4. KOHONEN'S NETWORK MODEL

This chapter describes Kohonen's network model. We will discuss how the cells of a neuron layer coordinate their sensitivity to sensory signals in such a way that their response properties to signal features vary in a regular fashion with their position in the layer, an organization observed in many parts of the brain. After some neurophysiological background information, a mathematical formulation of the model will be presented. Simulations will give a first impression of the main features of the model.

4.1 Neurophysiological Background

The model employs a neuron layer A, usually assumed to be a two-dimensional sheet. This layer is innervated by d input fibers (*axons*), which carry the input signal and excite or inhibit the neurons of the layer via synaptic connections, as illustrated schematically in Fig. 4.1. In the following, we consider conditions under which the excitation of the neurons is restricted to a spatially localized region in the layer. The location of this region is then determined by those neurons that respond most intensively to the given stimulus. The neuron layer acts as a *topographic feature map*, if the location of the most strongly excited neurons is correlated in a regular and continuous fashion with a restricted number of signal features of interest. Neighboring excited locations in the layer then correspond to stimuli with similar features. Of course, a single layer can only make a few important features visible in this way. In the simplest case, we may be dealing with the stimulus position on a sensory surface, such as the retina or the body's outer surface; simple examples of more abstract features are pitch and intensity of sound signals.

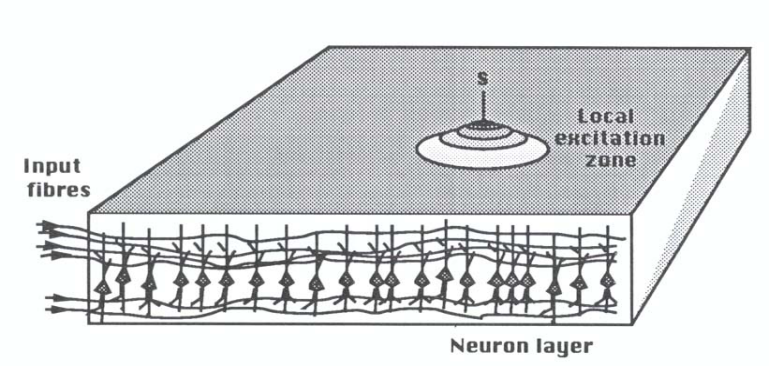

Abb. 4.1: Schematic representation of the neuron layer in Kohonen's model. The nerve fibers running horizontally provide the input signal ("stimulus") and excite the layer neurons via synaptic connections. Lateral interactions between the neurons constrain the reaction to a spatially bounded "excitation zone." The layer acts as a "topographical feature map" if the position s of the excitation zone varies in a continuous way with the presence of stimulus features of interest

We now describe the principles which enable the formation of such topographic feature maps in Kohonen's model by means of a self-organizing process. An incoming signal $\mathbf{v}$ is given by the average activities v_l of the individual incoming fibers $l = 1, 2, \ldots$. We identify the neurons of the layer by their two-dimensional position vectors $\mathbf{r} \in A$, with A being a two-dimensional grid. Every neuron $\mathbf{r}$ forms in its dendritic tree a weighted sum $\sum_l w_{\mathbf{r}l} v_l$ of the incoming activities v_l, where $w_{\mathbf{r}l}$ expresses the "strength" of the synapse between axon l and neuron $\mathbf{r}$. Here, $w_{\mathbf{r}l}$ is positive for an excitatory synapse and negative for an inhibitory synapse. The resulting excitation of an isolated neuron $\mathbf{r}$ is described by its average spike frequency $f^0_{\mathbf{r}}$. Usually, a relation

$$f^0_{\mathbf{r}}(\mathbf{v}) = \sigma\left(\sum_l w_{\mathbf{r}l} v_l - \theta\right) \tag{4.1}$$

is assumed for $f^0_{\mathbf{r}}$. Here, $\sigma(x)$ is a "sigmoid" function, increasing monotonically with x, with a qualitative behavior as shown in Fig. 3.10. In particular, $\sigma(x)$ tends asymptotically to the saturation values 0 or 1 for $x \to \pm\infty$. The quantity θ acts as an excitation threshold, below which the neuron responds weakly.
In addition to the coupling to the input fibers, the neurons are connected

to each other via synapses. Thus, the layer has internal feedback. If one designates by $g_{\mathbf{r}\mathbf{r}'}$ the coupling strength from neuron $\mathbf{r}'$ to neuron $\mathbf{r}$, any excitation $f_{\mathbf{r}'}$ of neuron $\mathbf{r}'$ provides a contribution $g_{\mathbf{r}\mathbf{r}'} f_{\mathbf{r}'}$ to the total input signal of neuron $\mathbf{r}$. The contributions of all neurons $\mathbf{r}'$ in the layer are additively superimposed onto the external input signal $\sum_l w_{\mathbf{r}l} \mathbf{v}_l$. In the stationary case, the neuron activities $f_{\mathbf{r}}$ are thus the solution of the nonlinear system of equations

$$f_{\mathbf{r}} = \sigma \left(\sum_l w_{\mathbf{r}l} v_l + \sum_{\mathbf{r}'} g_{\mathbf{r}\mathbf{r}'} f_{\mathbf{r}'} - \theta \right). \tag{4.2}$$

Frequently, the feedback accounted for by $g_{\mathbf{r}\mathbf{r}'}$ is due to excitatory synapses ($g_{\mathbf{r}\mathbf{r}'} > 0$) at small distances $\|\mathbf{r} - \mathbf{r}'\|$ and inhibitory synapses ($g_{\mathbf{r}\mathbf{r}'} < 0$) at larger distances $\|\mathbf{r} - \mathbf{r}'\|$. It can be shown that the effect of such "center-surround" organisation of synaptic interactions on the solutions of (4.2) consists in the formation of excitatory responses that are confined to a neighborhood around the neuron receiving maximal external excitation. In the following, we will not prove this in general, but we would like to demonstrate it using a simplified version of (4.2).

To this end, we consider the limiting case when the "sigmoid function" $\sigma(x)$ approximates a step function $\theta(x)$ (as defined in Section 3.1). Further, we restrict ourselves to a one-dimensional system without an external input signal (*i.e.*, $v_l = 0$) and with thresholds $\theta = 0$. We assume for $g_{rr'}$ the function

$$g_{rr'} = \begin{cases} 1 & \text{if } |r - r'| \leq a, \\ -g & \text{else.} \end{cases} \tag{4.3}$$

Here, we assume $g > 2a + 1$, *i.e.*, neurons at distances exceeding a act inhibitory, while neurons closer than a act excitatory; the strength of the inhibition is given by the value of g. Defining the quantities

$$M = \sum_r f_r, \tag{4.4}$$

$$m_s = \sum_{r=s-a}^{s+a} f_r, \tag{4.5}$$

we see that (4.2) becomes

$$f_r = \theta\Big([1+g] \sum_{r'=r-a}^{r+a} f_{r'} - g \sum_{r'} f_{r'} \Big), \tag{4.6}$$

or, by using the property $\theta(gx) = \theta(x)$ which holds for $g > 0$,

$$f_r = \theta\Big([1 + g^{-1}]m_r - M\Big). \tag{4.7}$$

Because of the θ-function, every neuron can be in only one of the two states $f_r = 0$ or $f_r = 1$. Equation (4.7), together with (4.4) and (4.5) represents a system of equations for the neuron activities f_r. We now show that, as a consequence of the "center-surround" organization of the lateral interactions, (4.7) only has solutions in which the total excitation is concentrated within a single, connected "cluster" of $a + 1$ consecutive neurons with $f_r = 1$. All of the neurons outside of this cluster are in the quiescent state ($f_r = 0$). To this end, we first prove the following lemma: *Lemma:* If the quantities f_r constitute a solution of (4.7), and if $g > 2a + 1$, then $f_r = 1$ always implies $f_s = 0$ for all $s > r + a$ and all $s < r - a$. *Proof:* From (4.7) it follows because of $f_r = 1$ that the inequality $m_r + g^{-1}m_r > M$ is satisfied. From the definitions (4.4) and (4.5) one also has $m_r \leq M$, and together

$$m_r \quad \leq \quad M < m_r + \frac{m_r}{g} \leq m_r + \frac{2a+1}{g} < m_r + 1.$$

Since M and all the m_r are integers, one has $M = m_r$ and, thus, the lemma is proven.
The lemma implies that two active neurons r, s can never be located more than a positions apart ($|r-s| \leq a$). From this, it follows that $M \leq a+1$, *i.e.*, at most $a+1$ neurons can be excited at the same time. If s is the leftmost of these neurons, then it follows for each of the a neurons $r \in [s, s+a]$ adjacent to s on the right

$$\begin{aligned} [1 + g^{-1}]m_r - M &= [1 + g^{-1}] \sum_{r'=r-a}^{r+a} f_{r'} - M \\ &= [1 + g^{-1}] \sum_{r'=s-a}^{s+a} f_{r'} - M \\ &= [1 + g^{-1}]m_s - M > 0. \end{aligned} \tag{4.8}$$

Here, the shift of the limits of summation in the next to last step is based on the vanishing of all the $f_{r'}$ for $r' < s$ and $r' > s + a$. For each of the $a + 1$ neurons $r = s, s + 1, \ldots, s + a$, (4.8) yields then $f_r = 1$, and since $M \leq a + 1$

all the remaining neurons satisfy $f_r = 0$. Every solution of (4.7) therefore consists of a cluster of $a + 1$ adjacent excited neurons.
Similarly, in higher dimension, a sufficiently strong lateral inhibition also leads to the production of a spatially localized excitatory response. In the case of a continuous sigmoid function $\sigma(.)$, the spatial behavior of the excitation is no longer that of a step function, but rather takes a maximum at a position $\mathbf{r}'$ and from there decreases to zero in all directions. The location $\mathbf{r}'$ of the excitatory center is dependent on the input signal v_l (not taken into account in the above derivation). We pay special attention to this position $\mathbf{r}'$, since by mapping every input signal to a position $\mathbf{r}'$, the layer provides the desired map of the space of input signals. One could obtain $\mathbf{r}'$ by solving the nonlinear system of equations (4.2). Instead of this tedious step, Kohonen suggests an approximation for $\mathbf{r}'$, replacing it with the position of maximum excitation *on the basis of the external signal* v_l *alone*, *i.e.*, $\mathbf{r}'$ is determined from

$$\sum_l w_{\mathbf{r}'l} v_l = \max_{\mathbf{r}} \sum_l w_{\mathbf{r}l} v_l. \tag{4.9}$$

Under the two assumptions that the "total synaptic strength" per neuron $\sqrt{\sum_l w_{\mathbf{r}l}^2}$, is constant and the same for every neuron, and that all of the input signals $\mathbf{v}$ have the same "intensity" $\|\mathbf{v}\| = 1$, the condition

$$\|\mathbf{w}_{\mathbf{r}'} - \mathbf{v}\| = \min_{\mathbf{r}} \|\mathbf{w}_{\mathbf{r}} - \mathbf{v}\|, \tag{4.10}$$

which often is more convenient from a mathematical point of view, yields the same result for $\mathbf{r}'$. Here, $\|\mathbf{x}\|$ indicates the Euclidean vector norm $\sqrt{\sum_l x_l^2}$, and vector $\mathbf{w}_{\mathbf{r}} \equiv (w_{\mathbf{r}1}, \ldots, w_{\mathbf{r}d})^T$ is a compact notation for the synaptic strengths of neuron $\mathbf{r}$.
Thus, we now see how the map is related to the synaptic strengths $w_{\mathbf{r}l}$. An input signal $\mathbf{v}$ is mapped to the position $\mathbf{r}'$ implicitly defined by (4.10). For fixed synaptic strengths, (4.10) defines a nonlinear projection of the space of input signals onto the two-dimensional layer. In the following, we will use the notation

$$\phi_{\mathbf{w}} : \mathbf{v} \mapsto \mathbf{r}' = \phi_{\mathbf{w}}(\mathbf{v}) \tag{4.11}$$

to refer to this mapping. The index $\mathbf{w}$ shall remind us of the mapping's dependence on the synaptic strengths of all neurons.
This leads to the second important issue, the determination of synaptic strengths $\mathbf{w}$ providing "useful" maps. In the nervous systems of higher animals, a detailed genetic specification of all synaptic strengths is not possible.

This specification would require an exact knowledge of the way input signals are coded, a condition which even for technical applications, for example due to tolerances, is difficult to satisfy. Moreover, a system with fixed values $w_{\mathbf{r}l}$ could not respond to subsequent changes of the coding, *e.g.*, due to drift or aging processes; this obviously would contradict the high capacity for adaptation of biological systems. Apparently, such flexibility requires that the neurons be able to find suitable synaptic strengths, starting from arbitrary or only roughly correct initial settings.

In the present model, the only source of information for this process is assumed to be a sequence of input stimuli entering the layer, occurring randomly according to some statistical probability distribution. Each stimulus causes at synapse $w_{\mathbf{r}l}$ the coincidence of a presynaptic activity v_l and the resulting postsynaptic activity of neuron $\mathbf{r}$. The postsynaptic activity of neuron $\mathbf{r}$ is just the value of the excitatory response of the layer at the position $\mathbf{r}$. Its magnitude includes all interaction effects within the layer and should be computed from (4.2). Kohonen's model now makes the simplifying assumption that this response can be written as a function $h_{\mathbf{r}\mathbf{r}'}$ of two position variables $\mathbf{r}$ and $\mathbf{r}'$, whose "shape" (with respect to variation of $\mathbf{r}$) is fixed, but whose position (denoted by the second variable $\mathbf{r}'$) depends on the stimulus. Specifically, the position $\mathbf{r}'$ is taken to be the position of the excitation maximum, *i.e.*, $\mathbf{r}'$ is defined by (4.9) or (4.10), and $\mathbf{r}$ is the location of the neurons whose response is to be described by $h_{\mathbf{r}\mathbf{r}'}$. The model then prescribes for the change of synaptic strengths $w_{\mathbf{r}l}$ the expression

$$\Delta w_{\mathbf{r}l} = \epsilon(h_{\mathbf{r}\mathbf{r}'} v_l - h_{\mathbf{r}\mathbf{r}'} w_{\mathbf{r}l}). \tag{4.12}$$

The first term corresponds to the "Hebbian learning rule" mentioned earlier, according to which a synapse is strengthened in the case of correlated pre- and postsynaptic activity. The second term is a decay term for the synaptic strengths, which is proportional to the postsynaptic activity. The relative scaling between the first term and the second (decay) term is normalized to unity by appropriate scaling of $\mathbf{v}$. Here, ϵ determines the size of a single adaptation step ($0 < \epsilon < 1$). If ϵ is chosen to be a function $\epsilon(t)$, decreasing gradually with the number t of learning steps from large initial values to small final values, then at the beginning the system is rapidly able to learn coarsely the correct synaptic strengths. However, for large ϵ, the fluctuation of the map caused by each learning step is also large. Hence, if the map is to stabilize asymptotically in an equilibrium state, one must let ϵ decrease to zero. On the other hand, a permanent "residual plasticity" can be realized

with low fluctuations of the map by means of a small, nonvanishing final value for ϵ.
Based on (4.12), every synaptic change is limited to a neighborhood zone about the excitation center. In this zone, the synaptic connections are changed such that a subsequent re-occurrence of the same or a similar stimulus will lead to an increased excitation. The shape of the function $h_{\mathbf{rr}'}$ controls the size of the neighborhood zone and, thus, of the number of neurons affected by a single adaptation step.

4.2 Simplification and Mathematical Definition

The precise form of the excitatory response $h_{\mathbf{rr}'}$ appears not to be critical for the qualitative behavior of the system under the learning rule (4.12) and could only be obtained by numerical solution of (4.2). Hence, in the present model, the exact solution is only approximated qualitatively by means of a given choice of $h_{\mathbf{rr}'}$. To this end, for $h_{\mathbf{rr}'} \geq 0$ a unimodal function depending only on the distance $\mathbf{r} - \mathbf{r}'$ with its maximum at $\mathbf{r} = \mathbf{r}'$ and approaching zero for large distances is assumed. An appropriate choice is given by the Gaussian

$$h_{\mathbf{rr}'} = \exp(-(\mathbf{r} - \mathbf{r}')^2/2\sigma_E^2). \tag{4.13}$$

The radius σ_E of this excitatory function determines the length scale on which the input stimuli cause corrections to the map. As a rule, it is better if the coarse structure of the map is allowed to form first, before the fine structure is incorporated into the map. This is made possible by choosing σ to be a function $\sigma(t)$ starting with a rather large initial value $\sigma(0)$ and decreasing slowly with the number of learning steps toward a small final value. This can be interpreted as gradually increasing the "selectivity" of the individual neurons in the course of the learning process.
Each learning step requires the arrival of an input stimulus $\mathbf{v}$. For the model, these input stimuli are treated as independent random variables from a vector space V, and their occurrence is determined by a probability density $P(\mathbf{v})$.
A final simplification is that the neuron positions $\mathbf{r}$ are taken to be the points of a discrete periodic lattice A.
Thus, Kohonen's model can be described by the following algorithm (Kohonen 1982a, 1984a):

1. *Initialization*: Start with appropriate initial values for the synaptic strengths $w_{\mathbf{r}l}$. In the absence of any a priori information, the $w_{\mathbf{r}l}$ can be chosen at random.

2. *Choice of Stimulus*: Choose, according to the probability density $P(\mathbf{v})$, a random vector $\mathbf{v}$ representing a "sensory signal."

3. *Response*: Determine the corresponding "excitation center" $\mathbf{r}'$ from the condition
$$\|\mathbf{v} - \mathbf{w}_{\mathbf{r}'}\| \leq \|\mathbf{v} - \mathbf{w}_{\mathbf{r}}\| \quad \text{for all } \mathbf{r} \in A. \tag{4.14}$$

4. *Adaptation Step*: Carry out a "learning step" by changing the synaptic strengths according to
$$\mathbf{w}_{\mathbf{r}}^{\text{new}} = \mathbf{w}_{\mathbf{r}}^{\text{old}} + \epsilon h_{\mathbf{r}\mathbf{r}'}(\mathbf{v} - \mathbf{w}_{\mathbf{r}}^{\text{old}}) \tag{4.15}$$
and continue with step 1.

The mapping
$$\phi_{\mathbf{w}} : V \mapsto A, \quad \mathbf{v} \in V \mapsto \phi_{\mathbf{w}}(\mathbf{v}) \in A, \tag{4.16}$$
where $\phi_{\mathbf{w}}(\mathbf{v})$ is defined through the condition
$$\|\mathbf{w}_{\phi_{\mathbf{w}}(\mathbf{v})} - \mathbf{v}\| = \min_{\mathbf{r} \in A} \|\mathbf{w}_{\mathbf{r}} - \mathbf{v}\| \tag{4.17}$$
which constitutes the *neural map* of the input signal space V onto the lattice A which is formed as a consequence of iterating steps 1.–3.
To illustrate this algorithm, the relationships are schematically shown again in Fig. 4.2. The ensemble of all possible input values forms the shaded manifold V, from which a point $\mathbf{v}$ is chosen as "stimulus" for the network in step 1. This leads to a selection (step 2) of an excitation center $\mathbf{s}$ among the neurons (lattice A). All neurons in the neighborhood of this center (highlighted) participate in the subsequent adaptation (step 3). It consists in a "shift" of the vectors $\mathbf{w}_{\mathbf{r}}$ towards $\mathbf{v}$. The magnitude of this shift is fixed by the learning step size ϵ and by the function $h_{\mathbf{r}\mathbf{s}}$.
Mathematically, the algorithm represents a so-called *Markov process*. A Markov process is defined by a set of states and a set of transition probabilities between states. These transition probabilities determine a stochastic process that, given some initial state, produces a sequence of states. This sequence is obtained by using the transition probabilities from the current state to

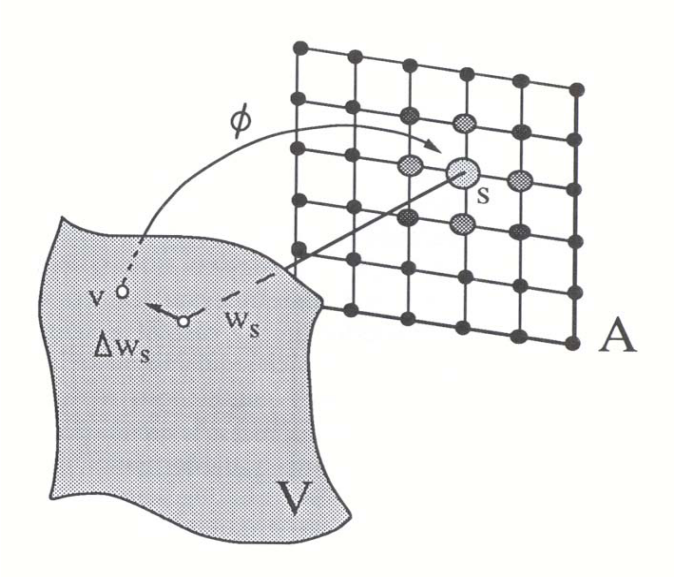

Abb. 4.2: The adaptation step in Kohonen's model. The input value $\mathbf{v}$ selects a center $\mathbf{s}$ in whose neighborhood all neurons shift their weight vectors $\mathbf{w_s}$ towards the input $\mathbf{v}$. The magnitude of the shift decreases as the distance of a unit from the center $\mathbf{s}$ increases. In the figure, this magnitude is indicated by different sizes and gray values. The shift of weights is only depicted, though, for unit $\mathbf{s}$

choose a successor, which then becomes the current state for the next step (for a thorough discussion of Markov processes see for example Gardiner 1985 or van Kampen 1981).

In the present model, each possible state is given by a set of values for all the synaptic strengths $\mathbf{w} \equiv (\mathbf{w}_{\mathbf{r}_1}, \mathbf{w}_{\mathbf{r}_2}, \ldots, \mathbf{w}_{\mathbf{r}_N})$ in the system (N denotes the number of neurons). The function $\phi_{\mathbf{w}}$ associates with each such state a mapping that, as we have discussed, has the interpretation of a "neural map" of some feature space. The update of a state $\mathbf{w}$ is obtained as a result of applying (4.15), *i.e.*, the decision for the update is caused by the input stimulus $\mathbf{v} \in V$. Each update represents a "learning step" and can be thought of as a local "distortion" of the associated "neural map." Beginning with an initial state that corresponds to a completely disordered map, the goal of the algorithm is to arrive at a state (more precisely, the system shall enter a subset of its state space comprising states differing only by small "statistical

fluctuations", see Chapter 14) that corresponds to an ordered, "topology-conserving map" of the stimulus space V, in which some relevant features of input stimuli are two-dimensionally (in the case of a neural sheet) represented. In order to reach such state and make it stationary asymptotically, the learning step length ϵ must slowly tend to zero.
The training process is qualitatively in good agreement with observed features of the formation of certain neural projections in the brain. The resulting maps predominantly represent those directions of the stimulus space V along which the input stimuli change most strongly. These directions, which often correspond to stimulus features of particular interest, may vary locally within V. Therefore, a good projection requires a nonlinear mapping. Usually, the map tries to maintain the neighborhood relationships between the input stimuli under this mapping process. Therefore, Kohonen named the resulting maps "topology-conserving feature maps." Furthermore, the map automatically takes into account the statistical weight $P(\mathbf{v})$ of the input stimuli. Regions of V from which many input stimuli occur become "magnified" and are thus projected with better resolution than regions of less frequently occurring signals. An appropriate choice for the rate of decrease of ϵ and σ with the number of learning steps is important for good results and rapid convergence. If the decrease is too rapid, the synaptic strengths "freeze" before the map has reached an equilibrium state. If the decrease is too slow, the process takes longer than necessary.
To illustrate the basic properties of this approach, we now consider a few simulation examples of the process.

4.3 Simulation Examples

In the first example, a neural network creates a map or image of an unknown region G with curved boundary. Only indirect sensory signals are available to the network. These come from a source of sound moving around in G. From time to time, the sound source emits a sound signal (of constant intensity), and the position in G of each sound emission is random. The sound signal is received by two microphones, each connected to an amplifier with logarithmic characteristics (Fig. 4.3). The two amplifier output signals v_1, v_2 are the "sensory signals," and they are fed via two "axons" to the 1600 "neurons"

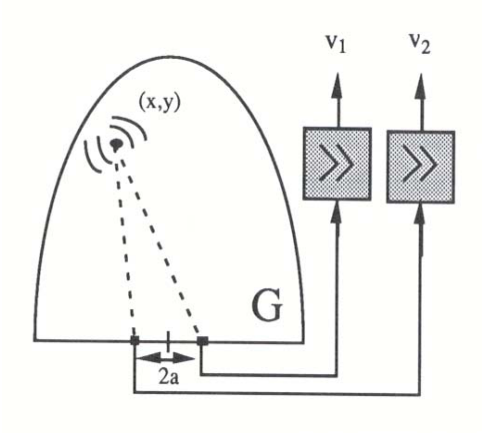

Abb. 4.3: Region G containing the sound source. The two microphone positions are marked at the lower boundary of G. The microphone signals are fed into two logarithmic amplifiers, whose output signals v_1, v_2 serve as input for the network.

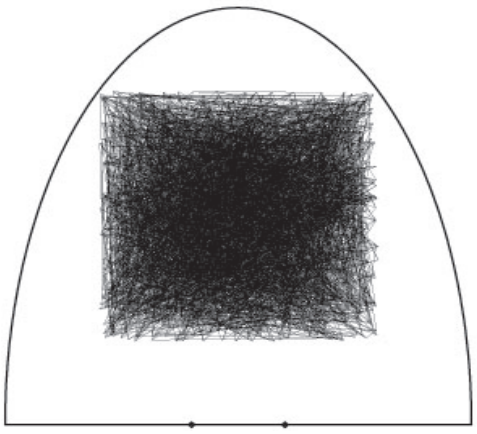

Abb. 4.4: Initial relation between neurons and points in G. Initially, each neuron is assigned to a point of G chosen randomly from the filled quadrant. This assignment ignores any neighborhood relations. This is evident from the completely irregular "embedding" of the lattice in the quadrant.

of a model network.[1] The "neurons" are arranged in a planar 40×40 lattice. Every single model neuron $\mathbf{r}$ is characterized by a two-component vector $\mathbf{w_r} = (w_{\mathbf{r}1}, w_{\mathbf{r}2}) \in G$ of "synaptic strengths." Each neuron is to adjust its vector $\mathbf{w_r}$ gradually in such a way as to become sensitive for a small subset of input signals $\mathbf{v} = (v_1, v_2)^T$. This subset corresponds to a small subarea of G within which the moving source may be located. This subarea constitutes the "receptive field" of the particular neuron in the "environment" G. The

[1] In the computer simulation, sound source, microphone, and amplifier are represented as follows: if the sound source is at the position (x, y), the output signals v_1 and v_2 of the two amplifiers are given by

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -\log[(x-a)^2 + y^2] \\ -\log[(x+a)^2 + y^2] \end{pmatrix}, \tag{4.18}$$

where $2a$ is the separation of the microphones.

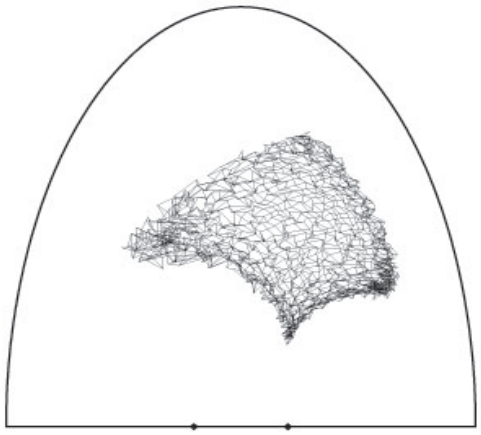

Abb. 4.5: After 100 learning steps, an assignment has already formed which roughly reproduces the neighborhood relations of points of G in the lattice. However, the distribution of "responsibilities" of neurons for the region G is still very inhomogeneous.

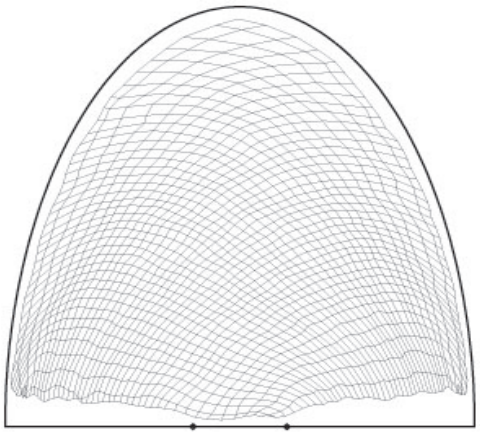

Abb. 4.6: After 40,000 learning steps, a good correspondence between lattice neurons and points of G has formed. This corresponds to the choice of curvilinear coordinates, mapping the region G onto the square neuron lattice.

neurons are to coordinate the formation of their receptive fields in such a way that — in the manner of a topographic map — the arrangement of neurons in the lattice reflects the arrangement of their respective receptive fields in the environment. This is achieved if each point of the region G corresponds to a point in the neural lattice such that the neighborhood relation between points is preserved under the correspondence, *i.e.*, the network becomes associated with a "continuous" image of G. This correspondence gives a simple example of a sensory map or sensory image of an environment, here the region in front of the two microphones. Similar "auditive maps" occur in the brain. However, this simulation example is only intended to serve as an illustration of the algorithm and makes no claim of corresponding to any brain map.
In Figs. 4.4–4.6, the evolution of the assignment of neurons to positions is shown in detail. For each neuron $\mathbf{r} \in A$, the location (x, y) of its receptive field in G has been marked, as assigned by the map. Marked locations are connected by a line if their corresponding neurons are adjacent on the lattice. (Thus, in place of the image itself, the embedding of the lattice A

in G is shown, from which the map can be obtained as its inverse.) Initially, the assignment is completely random, and there is no agreement between the arrangement of neurons and the corresponding locations (Fig. 4.4). After only a few signals, the coarse structure of the assignment has been found (Fig. 4.5), until finally after 40,000 sound signals a good assignment is achieved (Fig. 4.6). In this case, the algorithm has automatically found a nonlinear coordinate transformation mapping the region G with curved boundary onto a square lattice A. The resulting coordinate transformation takes the frequency distribution of the arriving signals into account, as illustrated in the simulation result shown in Fig. 4.7. Instead of a homogeneous distribution of source locations, the signals from the indicated circular region in G were now emitted with a three times higher probability than in the remaining part of G. Within both regions the probability density was constant. In all other respects the simulation was identical to that presented in Fig. 4.4–4.6. As a consequence of the inhomogeneous stimulus distribution, substantially more neurons are assigned to positions in the circular region. This corresponds to a higher resolution of the map for this part of G, which is a desirable result, since a concentration of assignments within regions where signals frequently occur leads to a more efficient use of the network.

However, the frequency with which a signal occurs is not always an indication of its importance. Varying importance of signals can also be taken into account by regulating the plasticity of the network. For example, one can adjust the size of a learning step according to an a priori importance attributed to the signals. This increases the "attentiveness" of the network for signals deemed more important and has the same effect as correspondingly more frequent occurrence. This is illustrated in Fig. 4.8, which shows the result of a simulation with sound emission probability again uniform throughout all of G. However, in contrast to Fig. 4.4–4.6, the network reacted to every sound event from within the circle with an adaptation step that was three times larger than for a sound event from the remaining part of G. The result thus obtained is practically identical to that of Fig. 4.7.

In the example presented, the space of stimuli G is mapped onto a lattice A of the same dimensionality. If the space of stimuli possesses a higher dimensionality, the map tries to project the higher-dimensional space as faithfully as possible by means of an appropriate "convolution." To illustrate this behavior, we consider a one-dimensional neural "net," *i.e.*, a neuron chain. For the input signal, we take a random sequence of two-dimensional vectors $\mathbf{v}$, whose values are homogeneously distributed in the unit square. For $h_{\mathbf{r}\mathbf{r}'}$, we

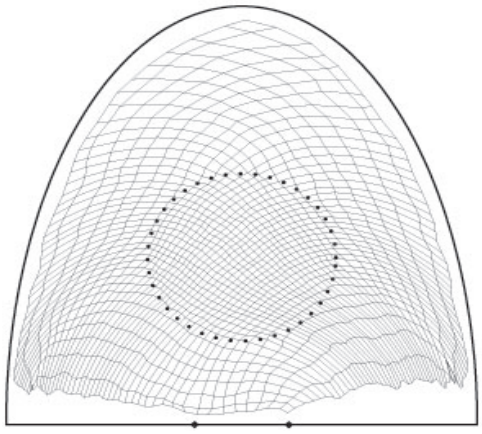

Abb. 4.7: Result of the same simulation as in Fig. 4.6, except that within the circular region marked by dots signals were emitted with a three times higher probability than in the remaining region of G. In this case, more neurons code positions in the circular region. This corresponds to a higher resolution of the map created for this region.

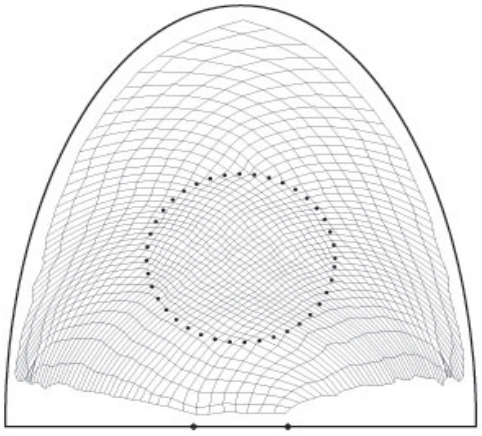

Abb. 4.8: The same effect as in Fig. 4.7 can be achieved by a signal-dependent adjustment of the plasticity of the neurons. In this simulation, the sound signals were again emitted as in Fig. 4.4–4.6 with a homogeneous probability everywhere in G, but the learning step size ϵ was increased by a factor of three if the sound source was located in the circular region.

choose the Gaussian (4.13) with $\sigma(t) = 100 \cdot (0.01)^{10^{-5}t}$. The correspondence between neurons and points of the square is again represented as an embedding of the neuron chain into the square, as in the previous example. This assignment is initially made at random as shown in Fig. 4.9a. After 200 iterations, the curve has attained a U-shaped configuration (Fig. 4.9b). At this time, the range σ of the function $h_{\mathbf{rr}'}$ is still large and, hence, structure has formed only at this length scale. As σ decreases further, structures gradually form at shorter length scales as well (Fig. 4.9c, 50,000 iterations). Eventually, after 100,000 iteration steps, the hierarchically convoluted graph of Fig. 4.9d has emerged. The network thus tries to fill the two-dimensional region while reproducing the neighborhood relations as well as possible. The degree of success is evident from the similarity of the curve created in this way to the

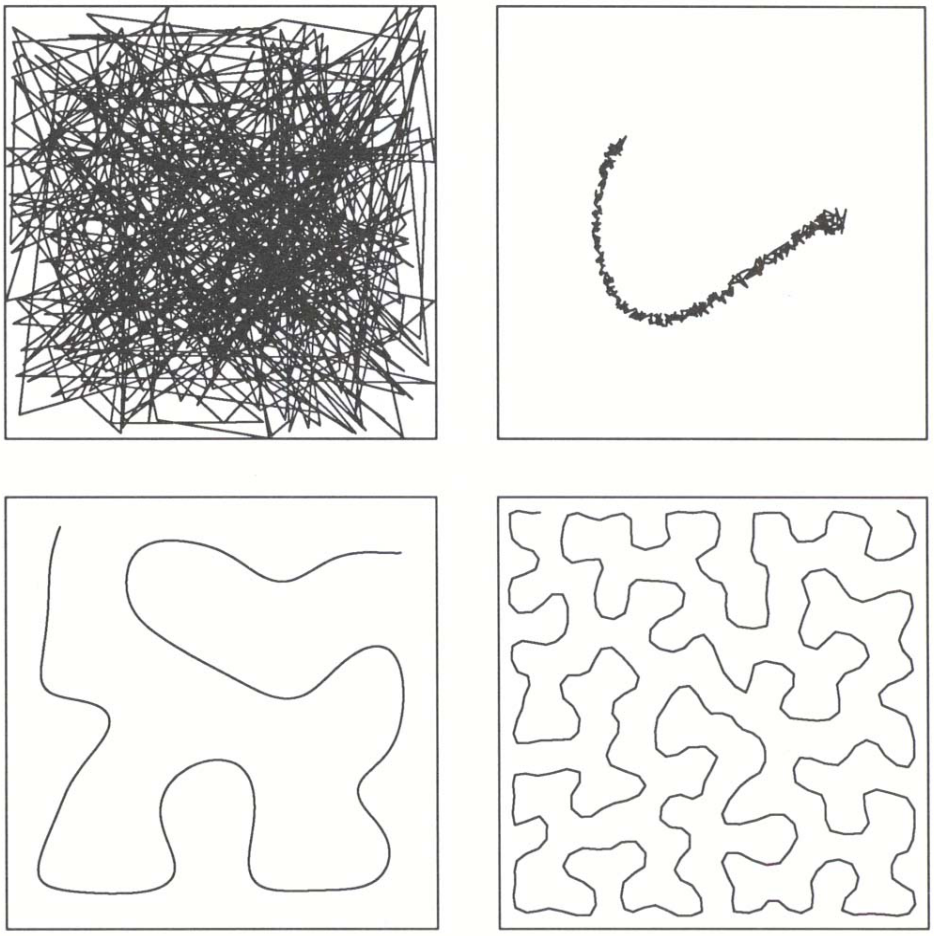

Abb. 4.9: Mapping between a "neural" chain and a squared stimulus space. From top left to bottom right: a) randomly chosen initial assignment; b) coarse assignment after 200 Markov steps; c) after 50,000 Markov steps; d) assignment obtained after 100,000 Markov steps resembling a "Peano curve."

finite approximation of a so-called "Peano curve." This is an infinitely, recursively convoluted fractal curve representing the solution of the problem of mapping a one-dimensional interval continuously onto a two-dimensional surface.
However, as a rule one is interested in mapping of higher-dimensional regions

onto a two-dimensional image. Indeed, Kohonen used the procedure successfully to map spectra of different speech sounds (phonemes) to separate map positions. Here, the tonal similarity relations between the individual phonemes are translated into locational relations in the image. This constitutes a very important preprocessing step for the problem of artificial speech recognition. The subsequent steps require the analysis of transitions between individual phonemes, *i.e.*, of time sequences. The possibility of employing the procedure also for such purposes shall be indicated in the following concluding example. At the same time, this example will clarify how in the course of the formation of a map hierarchic relations can also be represented. The source of the signal is a Markov process (here used as a simple model of a temporal signal and to be distinguished from the learning algorithm itself) with 10 states. The aim is to create a map of the possible transitions between states of the process. Transitions to the same successor state are to be adjacent in the map. A state i, $i = 0, \ldots, 9$, is assumed to have one of the five states $i-3$, $i-2$, $i-1$, $i+1$ or $i+2$ (modulo 10) as a possible successor. A transition from state i to state j is coded by a 20-component vector $\mathbf{v}$ with components $v_k = \delta_{k,i} + \delta_{k,j+10}$. A transition occurs at each time step, and all transition probabilities have the same value 0.2. A lattice consisting of 20×20 neurons is used, and the Gaussian (4.13) is chosen for $h_{\mathbf{rr}'}$. The remaining parameter values of the simulation are $\sigma(t) = 5 \cdot 0.2^{t/t_{max}}$, $\epsilon(t) = 0.9 \cdot (0.05/0.9)^{t/t_{max}}$ and $t_{max} = 5,000$ learning steps. Additionally, for the computation of the distances $\|\mathbf{v} - \mathbf{w}(\mathbf{r})\|$, a "metric" was used which weights the differences in the last 10 components of $\mathbf{v}$ twice as strongly as those of the first 10 components. In Fig. 4.10, the 20×20-lattice of neurons is represented. For each lattice site, two numbers $i, j \in \{0, \ldots, 9\}$ indicate the initial and final state of the transition assigned to the respective neuron. The initial distribution was again chosen randomly. Figure 4.10 shows the map obtained after 5,000 learning steps. For each of the 50 allowed transitions, an "island" of neurons responding to this transition has formed, and the islands are in turn arranged in such a way that islands corresponding to transitions to the same successor state form a larger cluster. This corresponds to a hierarchical arrangement and is a consequence of the described choice of weight, the successor state obtaining a higher weight than the predecessor state in the choice of the excitation center. This choice dominates the formation of the "large-scale" structure of the map, *i.e.*, the structure on the level of "clusters of islands." This illustrates that, by an appropriate choice of metric (the choice of weight corresponds to a choice of metric), it is possible to arrange

for certain features (here successors) to be grouped together hierarchically in the map.

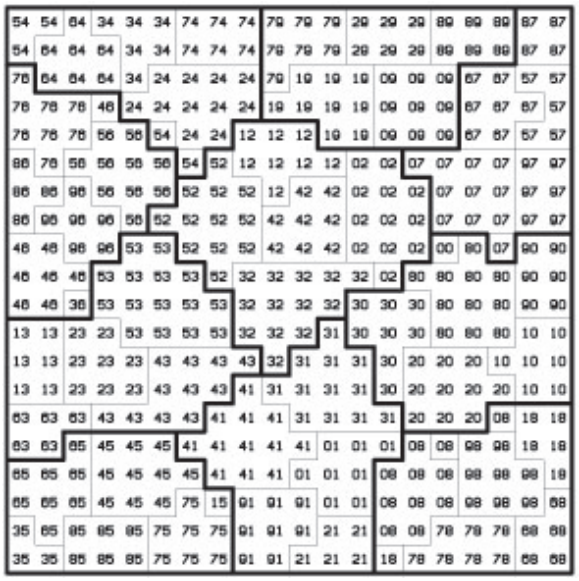

Abb. 4.10: Mapping of the transitions $i \rightarrow j$ of a Markov process with states $i, j = 0, ..., 9$ onto a lattice consisting of 20×20 neurons. For each lattice location, the transition to which the corresponding neuron best responds is indicated as jk. Neurons with the same transition are adjacent to one another within islands. Islands with the same successor in turn form "clusters." This corresponds to a hierarchical distribution of the neuron specificities over the lattice.

By the inclusion of *contextual information*, such a hierarchical grouping can emerge from the data itself. For example one can create "semantic maps" which arrange words in hierarchies of meaning. This ordering is gradually found by the system itself in the course of a learning phase, where simple English sentences can serve as "training data" (Ritter and Kohonen 1989). After this initial overview, we consider in the following chapters a series of information processing tasks, for which the choice is motivated by their significance for biological systems. At the same time, we investigate how self-organizing maps can be useful in solving such problems. While viewing biological examples as a guide, we will occasionally consider technical applications when appropriate. This applies particularly to Chapter 6, which

gives a solution to the "traveling salesman problem" and Chapters 10–13, which are concerned with applications to robotics.