a non-trivial circularly symmetric solution whose region of positivity is of finite extent R. Cases I and II correspond to $W_\infty > 0$ and $W_\infty < 0$ respectively.

(c) The response to stationary inputs of an R[2]-solution can also to treated as in [3]. Consider a small stationary input $\varepsilon I(\mathbf{x})$, which is not assumed to be circularly symmetric so that the asymptotic bubble will not be circularly symmetric either. The equation of constraint is, following (22),

$$d\mathbf{x}/dt \cdot \nabla \mathbf{u} + \partial u/\partial t = 0 \quad \text{on} \quad \partial D(t) \tag{24}$$

Replacing the time derivatives of u on the left-hand side of (24) by (1), it is now possible to derive the condition, for $\mathbf{x}(t)$ on the boundary of $D(t)$,

$$d\mathbf{x}/dt \cdot \nabla \mathbf{u} = (1/\tau)[h + \varepsilon I + G(|\mathbf{x}(t)|)] \tag{25}$$

On expanding in a perturbation series in the small quantity ε, with

$$\mathbf{x}(t) = \mathbf{x}(0) + \mu(t) \tag{26}$$

where the length of $\mu(t)$ is of order ε, the constraint arises

$$d\varepsilon/dt \cdot \nabla \mathbf{u} = (1/\tau)[\varepsilon I + \nabla G(|\mathbf{x}(t)|)\,\mu] \tag{27}$$

where the derivatives are evaluated at $\mu = 0$, so $\mathbf{x}(t) = \mathbf{x}(0)$.

The result from the constraint (27) is that the net radial movement of the boundary of $D(t)$ is towards the region of largest input. There will be a movement of regions of $\partial D(t)$ towards lower values of the input, if these are positive, but there will be a larger velocity of movement towards those regions of the boundary nearer the maxima of I. This is similar to the one-dimensional result of [3]. However, it is necessary now to deal with the whole boundary of $D(t)$, treated as a continuum and not as separate end-points. This aspect will become more important to control and quantify when we turn to learning later in the paper.

(d) The result (d) of [3] can also be extended in the same way, where the effect of one region (say $D1$) on another (say $D2$) is given, in terms of the lateral interaction term in (2), as the effective input to a neuron in $D2$ at the point $\mathbf{x}$ of amount

$$s(\mathbf{x}) = \int_{D1} w(|\mathbf{x} - \mathbf{x}'|)\, d\mathbf{x}' \tag{28}$$

This will have the same effect as in the one-dimensional case, with attraction between the bubbles at $D1$ and $D2$ if they are close enough (as determined by $s(x)$), repulsion if the two regions are further separated, and ultimately no interaction between the bubbles at all if they are beyond the range of the lateral interaction term w (if that is finite).

(e) The case of spatially homogeneous oscillations extends immediately to the two-dimensional case, since only ∞-solutions are being considered.

(f) This case involves temporal structure and will be considered more fully in the next section.

It is also of interest to consider possible time-dependent instabilites of the above two-dimensional bubble solutions with respect to time-dependent perturbations, in order to determine, for example, if a Hopf bifurcation might occur. This has been investigated in the one-dimensional case in [23] and [24], and only Hopf bifurcations to unstable finite oscillating modes discovered. The method is immediately extendable to the two-dimensional case using (25); we leave it to the reader to perform the corresponding extension in detail.

3 Temporal analysis

There are various temporal features which need to be considered as part of the analysis of bubbles, and which have not been treated previously by Amari or other authors. Amongst these are: (i) the temporal details of the formation of bubbles by external stimuli, (ii) the decay characteristics of bubbles after their initial formation, (iii) the temporal history of the interaction between two bubbles.

For the first of these, it is of interest to determine the dependence of the speed of bubble formation on the various parameters of the continuum neural field. This would allow predictions to be made which may ultimately be checked by the use of measuring devices with fast response, such as by MEG or optical dye techniques [18].

The second question above, that of the determination of the dependence of the lifetimes of bubbles on the CNFT parameters, depends on further structure being

added to the CNFT than that present so far. In particular, it is necessary to include decay terms for the membrane potentials, as would be brought about by adapting or habituating currents, so that the bubbles do not have an infinite lifetime. Such terms are necessary to be able to relate the presence of bubbles to actually observed activity in cortex; as noted explicitly in [25], there is no observed stimulus-led continued activity in the cortex (at least in the primary visual cortex). There must be a decay mechanism for bubbles once they are formed. Adaptation is a natural mechanism for this. It will be explored later in this section.

The third question is of relevance in relating bubble phenomena to observed features of visual processing, such as in the possible occurrence of bubble fusion or splitting so as to explain apparent motion (AM) [26]. Bubble formation may help clarify some of the puzzling aspects of the observed phenomena of AM.

3.1 The temporal aspects of bubble formation

The purpose of this subsection is to analyse the behaviour in time of the development of a bubble as it first forms and as it develops into its asymptotic stationary state, as brought about by an external input $S(\mathbf{x}, t)$. This will be done initially for one dimension, and also only for fixed afferent synaptic weights. The appropriate equation, from (1), is then

$$\tau \partial u(x,t)/\partial t = -u(x,t) + \int dx'\, w(x-x')\theta[u(x',t)] + S(x,t) + h \tag{29}$$

It will be assumed that a bubble of activity is created by the input $S(x,t)$, which is itself turned on a time $t = 0$. For simplicity we will take $S(x,t)$ to have the separable form

$$S(x,t) = S(x)g(t) \tag{30}$$

where $S(x)$ is a symmetric function about the origin with a peak at the origin and finite support, whilst $g(t)$ is taken to be a sigmoidal function of time which is zero before $t = 0$.

The resulting bubble will then have support $(-l(t), l(t))$, and our task is to determine the time dependence of $l(t)$. From (30) the equation satisfied by $u(t)$ is

$$u(x,t) = S(x)K(t) + \{1 - \exp(-t/\tau)\}h + \int_0^t \exp[-(t-t')/\tau]\, dt' \int_{-l(f)}^{l(t)} w(x-y)\, dy \tag{31}$$

where $K(t)$ is defined as

$$K(t) = \int_0^t \exp[-(t-t')/\tau] g(t')\, dt' \tag{32}$$

and it has been assumed that $u(0) = 0$ as the initial condition. It might be thought more natural to choose the initial condition $u(0) = -h$, so that the bubble would then take longer to grow (for $h > 0$). However, this latter choice can be modified to our boundary condition $u(0) = 0$ by shifting the time origin to the point when u vanishes (which is possible for either sign of h). Thus, we proceed with the boundary condition $u(0) = 0$.

The requirement for the bubble to exist is that

$$u(l(t), t) = 0, \quad \forall\ t > 0 \tag{33}$$

which leads to the integral equation for the function $l(t)$ as

$$0 = S(l(t))K(t) + \{1 - \exp(-t/\tau)\}h + \int_0^t \exp[-(t-t')/\tau] dt' W(2l(t')) \tag{34}$$

in terms of the integral $W(x)$ of $w(x)$ defined by (6).

In general, it is not possible to solve the integral equation (34) in explicit terms for the function $l(t)$. However, it is possible to deduce certain properties of its solution, both for short and long times.

In order to be specific about how the external input is turned on, we will first assume that

$$g(t) = [1 - \exp(-t/\alpha)] \tag{35}$$

for some positive constant α, which is the time taken to turn on the input $S(x,t)$. Then for small t,

$$K(t) = O(t^2) \tag{36}$$

Let us consider the initial time development of $l(t)$ from (31). There is a value t_0 of t at which bubble creation can commence, this value being defined as that for which the bubble has zero size at $t = t_0$, and then starts to grow. The value of t_0 is positive, as seen from (31), since initially (at $t = 0$) the first and third terms on the right-hand side of (31) are zero to $O(t)$, both being of order $O(t^2)$, so the second term, which is negative and $O(t)$, dominates. Thus, u is negative everywhere for very short time (to $O(t)$), and only begins to acquire a positive value, at $x = 0$, at some non-zero time t_0, as the contribution from the external input begins to dominate the negative value of h.

At that time the value of $l(t)$ is required to be zero, so that from (34) the condition on t_0 is

$$0 = S(0)K(t_0) + \{1 - \exp(-t_0/\tau)\}h \tag{37}$$

The first-order coefficient l_1 in the development of $l(t)$,

$$l(t) = l(t_0 + \varepsilon) = \varepsilon l_1 \tag{38}$$

can now be calculated from (34) as

$$l_1 = [K'(t_0)S(0) + \exp(-t_0/\tau)h/\tau]/[S'(0)K(t_0)] \tag{39}$$

provided $S'(0) \neq 0$. If $S'(0) = 0$, then an expansion of $l(t)$ in powers of $t - t_0$ would not appear possible.

If, alternatively, $g(t)$ is the step function at $t = 0$,

$$g(t) = \theta(t) = 1\ (t > 0), = 0 (t < 0) \tag{40}$$

then a bubble is created at $t = 0$, and the initial size $l(0)$ of the bubble must satisfy

$$S(l(0)) + h + W(2l(0)) = 0 \tag{41}$$

The bubble will then develop in time according to the solution of the equation

$$0 = [S(l(t)) + h][1 - \exp(-t/\tau)] + \int_0^t \exp[-(t-t')/\tau]\, dt' W(2l(t')) \tag{42}$$

This equation can be solved to the second order in t, for small t, in terms of the expansion

$$l(t) = l(0) + tl_1 + (1/2)t^2 l_2 \tag{43}$$

to give

$$l_1 = 0 \text{ (provided } S'(0) + \tau w(2l(0)) \neq 0)$$
$$l_2 = -3S'(l(0))/[4\tau w(2l(0))] \tag{44}$$

Higher-order terms in t extending (43), (44) can also be obtained in a similar manner. Note the quadratic temporal expansion of $l(t)$ due to the absence of the first-order term in (43), as in agreement with the results of the simulation in Fig. 2.

One may also consider the asymptotic behaviour of $l(t)$ for large t. Using that for large t,

$$\int_0^t \exp[-(t-t')/\tau]\, dt' F(t') = \tau F(t) + O(t^{-2}) \tag{45}$$

one arrives again at (41), but now with $l(0)$ replaced by the limiting value $l(\infty)$ of $l(t)$, for large t; this modified equation is also valid to $O(t^{-2})$. Hence the asymptotic expansion of $l(t)$ about $l(\infty)$ is

$$l(t) = l(\infty) + O(t^{-2}) \tag{46}$$

If the input is turned off at some finite time, then the asymptotic size of the bubble now satisfies the original equations (7) or (20).

The above discussion can also be extended to the two-dimensional case, when (31) becomes

$$u(\mathbf{x}, t) = S(\mathbf{x})K(t) + \{1 - \exp(-t/\tau)\}h + \int_0^t \exp[-(t-t')/\tau]\, dt' \int_{D(t')} w(\mathbf{x} - \mathbf{y})\, d\mathbf{y} \tag{47}$$

where $D(t)$ is the support of the bubble at time t. For a radially symmetric input, $S(\mathbf{x}) = S(|\mathbf{x}|)$, and for an $R[2]$ bubble of radius $R(t)$ at time t, (47) becomes

$$0 = S(R(t))K(t) + \{1 - \exp(-t/\tau)\}h + \int_0^t \exp[-(t-t')/\tau]\, dt' G(R(t')) \tag{48}$$

where $G(R)$ is defined in (16).

Analysis for the cases that the input is turned on in time according to (35) or (40) lead to the initial temporal development of $R(t)$ given by (39) or (43), (44), where l is replaced by R, W by G and w by G' in those formulae. The large time development of $R(t)$ is also given by (46), where $R(\infty)$ satisfies (41), now written as

$$S(R(\infty)) + h + G(R(\infty)) = 0 \tag{49}$$

3.2 The decay of bubbles

In order to consider bubble decay, an adaptation term proportional to a summation of past neuron activity, exponentially weighted, will be subtracted from the membrane activity. Such a term is to be seen as an approximation to the long-lasting after-hyperpolarisation current observed in [27]. The resulting expression replacing (1) is then

$$\tau \partial u(\mathbf{x}, t)/\partial t = -u(\mathbf{x}, t) + \int d\mathbf{x}' w(\mathbf{x} - \mathbf{x}')\theta[u(\mathbf{x}', t)] + \int d\mathbf{y}\, s(\mathbf{x}, \mathbf{y})I(\mathbf{y}, t) + h - \lambda \int \exp[-(t-t')/\tau']f[u(\mathbf{x}, t')]\, dt' \tag{50}$$

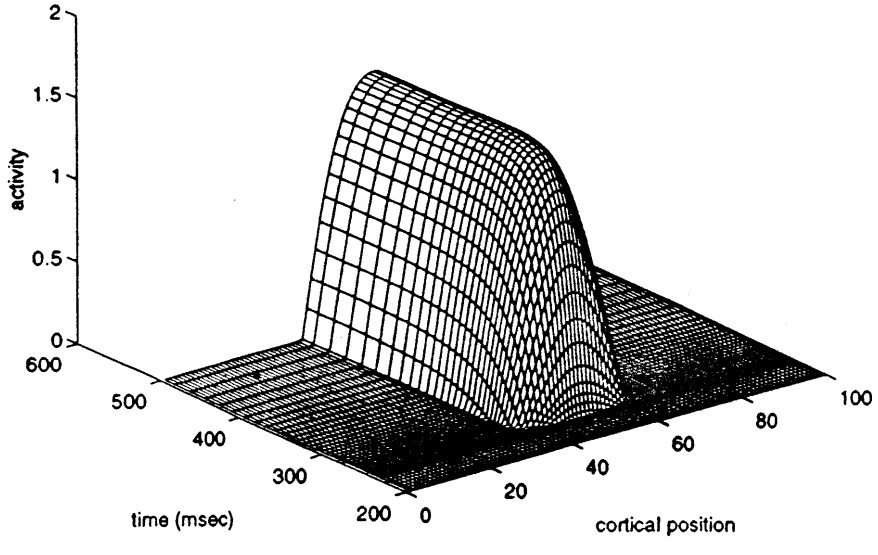

Fig. 2. The emergence of a one-dimensional bubble in time from position 50 along the line at time zero due to a non-zero input impressed on the neurons at time zero, and then being removed. *Vertical axis* denotes the output activity of the neurons at a particular place (only a measure of positive values of the membrane potential). The Mexican hat lateral connection weight function is the difference of the Gaussians of (9)

where τ' is a measure of the lifetime of the adaptation current, and λ denotes its strength; the range of integration in the last term on the right-hand side of (50) is the interval $[0, t]$. It is now necessary to calculate the lifetime of a bubble created using (50). Let us first consider the one-dimensional case; that for two dimensions will follow straightforwardly.

A particular case of interest is when a bubble has initially been created by an input which is then removed. That could be due, for example, to the neural module acting as the source of the input having a shorter lifetime for the persistence of bubbles than the one under consideration. It would also occur if the bubble is created in a primary sensory module and the input itself has been modified.

To discuss this case, it is appropriate to first reduce even further to a single recurrent neuron. For that case the membrane potential equation, from (50), is:

$$\tau \partial u(t)/\partial t = -u(t) + w\theta[ut)] + h - \lambda \int \exp[-(t-t')/\tau']\theta[u(t')]\,dt' \tag{51}$$

where a step function response has been taken for the neuron. From (51)

$$u(t) = u(0) + [1 - \exp(-t/\tau)][h + w] - \lambda \int_0^t \exp[-(t-t')/\tau]dt' \times \int_0^{t'} \exp[-(t'-t'')/\tau']\,dt'' \tag{52}$$

From (52), with $u(0) > 0$, $u(t)$ will remain positive initially in time. Moreover, (52) reduces to the expression

$$u(t) = u(0) + [1 - \exp(-t/\tau)][h + w - \lambda\tau\tau'] + \lambda\tau(\tau')^2[\exp(-t/\tau') - \exp(-t/\tau)]/(\tau' - \tau) \tag{53}$$

where the last term on the right-hand side of (53) is replaced, for $\tau = \tau'$, by the expression

$$\lambda\tau\tau' \exp(-t/\tau)$$

The last term in (53) may be neglected if $\tau' << \tau$, so that if

$$\lambda\tau\tau' > h + w + u(0) \tag{54}$$

then for suitably large t, $u(t)$ will become negative, and the firing of the neuron will then cease. If no new input arrives, then no further activity will ensue from the neuron.

The initial lifetime of the bubble is given by equating the right-hand side of (53) to zero. Using the assumption that $\tau' >> \tau$ in (53) gives the approximate value for the lifetime T as

$$T = -\tau' \ln[1 - u(0)/\tau[\lambda\tau\tau' - h - w] \tag{55}$$

where the factor $[\lambda\tau\tau' - h - w]$ is positive by (54). Equation (55) is the formula we wish to extend to the case of a one- and then a two-dimensional CNFT.

Firstly, the case of an ∞-solution in either dimension reduces to the above analysis with the constant w in the single neuron case being replaced by the quantities

$$w = \int w(x)\,dx, w = \int w(\mathbf{x})\,d\mathbf{x}$$

in the one- and two-dimensional cases, respectively.

The relevant equation in one dimension for the a-solution is (34) with the added adaptation term

$$-\lambda\tau\tau' \tag{56}$$

[dropping the term of $O(\tau')$ in (53)] and the added initial value $u(l(t), 0)$; the input term involving S also has to be dropped. The bubble will have a finite lifetime if the adaptation term is so negative that a solution exists to the resulting equation for the asymptotic size of the bubble

$$h - \lambda\tau\tau' + W(2l(\infty)) = 0 \tag{57}$$

This could arise if

$$\lambda\tau\tau' - h > W_m \tag{58}$$

as discussed in theorem 2. Thus, if (58) is true, then the bubble will have a finite lifetime given, under the same approximation as for the single neuron, by

$$T = \tau' \ln[-h/(\lambda\tau\tau' - h - W_m)] \tag{59}$$

This approximation should hold for both the one- and two-dimensional cases, when in the latter W_m is replaced by G_m. In both cases we note that as W_m or G_m is increased, say by increase of cell density, the corresponding lifetime increases.

For the other extreme $\tau' >> \tau$, then τ and τ' must be interchanged in the lifetime formulae (55) and (59).

In conclusion, for the case $\tau' >> \tau$ the bubble lifetime is effectively proportional to τ, and thus dependent on whatever mechanism produces the bubble itself. In the opposite case $\tau >> \tau'$ the bubble lifetime is proportional to τ'. The later quantity is expected to be an intrinsic characteristic of the single (pyramidal) neuron, and thus very likely constant throughout the cortex.

Experimental data on lifetimes of such trace activity in cortical sites is presently only indirect; the alteration of peak responses (say 100 ms after stimulus onset) to repetitive auditory [19] and visual [20] stimuli has been found to vary in a manner consistent with our above result on the lifetime of recurrent activity. Thus, the main results of [19] and [20] are that the peak amplitudes can be fitted as a function of the interstimulus interval (ISI) by an expression of the form

$$\text{Amplitude} \propto [1 - \exp(-t/\tau)] \tag{59a}$$

where τ is the decay time for the peak activity and t is the ISI. It is possible, from the above results, to obtain (59a), with the decay time τ being the decay of adaptation τ' in the extended CNFT equation (50) above. This is seen to arise for the approximation that the time constant of the neuron is neglected so that the inhibitory

adaptation term is most important in the dependence of the cell activity on the ISI. The resulting term, assuming the membrane potential is positive throughout the ISI, is therefore

$$\text{constant} - \lambda \exp[-(t + \text{ISI})/\tau']$$

where t is the peak time chosen, say at 100 ms. Since at ISI = 0 there will be no amplitude peak, then the constant is fixed, and the resulting expression for the amplitude at the peak time t is

$$A[1 - \exp(-\text{ISI}/\tau')] \tag{59b}$$

for a suitable constant A. It is exactly this dependence on ISI which was observed in the experiments of [19] and [20].

In the experiments of [19], a subject listened to a sequence of tones at a given frequency delivered with an ISI which varied from one series of runs (each about 50–100 repetitions, which were then averaged) to another, with ISI values of 1, 2, 4, 10 s being used. The value of the peak at about 100 ms after stimulus onset (the so-called N100 peak) in different regions of the cortex was then determined by using an equivalent current dipole fit to the N100 peak data averaged over the repetitions. In the experiment reported in [20], a checkerboard pattern was shown repetitively, for 70 ms for each exposure, to each subject with an ISI varying over different trials from 150 ms to 40 s. Again, the peak value of the response at differing cortical positions was fitted with a current dipole. In both cases the data were found to fit (59a) well, with values of time constant of the order of seconds.

A variability of the time constant of (59a) was observed experimentally in its dependence on the region of the brain being analysed [19, 20]; this is in partial support of the adaptive explanation given above of the formula (59b) and its understanding as the production of a bubble of lifetime given, in its dependence on the other parameters of the model, by (59). For, as more fully discussed in Part II, the bubble lifetime (59) will be larger the greater the cell density in the region of the creation of the bubble. Such regions of great cell density are those at the highest level of the cortical processing hierarchy, such as areas 46/9 in the prefrontal cortex, 39/40 in the parietal lobe or 37 in the temporal lobe [28, 29]. However, this interpretation is still premature, and further experiments need to be performed to explore the nature of the experimental results of [19, 20] more fully, and in particular to investigate the data at the single trial level.

4 Receptive fields in two dimensions

4.1 The learning equations

The extension to two dimensions of the one-dimensional learning equations of [4] is straightforward, and only the barest explanation will be given here, together with the equations. An important feature is the spread of inputs from presynaptic field points $\mathbf{y}$ to laterally positioned points $\mathbf{y}'$ on the same presynaptic sheet Y, with induced firing rate $a(\mathbf{y} - \mathbf{y}')$. As in Sect. 2 the afferent synaptic weight to the position $\mathbf{x}$ on the cortical field X from the presynaptic field position $\mathbf{y}$ is $s(\mathbf{x}, \mathbf{y})$. For any presynaptic field input $I(\mathbf{y}, t)$ equation (1) results, modified by the presynaptic spread of the input by the function $a(\mathbf{y} - \mathbf{y}')$:

$$\tau \partial u(\mathbf{x}, t)/\partial t = -u(\mathbf{x}, t) + \int d\mathbf{x}' w(\mathbf{x} - \mathbf{x}') f[u(\mathbf{x}', t)] + \int d\mathbf{y}\, s(\mathbf{x}, \mathbf{y}') a(\mathbf{y} - \mathbf{y}') I(\mathbf{y}', t) + h \tag{60}$$

The Hebbian learning equation for the afferent synaptic weights is the immediate two-dimensional extension of the one-dimensional case of [4]:

$$\tau \partial s(\mathbf{x}, t)/\partial t = -s(\mathbf{x}, t) + \varepsilon a(\mathbf{y} - \mathbf{y}') \theta[u(\mathbf{x}, \mathbf{y}', t)] \tag{61}$$

where $u(\mathbf{x}, \mathbf{y}, t)$ is the membrane potential at the cortical point $\mathbf{x}$ due to a presynaptic input at the point $\mathbf{y}$ and ε is a learning rate constant. As in [4] it is assumed that the time to achieve relaxation to the asymptotic solution to (60) is much less than that for solving (61), with the resulting solution $u(\mathbf{x}, \mathbf{y}, s)$ brought about by an input at the presynaptic field point $\mathbf{y}$ with input synapse s. It is therefore legitimate to insert this asymptotic solution into the second term on the right-hand side of (61). Choosing a distribution $p(\mathbf{y})$ of such presynaptic inputs allows, for suitably long τ, for averaging to be performed over $\mathbf{y}'$ in (61), to give

$$\tau \partial s(\mathbf{x}, t)/\partial t = -s(\mathbf{x}, t) + \varepsilon \int p(\mathbf{y}')\, d\mathbf{y}' a(\mathbf{y} - \mathbf{y}') \theta[u(\mathbf{x}, \mathbf{y}', s)] \tag{62}$$

As in [4] this equation may be rewritten in terms of the total stimulus $S(\mathbf{x}, \mathbf{y}, t)$ received at the point $\mathbf{x}$ due to the stimulus applied at $\mathbf{y}$:

$$S(\mathbf{x}, \mathbf{y}, t) = \int d\mathbf{y}' s(\mathbf{x}, \mathbf{y}', t) a(\mathbf{y}' - \mathbf{y}) \tag{63}$$

as

$$\tau \partial S(\mathbf{x}, \mathbf{y}, t)/\partial t = -S(\mathbf{x}, \mathbf{y}, t) + \varepsilon \int p(\mathbf{y}')\, d\mathbf{y}' k(\mathbf{y} - \mathbf{y}') \theta \times [u(\mathbf{x}, \mathbf{y}', s)] \tag{64a}$$

where

$$k(\mathbf{y} - \mathbf{y}') = \int a(\mathbf{y}'' - \mathbf{y}) a(\mathbf{y}'' - \mathbf{y}')\, d\mathbf{y}'' \tag{64b}$$

Finally, we note the equation for u:

$$u(\mathbf{x}, \mathbf{y}, s) = S(\mathbf{x}, \mathbf{y}) + \int d\mathbf{x}' w(\mathbf{x} - \mathbf{x}') \theta[u(\mathbf{x}, \mathbf{y}', s)] + h \tag{65}$$

4.2 Receptive fields in two dimensions

The receptive field (RF) R_x of a neuron at the cortical point $\mathbf{x}$ is the subset of the presynaptic field Y such that

the neuron at $\mathbf{x}$ is excited when a stimulus is applied to a point of the RF:

$$R_x(t) = [\mathbf{y}: u(\mathbf{x}, \mathbf{y}, t) > 0, \mathbf{y} \in Y] \tag{66}$$

We will consider the case when the RF is a connected region on the cortical sheet X. Moreover, the RF will be assumed to be topologically isomorphic to the disc $|\mathbf{x}| < 1$ in the plane; the boundary of the RF is thus an S^1. We can therefore write

$$\partial R_x(t) = \{\mathbf{y}(\mathbf{x}, \lambda, t)\}, \lambda \in S^1 \tag{67}$$

with the periodicity conditions

$$\mathbf{y}(\mathbf{x}, 0, t) = \mathbf{y}(\mathbf{x}, 1, t)$$

This generalises the one-dimensional interval, for which the parameter λ in (67) is reduced to the discrete set $\{0, 1\}$ and R_x reduces to an interval with end-points $y(x, \lambda, t)$, with $\lambda \in \{0, 1\}$. The higher-dimensional extension of (67) is that the boundary of R_x is an S^n, which can be parametrised by an extension of the case (67).

Effective use was made in [4] of the end points $y(x, \lambda)$ for $\lambda = 0, 1$, of the interval R_x in the one-dimensional case, which were denoted $r_1(x), r_2(x)$. The generalisation to two dimensions of this is thus to the function on $X \times S^1, r(\mathbf{x}, \lambda)$, which maps the cortical point $\mathbf{x}$ onto the $S^1 r(\mathbf{x}, \lambda)$ of ∂R_x. In addition, the distances

$$\begin{aligned} &[r_2(x) - r_1(x)] = r(x) \\ &r_1^{-1}(y) - r_2^{-1}(y) = l(y) \end{aligned} \tag{68}$$

were an important and simplifying tool in the analysis of [4]. It is now useful to develop the two-dimensional version of those concepts.

The two-dimensional analogue of $r(x)$ seems most natural to define as the diameter of R_x

$$r(\mathbf{x}) = \max \|\mathbf{r}(\mathbf{x}, \lambda) - \mathbf{r}(\mathbf{x}, \lambda')\| \tag{69}$$

where the maximum is taken over all possible choices of λ and λ'. This is to be seen as the maximum of the two-parameter family of distances

$$r(\mathbf{x}, \lambda, \lambda') = \|\mathbf{r}(\mathbf{x}, \lambda) - \mathbf{r}(\mathbf{x}, \lambda')\| \tag{70}$$

(which reduces immediately to $r(x)$ in the one-dimensional case).

For a stimulus at the presynaptic point $\mathbf{y}$, the excited region in X, or the so-called projective field of $\mathbf{y}$, is the set of points

$$L_y = \{\mathbf{x} : \mathbf{y} \in R_x\} \tag{71}$$

The boundary of this region is easily seen to be given by

$$\partial L_y = \{(\mathbf{x}, \lambda) : \mathbf{y} = \mathbf{r}(\mathbf{x}, \lambda)\} \tag{72}$$

In particular, if $\mathbf{y} = \mathbf{r}(\mathbf{x}, \lambda)$ then

$$\partial L_y = \bigcup_t \mathbf{r}_t^{-1}(\mathbf{r}(\mathbf{x}, \lambda)) \tag{72a}$$

where $\mathbf{r}_t(\mathbf{x})$ denotes $\mathbf{r}(\mathbf{x}, t)$, and is clearly an S^1, as shown in Fig. 3. There is a similar analogue to the length function $l(y)$ of (68) to the definition (69) above, as

$$l(\mathbf{y}) = \operatorname{diam} \partial L_y = \max \|\mathbf{r}(\mathbf{x}, \lambda) - \mathbf{r}(\mathbf{x}, \lambda')\| \tag{73}$$

where the maximum is taken over the pairs $(\mathbf{x}, \lambda)$ and $(\mathbf{x}, \lambda') \in \partial L_y$. A two-parameter set of translations, together with their associated lengths, can be defined as extending those in one dimension, which were

$$l_1(x) = x - r_2^{-1}(r_1(x)), \quad l_2(x) = r_1^{-1}(r_2(x)) - x$$

to the form

$$\mathbf{l}_{rs}(\mathbf{x}) = \mathbf{x} - \mathbf{r}_s^{-1}(\mathbf{r}_r(\mathbf{x})), \quad \mathbf{l}_{sr}(\mathbf{x}) = \mathbf{r}_r^{-1}(\mathbf{r}_s(\mathbf{x})) - \mathbf{x} \tag{74}$$

The geometry of this system is shown in more detail in Fig. 4.

Similar extensions can be given to higher dimensions.

4.3 Dynamical equations for receptive fields

Following [4] differentiation of (65) and use of (64) leads to the dynamical equation

$$\begin{aligned} &\tau \partial u(\mathbf{x}, \mathbf{y}, t)/\partial t - \tau(\partial)/\partial t \int d\mathbf{x}' w(\mathbf{x} - \mathbf{x}') \theta[u(\mathbf{x}', \mathbf{y}, t)] \\ &\quad = -u(\mathbf{x}, \mathbf{y}, t) + \int d\mathbf{x}' w(\mathbf{x} - \mathbf{x}') \theta[u(\mathbf{x}', \mathbf{y}, t)] \\ &\qquad + \int d\mathbf{y}' p(\mathbf{y}') k(\mathbf{y} - \mathbf{y}') \theta[u(\mathbf{x}, \mathbf{y}', t)] + h \end{aligned} \tag{75}$$

The conditions on u on the boundary of R_x are

$$u(\mathbf{x}, \mathbf{y}, t) = 0, \quad \forall\ \mathbf{y} \in \partial R_x \tag{76}$$

This equation has already been used in Sect. 2.3 to deduce the differentiated constraint (24), leading to (25). However, time in (76) now enters implicitly through the variable $\mathbf{y}$ as

$$\mathbf{y} = \mathbf{r}_s(\mathbf{x}, t) \tag{77}$$

for each $s \in [0, 1]$. Thus, we arrive at the S^1-fold infinity of equations by differentiating (76) with respect to time:

$$\partial u(\mathbf{x}, \mathbf{r}_s, t)/\partial \mathbf{y},\ \partial \mathbf{r}_s(\mathbf{x}, t)/\partial t + \partial u(\mathbf{x}, \mathbf{r}_s, t)/\partial t = 0 \tag{78}$$

By symmetry, each R_x will be a circular disc at each instant of time, so conjugate pairs s, $s' \in [0, 1]$ (with $s + s' = 1$) will exist, so that $\mathbf{r}_s(\mathbf{x}, t)$ and $\mathbf{r}_{s'}(\mathbf{x}, t)$ will be at opposite ends of the diameter of R_x and the diameters of $L(\mathbf{r}_s)$ and $L(\mathbf{r}_{s'})$ are l_s and $l_{s'}$ respectively, equal to the lengths of the vectors

$$\begin{aligned} &\mathbf{l}_s(\mathbf{x}) = \mathbf{x} - \mathbf{r}_{s'}^{-1}(\mathbf{r}_s(\mathbf{x})),\ \mathbf{l}_{s'}(\mathbf{x}) = \mathbf{r}_s^{-1}(\mathbf{r}_{s'}(\mathbf{x})) - \mathbf{x} \\ &l_s = \|\mathbf{l}_s(\mathbf{x})\|, \quad l_{s'} = \|\mathbf{l}_{s'}(\mathbf{x})\| \end{aligned} \tag{79}$$

Then the argument of [4] proceeds in this case, on use of the constant input probability distribution $p(\mathbf{y}) =$ constant. One obtains successively

$$\begin{aligned} \int d\mathbf{y}' k(\mathbf{r}_s(\mathbf{x}) - \mathbf{y}') \theta[u(\mathbf{x}, \mathbf{y}', t)] &= \int d\mathbf{y}' k(\mathbf{r}_s(\mathbf{x}) - \mathbf{y}') \\ &= K(r(t)) \end{aligned} \tag{80}$$

where the integral in the middle term in (80) is over R_x, where $r(t)$ is defined by (69), and

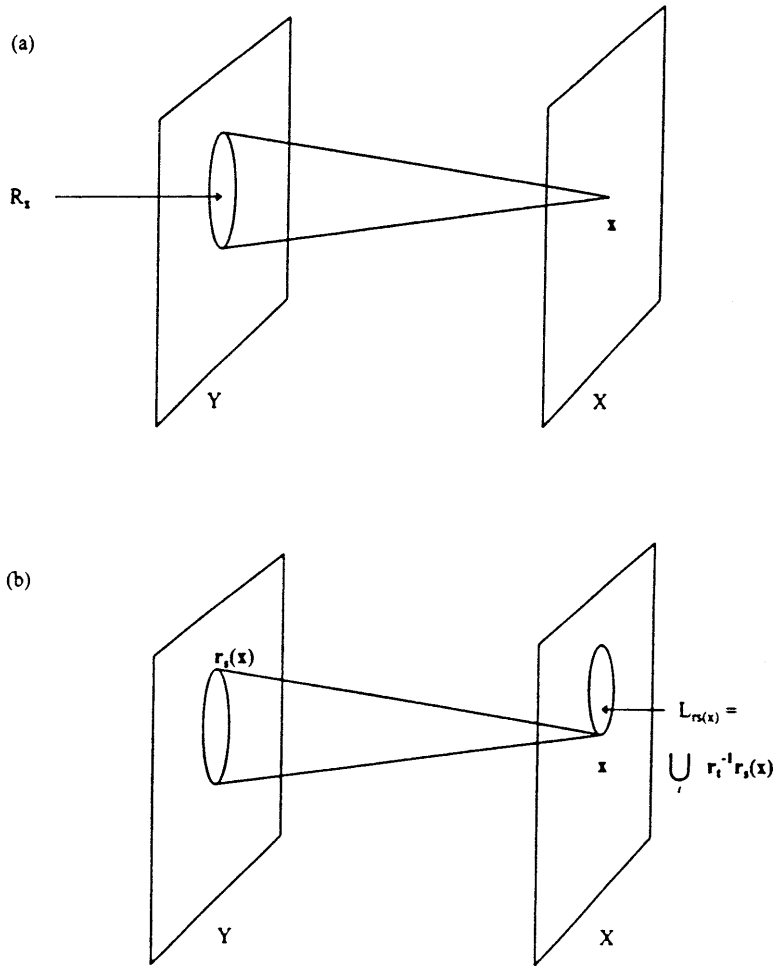

Fig. 3. **a** The nature of the receptive field R_x on the thalamic field Y of the neuron at the position $\mathbf{x}$ on the cortical sheet X. **b** Definition of the boundary of R_x by means of the parameterised function $\mathbf{r}_s(\mathbf{x})$ and the boundary of the resulting projective field of $\mathbf{r}_s(\mathbf{x})$, denoted by $L_{rs(x)}$

$$\int d\mathbf{x}' w(\mathbf{x}-\mathbf{x}')\theta[u(\mathbf{x}', \mathbf{r}_s(\mathbf{x}), t)] = W(l_s(t)) \tag{81}$$

with l_s defined by (73), so that from (75)

$$\tau \partial u(\mathbf{x}, \mathbf{r}_s(\mathbf{x}), t)/\partial t = \tau \partial W/\partial l_s \cdot \partial l_s/\partial t + W(l_s(t)) + K(r(t)) + h \tag{82}$$

Following the methods of [4], and using (79), (80), (81) and (82), (78) can be reduced to the form

$$\tau \alpha_1(\mathbf{x}, t) \cdot \partial \mathbf{r}_s/\partial t + X_1 = -[W(l_s(t)) + K(r(t)) + h] \tag{83}$$

$$\tau \alpha_2(\mathbf{x}, t) \cdot \partial \mathbf{r}_{s'}/\partial t + X_2 = -[W(l_{s'}(t)) + K(r(t)) + h] \tag{84}$$

where

$$\begin{aligned} &\alpha_1(\mathbf{x}, t) = \partial u(\mathbf{x}, \mathbf{r}_s(\mathbf{x}), t)/\partial \mathbf{y}, \alpha_2(\mathbf{x}, t) = \partial u(\mathbf{x}, \mathbf{r}_{s'}(\mathbf{x}), t)/\partial \mathbf{y} \\ &X_1 = l_s^{-1} \mathbf{l}_s(\mathbf{x}) \cdot \mathbf{A}_1 \\ &\mathbf{A}_1 = \mathbf{L}_1^{-1} \mathbf{B}_1, \quad \mathbf{B}_1 = \partial \mathbf{r}_{s'}(\mathbf{x} - \mathbf{l}_s(\mathbf{x})), t)/\partial t \\ &\qquad - \partial \mathbf{r}_s(\mathbf{x}, t)/\partial t \\ &L_{1ij} = \partial r_{s'i}(\mathbf{x} - \mathbf{l}_s(\mathbf{x})), t)/\partial x_j \end{aligned} \tag{85}$$

$$\begin{aligned} &X_2 = l_s^{-1} \mathbf{A}_2 \\ &\mathbf{A}_2 = \mathbf{L}_2^{-1} \mathbf{B}_2, \quad \mathbf{B}_2 = -\partial \mathbf{r}_s(\mathbf{x} + \mathbf{l}_{s'}(\mathbf{x}), t)/\partial t \\ &\qquad + \partial \mathbf{r}_{s'}(\mathbf{x}, t)/\partial t \\ &L_{2ij} = \partial r_{si}(\mathbf{x} + \mathbf{l}_{s'}(\mathbf{x})), t)/\partial x_j \end{aligned} \tag{86}$$

Note that the extension (83), (84) of (24) and (25) of [4] are still linear in the time derivatives $\partial \mathbf{r}_s(\mathbf{x}, t)/\partial t$, $\partial \mathbf{r}_{s'}(\mathbf{x}, t)/\partial t$. It is to be expected that a corresponding extension of the stability analysis of [4] can be achieved. We will return to that after discussing the equilibrium solutions to these equations.

4.4 Equilibrium solutions

The equations of temporal equilibrium resulting from (83) and (84) are

$$[W(l_s(t)) + K(r(t)) + h] = 0 \tag{87}$$

Again, as in [4], only the solution

$$l_s(\mathbf{x}) = l(\mathbf{x}) \tag{88}$$

where $l(\mathbf{x})$ is independent of s is possible. Since also

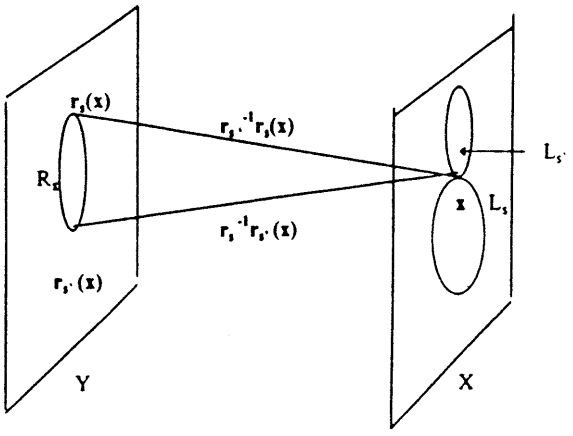

Fig. 4. More complete delineation of the relation between the parametrised boundary of R_x and the mappings which generalise those in the one-dimensional case of [3] and [4]

$$\mathbf{l}_{s'}(\mathbf{x}) = \mathbf{l}_s(\mathbf{x} + \mathbf{l}_{s'}(\mathbf{x})), \quad \mathbf{l}_s(\mathbf{x}) = \mathbf{l}_{s'}(\mathbf{l}_s(\mathbf{x}) - \mathbf{x}) \tag{89}$$

then from (88) it follows that

$$l(\mathbf{x}) = l(\mathbf{x} + \mathbf{l}_{s'}(\mathbf{x})) = l(\mathbf{l}_s(\mathbf{x}) - \mathbf{x}) \tag{90}$$

A simple solution to (90) is the constant one:

$$l(\mathbf{x}) = l_0 \tag{91}$$

so from (87) and (88), since $K(r(t))$ is continuous in $r(t)$, then also $r(t)$ is constant

$$r(t) = r_0 \tag{92}$$

with

$$K(r_0) + W(l_0) + h = 0 \tag{93}$$

A solution of the above uniform distribution of receptive fields is

$$\mathbf{r}_s(\mathbf{x}) = p\mathbf{x} + \mathbf{r}_s$$

$$\mathbf{r}_{s'}(\mathbf{x}) = p\mathbf{x} + \mathbf{r}_{s'}$$

$$\mathbf{r}_{s'} = (1/2)r_0(\cos s', \sin s')$$
$$\mathbf{r}_s = -(1/2)r_0(\cos s, \sin s) \tag{94}$$

where p is the cortical magnification factor $p = r_0/l_0$.

It is also possible to generalise solution (94) to the periodic solution

$$\mathbf{r}_s(\mathbf{x}) = p\mathbf{x} + \mathbf{g}_s(\mathbf{x})$$
$$\mathbf{r}_{s'}(\mathbf{x}) = p\mathbf{x} + \mathbf{g}_s(\mathbf{x}) + p\mathbf{l}_{s'}(\mathbf{x}) \tag{95}$$

where $\mathbf{g}_s(\mathbf{x})$ is periodic in $\mathbf{x}$ with period $\mathbf{l}_s(\mathbf{x})$, where from (94) we obtain

$$\mathbf{l}_s(\mathbf{x}) = (1/p)(\mathbf{r}_s - \mathbf{r}_{s'}) = -\mathbf{l}_{s'}(\mathbf{x}) \tag{96}$$

and both $\mathbf{l}_s(\mathbf{x})$ and $\mathbf{l}_{s'}(\mathbf{x})$ are independent of $\mathbf{x}$. Thus, $\mathbf{g}_s(\mathbf{x})$ is periodic under translations in two dimensions along the direction given by the angle s.

4.5 Stability of the equilibrium solutions

It is possible to extend the treatment of [4] to the two-dimensional case by considering small variations about the equilibrium solutions (94), (95); we will only consider the constant solution (94). For the small variations

$$\mathbf{r}_s(\mathbf{x}, t) = \mathbf{r}_s + \varepsilon \mathbf{v}_s(\mathbf{x}, t) \tag{97}$$

for all $s \in [0, 2\pi]$, then

$$\mathbf{r}_{s'}(\mathbf{x}, t) - \mathbf{r}_s(\mathbf{x}, t) = 2\mathbf{r}_s + \varepsilon[\mathbf{v}_{s'}(\mathbf{x}, t) - \mathbf{v}_s(\mathbf{x}, t)] \tag{98}$$

Let

$$\mathbf{l}_s(\mathbf{x}, t) = \mathbf{l}_s + \varepsilon \mathbf{l}'_s(\mathbf{x}, t) + O(\varepsilon^2) \tag{99}$$

From (79) it follows that

$$\mathbf{r}_{s'}(\mathbf{x} - \mathbf{l}'_s(\mathbf{x}, t) - \mathbf{l}_s, t) = \mathbf{r}_s(\mathbf{x}) + \varepsilon \mathbf{v}_s(\mathbf{x}, t) \tag{100}$$

so that one can show that

$$l'_{si} = (\mathbf{L}^{-1}\mathbf{A})_i$$

where

$$L_{ji} = \partial r_{s'j}(\mathbf{x} - \mathbf{l}_s, t)/\partial x_i,$$
$$A_j = \varepsilon[v_{s'j}(\mathbf{x} - \mathbf{l}_s(\mathbf{x}), t) - v_{sj}(\mathbf{x}, t)] \tag{101}$$

Moreover, $\mathbf{B}_1$ and $\mathbf{B}_2$ of (85) and (86) are already of $O(\varepsilon)$, so that only the constant solutions for $\mathbf{r}_s(\mathbf{x})$ need be used in calculating $\mathbf{L}_1$ and $\mathbf{L}_2$, resulting in

$$\mathbf{L}_1 = \mathbf{L}_2 = p\mathbf{I} \tag{102}$$

and so

$$\mathbf{A}_1 = [\partial \mathbf{v}_{s'}(\mathbf{x} - \mathbf{l}_s, t)/\partial t - \partial \mathbf{v}_s(\mathbf{x}, t)/\partial t]/p$$
$$\mathbf{A}_2 = [\partial \mathbf{v}_s(\mathbf{x}, t)/\partial t - \partial \mathbf{v}_{s'}(\mathbf{x} + \mathbf{l}_{s'}, t)/\partial t]/p \tag{103}$$

and from (101)

$$\mathbf{L} = p\mathbf{I}$$

so that

$$l'_{si} = \varepsilon[v_{s'j}(\mathbf{x} - \mathbf{l}_s(\mathbf{x}), t) - v_{sj}(\mathbf{x}, t)]/p \tag{104}$$

Collecting together (103) and the expansions

$$K(r(t)) = K(r_0) + (\varepsilon/pr_0)K'(r_0)[\mathbf{v}_{s'}(\mathbf{x}, t) - \mathbf{v}_s(\mathbf{x}, t)]\mathbf{r}_s$$
$$W(l) = W(l_0) + (\varepsilon/pr_0)[\mathbf{v}_{s'}(\mathbf{x}, t) - \mathbf{v}_s(\mathbf{x}, t)] \cdot \mathbf{l}_s \tag{105}$$

leads to the first-order equation

$$\tau(\alpha_s + \beta_s) \cdot \partial \mathbf{v}_s(\mathbf{x}, t)/\partial t - \tau\beta_s \cdot \partial \mathbf{v}_s(\mathbf{x} - \mathbf{l}_s, t)/\partial t$$
$$= -\beta_1 \cdot [\partial \mathbf{v}_s(\mathbf{x}, t)/\partial t - \partial \mathbf{v}_{s'}(\mathbf{x} - \mathbf{l}_s, t)/\partial t]$$
$$- K'(r_0)[\mathbf{v}_{s'}(\mathbf{x}, t) - \mathbf{v}_s(\mathbf{x}, t)] \cdot \mathbf{r}_s(1/pr_0) \tag{106}$$

where α_s is equal to α_1 of (85), but now has its dependence on the parameter s emphasised, and

$$\beta_s = -\mathbf{l}_s/(pl_0) \tag{106a}$$

Equation (106) and its companion with s and s' interchanged throughout are identical in form to the one-dimensional variational equation of [4].

The stability of the soluble solutions can now be investigated further by the use of Fourier expansions, as in [4]. Thus, with the expansion

$$\mathbf{v}_s(\mathbf{x},t) = \sum_n \mathbf{V}_s(\mathbf{n},t)\exp[iv\mathbf{n}\cdot\mathbf{x}] \tag{107}$$

with $v = 2\pi/L_x$ where L_x is the length of the cortical field, assumed to be a square, then (106) becomes

$$\tau\mathbf{A}_n\cdot\partial\mathbf{W}_s/\partial t = \tau\mathbf{B}_n\cdot\mathbf{W}_s \tag{108}$$

where $\mathbf{W}_s$ is the 4-component column vector with row components $[\mathbf{V}_s^T(\mathbf{x},t)\mathbf{V}_{s'}^T(\mathbf{x},t)]$ and $\mathbf{A}_n$, $\mathbf{B}_n$ are the 2×4 matrices

$$\mathbf{A}_n = \begin{pmatrix}(\alpha_s+\beta_s)^T & -\beta_s^T z^*\\ (-\beta_s^T z & (\alpha_s+\beta_s)^T)\end{pmatrix} \tag{109a}$$

$$\mathbf{B}_n = \begin{pmatrix}(\mathbf{k}_s-\beta_s)^T & \beta_s^T z^*-\mathbf{k}_s^T\\ (\beta_s^T z-\mathbf{k}_s^T & (\mathbf{k}_s-\beta_s)^T)\end{pmatrix} \tag{109b}$$

with $z=\exp[iv\mathbf{n}\cdot\mathbf{l}_s]$, $\mathbf{k}_s = K'(r_0)\mathbf{r}_s$. As usual, we search for the time development of $\mathbf{W}_s$ as

$$\mathbf{W}_s = \exp[\lambda t/\tau]\mathbf{w}_s \tag{110}$$

where $\mathbf{w}_s$ is assumed to be time-independent. We obtain the usual eigenvalue condition on λ, from (108) and (110), that

$$[\lambda\,\mathbf{A}_n - \mathbf{B}_n]\mathbf{w}_s = 0 \tag{111}$$

As in [4], the coefficients α_s and β_s are independent of position and time, α_s from the homogeneity of the solution being considered (the translation-invariant one) and β_s from (106a).

A more detailed analysis of these vector-valued constants α_s and β_s indicates that they are all proportional to the vector $\mathbf{r}_s$. This is immediately so for the latter, from (96) and (106a). For the former it follows from the definition of (85), the translation invariance of the solution $u(\mathbf{x},\mathbf{r}_s+p\mathbf{x})$, and the fact that the gradient of u, for an input translated by the vector $\mathbf{r}_s$, will be along the direction of the translate. Finally, the vector $\mathbf{k}_s$ lies along $\mathbf{r}_s$ as well. Thus, in total results the vector decomposition

$$\mathbf{A}_n = \mathbf{r}_s\otimes\begin{pmatrix}(\alpha_s+\beta_s) & -\beta_s z^*\\ (-\beta_s^T z & (\alpha_s+\beta_s))\end{pmatrix}$$

$$\mathbf{B}_n = \mathbf{r}_s\otimes\begin{pmatrix}(k_s-\beta_s) & \beta_s^T z^*-k_s\\ (\beta_s z-k_s & (k_s-\beta_s))\end{pmatrix} \tag{112}$$

where the 2×2 matrices in (112) are identical to those arising in the one-dimensional analysis of [4] (with the change that $k(r_0)$ in the one-dimensional case has now become $K'(r_0)$. The small perturbations $\mathbf{w}_s$ can now be expressed as

$$\mathbf{w}_s = \mathbf{r}_s\otimes(u,v) \tag{113}$$

and the resulting equation for the two-component vector (u,v) from (111), (112) and (113) is identical to the one-dimensional case of [4] (given the difference noted above). The following theorem extending that of the one-dimensional case then results, using the same arguments as in [4]:

Theorem 3

The constant equilibrium solution is stable when $K'(r_0)<0$ and unstable when $K'(r_0)>0$.

It is still to be shown that theorem 3 holds for the periodic solutions; the theorem is expected to be true in that case also, following the lines of [4].

5 Learning a general pattern set

5.1 General formalism

The above discussion was only for point inputs to the presynaptic field Y. However, there are more general classes of inputs which need to be discussed. Let us consider a pattern set Π composed of patterns p in space. These could be of the form of bars at various orientations, oriented edges or other shapes distributed in the plane. The elements of Π have a set of parameters associated with them, say the centres of the bars and their orientations, their lengths and so on. This set of parameters for the pattern p of the set Π will be denoted by $\mathbf{p}$, and their associated probability density for the choice of the pattern with those parameters by $P(\mathbf{p})$. The input $S(\mathbf{x},\mathbf{p})$ to the cortical sheet will therefore have the value obtained by extending (63) to the specific visual input corresponding to the pattern p,

$$S(\mathbf{x},\mathbf{p}) = \int s(\mathbf{x},\mathbf{y}')a(\mathbf{y}'-\mathbf{y})I(\mathbf{y},\mathbf{p})\,d\mathbf{y}'\,d\mathbf{y} \tag{114}$$

The learning equation for the afferent synapses, replacing (61), becomes (including summation over the pattern set Π):

$$\tau\partial s(\mathbf{x},\mathbf{y})/\partial t = -s(\mathbf{x},\mathbf{y}) + \varepsilon\int P(\mathbf{p})a(\mathbf{y}'-\mathbf{y})I(\mathbf{y}',\mathbf{p}) \times\theta[u(\mathbf{x},\mathbf{p})]\,d\mathbf{y}'\,d\mathbf{p} \tag{115}$$

where $u(\mathbf{x},\mathbf{p})$ is the extension of the asymptotic value of the membrane potential of (65) to the case when inputs are chosen from the pattern set Π with probability distribution $P(\mathbf{p})$:

$$u(\mathbf{x},\mathbf{p}) = S(\mathbf{x},\mathbf{p}) + \int w(\mathbf{x}-\mathbf{x}')\theta[u(\mathbf{x}',\mathbf{p})]\,d\mathbf{x}' + h \tag{116}$$

Using the asymptotic state of the afferent synapses from the learning equation (115), then the asymptotic state of the membrane potential satisfied the condition

$$u(\mathbf{x},\mathbf{p}) = \varepsilon \int P(\mathbf{q})k(\mathbf{y}-\mathbf{y}')I(\mathbf{y},\mathbf{p})I(\mathbf{y}',\mathbf{q})\theta[u(\mathbf{x},\mathbf{q})]\,d\mathbf{y}\,d\mathbf{y}'\,d\mathbf{q} + \int w(\mathbf{x}-\mathbf{x}')\,\theta[u(\mathbf{x}',\mathbf{p})]\,d\mathbf{x}' + h \tag{117}$$

This can be simplified by defining the kernel

$$K(\mathbf{p},\mathbf{q}) = \varepsilon \int k(\mathbf{y}-\mathbf{y}')I(\mathbf{y},\mathbf{p})I(\mathbf{y}',\mathbf{q})\,d\mathbf{y}\,d\mathbf{y}' \tag{118}$$

and then (117) may be written more compactly as

$$u(\mathbf{x},\mathbf{p}) = \int_{RF(x)} P(\mathbf{q})K(\mathbf{p},\mathbf{q})\,d\mathbf{q} + \int_{PF(p)} w(\mathbf{x}-\mathbf{x}')\,d\mathbf{x}' + h \tag{119}$$

where RF(**x**) is the receptive field of cortical neural site **x** and PF(**p**) is the projective field of the pattern **p** as defined earlier:

$$\mathrm{RF}(\mathbf{x}) = \{\mathbf{q}: u(\mathbf{x},\mathbf{q}) > 0\}$$
$$\mathrm{PF}(\mathbf{p}) = \{\mathbf{x}' : u(\mathbf{x}',\mathbf{p}) > 0\} \tag{120}$$

Equation (119) has to be solved for a given pattern set Π and is extremely complicated in general in spite of its deceptively simple appearance.

5.2 A finite pattern set

It is possible to analyse (119), as in [3], by reduction of Π to a finite set. This approach will be considered briefly now. For the finite set of patterns $\{p_i\}$ (for $1 \le i \le P$), we obtain from (119) the system of P equations

$$u(\mathbf{x},\mathbf{p}_i) = \sum_j K_{ij}\theta[u(\mathbf{x},\mathbf{p}_j)] + \int w(\mathbf{x}-\mathbf{x}')\theta[u(\mathbf{x}',\mathbf{p}_i)]\,d\mathbf{x}' + h \tag{121}$$

or, more compactly,

$$u_i(\mathbf{x}) = \sum_j K_{ij}\theta[u_j(\mathbf{x})] + \int w(\mathbf{x}-\mathbf{x}')\theta[u_i(\mathbf{x}')]d\mathbf{x}' + h \tag{122}$$

where $u_i(\mathbf{x}) = u(\mathbf{x},\mathbf{p}_j)$ and $K_{ij} = K(\mathbf{p}_i,\mathbf{p}_j)$. As argued in [5], when $w = 0$ there is no coupling between different neurons on the cortical sheet, and the pattern excited for the neuron at **x** will be p_i if

$$K_{ii} + h > 0, \quad \lambda_I \equiv -h/K_{ii} < 1 \tag{123}$$

Then the set of patterns which excite the neuron at **x** will be

$$Rf_x = \{\mathbf{p} : K(\mathbf{p},\mathbf{p}_i) > \lambda_I k_{ii} \tag{124}$$

Equation (124) gives the approximate resolution of the map, so extending theorem 3 of [5].

It is also possible to obtain an approximate value for the boundary of the projective field of a pattern by including the lateral connection term w in (123), so that when the projective fields of two patterns have a negligible overlap, we obtain

$$u_i(\mathbf{x}) = K_{ij} + \int w(\mathbf{x}-\mathbf{x}')\theta[u_j(\mathbf{x}')\,d\mathbf{x}' + h \tag{125}$$

The boundary of the projective field of the pattern p_i, $\partial\mathrm{PF}(p_i)$, is therefore defined by setting (125) to zero, so leading to

$$W_i(\mathbf{x}) = h - K_{ij} \tag{126}$$

so that

$$\partial\mathrm{PF}(p_i) = W_i^{-1}(-\infty, h - K_{ij}] \tag{127}$$

where

$$W_i(\mathbf{x}) = \int w(\mathbf{x}-\mathbf{x}')\theta[u_j(\mathbf{x}')]\,d\mathbf{x}' \tag{128}$$

The result (128) extends theorem 4 of [5].

6 Discussion

The paper has given for the first time an extension of the CNFT theory of [4] and [5] to two dimensions. In particular, a two-dimensional theory has been developed of the following:

1. A complete catalogue of the form of circularly symmetric bubbles, and the corresponding threshold parameter range under which they exist and are stable.
2. The detailed temporal evolution of the size of a two-dimensional bubble of finite size.
3. The lifetime for the decay of bubbles due to the presence of habituating ionic currents, in terms of the parameters entering the habituation term,
4. Definition of receptive and projective fields in two dimensions,
5. Extension of the detailed analysis of learning laws from one to two dimensions, with development of the system of equations for the dynamics of the receptive fields, the nature of their equilibrium solutions, and their stability analysis,
6. Formulation of the learning equations for a general pattern set.

There are still many unanswered aspects of this area, but the most important concern, what the theory developed so far is good for in cortical modelling.

Various applications of CNFT were described in the introduction. These included analysis and learning of visual orientation sensitivity, of bubbles on the superior colliculus guiding saccades and of changes of somatosensory receptive fields due to training experience. The

Table 1. Comparison between Miller's [33] and the CNFT models

	Miller et al.	CNFT
Nature of cortical neurons	Linear	Threshold response
Mechanism for limitation of synaptic strength	Clipping and normalisation by subtraction	Inhibitory and adaptive threshold
Nature of lateral cortical connectivity	Either purely excitatory or of Mexican hat form	Mexican hat

following topics are presently accessible to an approach by means of CNFT [30]: binocularity, orientation sensitivity, bubbles in coupled networks, learning in coupled networks, apparent motion, colour stabilisation, stabilised objects, higher level visual processing, bubbles in auditory processing, a general CNFT framework for sensory processing.

The above list indicates that the approach of CNFT can provide a general framework to incorporate some of the principles of sensory processing. That is not surprising since the approximation of the cortical sheet as a continuum should be able to be used to construct a neural approach to sensory processing. What is unclear from the beginning is that this approach can provide a mathematically tractable analytic framework. In particular, it will lead into difficult non-linear problems arising from the analysis of interacting CNFT modules.

The development of mathematical solutions to the various topics in the above list, including some of those arising from coupled CNFT modules, shows this possibility to be feasible. There will be many difficult mathematical problems still left unanswered at the end of the next part of my series, but at least the framework will have been erected and various insights gained into the manner in which the coupled sensory CNFT networks combine together to solve the problem of making sense of the visual world.

There is also the question as to the nature of the 'bubbles' of neural activity which have been considered at some length here and which will be further considered in the next paper of my series. Let me attempt to summarise here the nature of these bubbles and their importance for the whole approach to learning. The basic thesis is that there are activity bubbles created by an input, which may or may not remain after input has been turned off, depending on the parameters of the network. Whatever the final status of the bubbles, during the learning process in the presence of inputs, the activity bubbles so formed allow Hebbian learning to proceed. The bubbles themselves can be analysed mathematically to determine the influence of the spatial stimulus correlations on the resulting receptive fields of the neurons. One can thereby proceed to develop a complete theory for the formation of a topographic map and compare this to other approaches; this has been especially emphasised in [31], where a detailed relation was made to the considerable body of work on this elaborated in the review in [32].

There is one particular approach which has been considerably successful in its explanation of a considerable body of experimental data, that of Miller and his colleagues [33]. The relation of this work to our approach using CNFT can only be considered at the level of assumptions, not of results. At that level we see that both use a continuum cortical field, with the more detailed comparisons given in Table 1.

It is clear from the table that the two approaches are somewhat different in their basic assumptions. Yet both arrive at topographic maps and binocular disparity maps. However, an important difference between the two approaches is that for the model of [33] the cortical neuronal response is linear, and thus is not restricted in its spatial extent. For CNFT the activity is in a limited spatial region only, and the analysis is in terms of the boundary regions at the edges of bubbles and their dynamics. It is in any case necessary to wait to be able to compare the two models until the results of detailed analysis is presented of the CNFT approach for this specific problem.

There is finally the question of how good is the basic CNFT framework, with mean firing rate neurons and no delay times, refractory times or synaptic noise. This will be considered elsewhere.

Acknowledgement. I would like to thank the Director of the Institute of Medicine, Research Centre Juelich for hospitality during the performance of this work, and the stimulating atmosphere he has created there.

Appendix A

Rewriting w as an explicit function of its squared variable, $w = w(x^2)$ leads to the explicit formulae for the first and second derivatives of G of (16), when written in radial co-ordinates r, θ as

$$G(R) = \int_0^R x\,dx \int_0^{2\pi} d\theta\, w(R^2 + x^2 - 2Rx\cos\theta) \tag{129}$$

$$\begin{aligned} dG/dR = R &\int_0^{2\pi} d\theta\, w(2R^2(1-\cos\theta)) \\ &+ \int_0^R x\,dx \int_0^{2\pi} d\theta\, 2[R - x\cos\theta]w'(R^2 + x^2 - 2Rx\cos\theta) \end{aligned} \tag{130}$$

where w' is the derivative of w with respect to its variable as written in (129). Furthermore

$$\begin{aligned} d^2G/dR^2 = &\int_0^{2\pi} d\theta\ w(2R^2(1-\cos\theta)) \\ &+ 6R^2 \int_0^{2\pi} d\theta[1-\cos\theta]w'(2R^2(1-\cos\theta)) \\ &+ 2\int_0^R x\,dx \int_0^{2\pi} d\theta\ w'(R^2 + x^2 - 2Rx\cos\theta) \\ &+ \int_0^R x\,dx \int_0^{2\pi} d\theta\ 4[R - x\cos\theta]w''(R^2 + x^2 - 2Rx\cos\theta) \end{aligned} \tag{131}$$

On setting $R = 0$ in equation (130) and (131), the results of (17) immediately ensue.

Appendix B

Proof of theorem 1

1. For a $\oslash$ solution, from (2) such a solution will be $u(x) = h$, so $h < 0$. Conversely if $h < 0$, then the null solution $u(x) = h$ will satisfy (2).
2. An ∞-solution will have the form, from (2), of

$$u(x) = \int_{R\mathbf{x}R} d^2x' w(|\mathbf{x} - \mathbf{x}'|^2) + h = W_\infty + h \tag{132}$$

so that for this to be a solution will require

$$W_\infty > -h \tag{133}$$

Conversely if (133) is valid, then $u(\mathbf{x}) = W_\infty + h$ is a solution to (2) which is an ∞-solution.
3. An $R2$-solution satisfies (11) and must satisfy

$$G(R) + h = 0 \tag{134}$$

with G defined by (16). Conversely, if (16) is satisfied, then the putative solution of (11) will vanish on the boundary of the disc of radius R and centre zero. Further, as in [3] it is possible to see, from the shape of $W(x)$, that provided that $h < 0$, then u is positive inside the disc and negative outside.

Proof of theorem 2

This is very similar to that in [3] for the one-dimensional case, except for the spatially periodic solution of that reference in the parameter range $0 < h < -W_\infty$, when no $\oslash$, ∞ or $R[2]$ solution can exist. Neglecting the periodic solutions in one variable (which were noted earlier in Sect. 2), the two-dimensional version of the one-dimensional periodic solutions are expected to possess annular regions of inner and outer radii $nb, nb + a$, for $n = 0, 1, 2, \ldots$, provided there is a solution to the conditions

$$\sum_n \; [W(x, nb + a) - W(x, nb)] + h = 0 \tag{135}$$

for any $x = mb$ or $mb + a$, for any $m = 0, 1, 2, \ldots$ Equation (135) will have a set of solutions for a, b when $0 < h < -W_\infty$, since the left-hand side of (135) is a function $H(a)$ of a which has the values

$$H(0) = h > 0, H(b) = W_\infty, +h < 0 \tag{136}$$

Thus, $H(a)$ must have a zero for a in the interval $(0, b)$, for each b.
The stability of the solutions of (134) was shown in Sect. 2.

References

1. Wilson HR, Cowan JD (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13:55–80
2. Feldman J, Cowan JD (1975) Large-scale activity in neural nets I. Theory with applications to motoneuron pool responses. Biol Cybern 17:29–38
3. Amari S-I (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biol Cybern 27:77–87
4. Takeuchi A, Amari S-I (1979) Formation of topographic maps and columnar microstructures in nerve fields. Biol Cybern 35:63–72
5. Amari S-I (1989) Dynamical study of formation of cortical maps. In: Arbib MA, Amari S-I (eds) Dynamic interactions in neural networks: models and data. Springer, Berlin Heidelberg New York
6. Burns D (1950) The uncertain nervous system. Cambridge University Press, Cambridge, UK
7. Griffiths JS (1963) A field theory of neural nets I. Bull Math Biophys 25:111–120; ibid (1965) A field theory of neural nets II. Bull Math Biophys 27:187–195
8. Beurle RL (1956) Properties of a mass of cells capable of regenerating pulses. Trans R Soc Lond Biol 240:55–94
9. Ermentrout GB, Cowan JD (1980) Large scale spatially organized activity in neural nets. SIAM J Appl Math 38:1–21
10. Ermentrout GB, Cowan JD (1979) A mathematical theory of visual hallucination patterns. Biol Cybern 34:136–150
11. Kopecz K, Schoner G (1995) Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. Biol Cybern 73:49–60
12. Kuroiwa J, Miyake S, Inawashiro S, Aso H, Tampri Y (1996) Self-organization in a formal neuron model and self-consistent Monte-Carlo simulation. Proceedings International Joint Conference on Neural Networks'96
13. Zhang J (1991) Dynamics and formation of self-organising maps. Neural Comput 3:54–66
14. Zhang K (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. J Neurosci 16:2112–2126
15. Petersen R, Taylor JG (1996) Reorganisation of somatosensory cortex after tactile training. In: Touretzky DS, Mozer MC, Hasselmo ME (eds). Advances in Neural Information Processing Systems. MIT Press, Cambridge, Mass. pp 82–88
16. Horn D, Opher I (1997) Solitary waves of integrate and fire neural fields. Neural Comput 9:1677–1690
17. Nowak LG, Bullier J-P (1997) The timing of information transfer in the visual system. In: Kaas J, Rockland K, Peters A (eds) Extrastriate explorations. (Cerebral cortex, Vol 12) (in press)
18. Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868–1871
19. Lu Z-L, Williamson SL, Kaufmann L (1992) Human auditory primary and association cortex have different lifetimes for activation traces. Brain Res 572:236–241
20. Uusitalo M, Williamson SJ, Seppa MT (1996) Dynamical organisation of the human visual system revealed by lifetimes of activation traces. Neurosci Lett 213:149–152
21. Gerstner W, Hemmen JL van (1996) Coherence and incoherence in a globally coupled ensemble of pulse-emitting units. Phys Rev Lett 71:312–315
22. White EL (1989) Cortical circuits: synaptic organisation of the cerebral cortex. Structure, function and theory. Birkhauser, Boston
23. Ermentrout B (1997) Neural nets as spatio-temporal pattern-forming systems. University of Pittsburgh preprint
24. Pinto D (1997) Computation, experimental and analytical explorations of neuronal circuits in the cerebral cortex. PhD Thesis, University of Pittsburgh, Department of Mathematics
25. Douglas RJ, Koch C, Mahowald M, Martin KAC, Suarez HH (1995) Recurrent excitation in neocortical circuits. Science 269:981–985
26. Ulman S (1982) The interpretation of visual motion. MIT Press, Cambridge, Mass.
27. Schwindt PC, Spain WJ, Foehring RC, Stafstrom CE, Chubb MC, Crill WE (1988) Multiple potassium conductances and their functions in neurons from cat sensorimotor cortex in vitro. J Neurophysiol 59:424–449; ibid Slow conductances in neurons from cat sensorimotor cortex in vitro and their role in slow excitability changes. J Neurophysiol 59:450–467

28. Barbas H, Pandya DN (1992) Patterns of connections of the prefrontal cortex in the rhesus monkey associated with cortical architecture. In: Levin HS, Eisenberg HM, Benton AL (eds) Frontal function and dysfunction. Oxford University Press, Oxford
29. Mesulam MM (1981) Patterns of behavioural anatomy: association areas, the limbic system and hemispheric specialisation. Ann Neurol 10:309–325
30. Taylor JG (1997) Invited talk. World Joint Conference on Neural Networks, Washington, DC
31. Petersen RS (1997) The neural field theory approach to cortical self organisation. PhD Thesis, University of London, King's College, Department of Mathematics
32. Swindale NV (1996) The development of topography in the visual cortex: a review of models. Network: computation in neural systems 7:161–247
33. Miller KD, Keller JB, Stryker MP (1989) Ocular dominance column development: analysis and simulation. Science 245:605–615