ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

Self-Organisation, Metastable States and the ODE Method in the Kohonen Neural Network

John A. Flanagan and Martin Hasler
Department of Electrical Engineering
Swiss Federal Institute of Technology
CH-1015 Lausanne, Switzerland

Abstract - Metastable states are analysed for a simple Kohonen Neural Network (KNN) using the ODE method. Normally the ODEs are determined for each possible configuration of the neuron weights. By comparing the trajectories of the ODEs of the neuron weights and those obtained from simulation we are forced to rethink how the ODEs should be formed for the KNN before a complete analysis of the self-organisation process in the KNN can be analysed with the ODE method. It is shown that the ODEs predict better what happens in practise if the general ODEs are formed by using a weighted average of the ODEs for each possible configuration of the neuron weights. The weight accorded to each configuration depends amongst other things on the type of gain function used.

1 Introduction

The KNN [1] is a biologically inspired algorithm which models the biological phenomena of retinotopy, or the formation of self-organised maps between the retina in the eye and the cortex in the brain. In general self-organisation implies that adjacent inputs to the retina are coded to adjacent regions of the cortex.

In one dimension the process is modelled as follows ; each neuron has a neuron weight X_i with $1 \leq i \leq N$. First the neuron weights are randomly initialised and then a series of random inputs ω are presented to the network. At iteration t a *winner* neuron v is chosen such that ,

$$| X_v(t) - \omega(t) | \leq | X_k(t) - \omega(t) | \quad \forall\, k \tag{1}$$

Each neuron weight is then updated as,

$$X_k(t+1) \;=\; X_k(t) \;+\; \alpha(t)h(v,k)(X_k(t) - \omega(t)) \qquad \forall\, k \tag{2}$$

The gain function $\alpha(t) \in (0,1)$ and normally $\alpha(t) \to 0,\ t \to \infty$. The function $h(v,k)$ is referred to as the neighbourhood function, and normally is one for $v = k$

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

and decreases as $| v-k |$ increases. If the algorithm is allowed continue for many inputs the neuron weights will organise, in the one dimensional case this corresponds to the neuron weights being in one of the two following *configurations* $X_1 < X_2 < \ldots < X_N$ or $X_1 > X_2 > \ldots > X_N$ [2]. In general a configuration is denoted by C_k and defined as being one particular *ordering* of the neuron weights. Sometimes if $h(v,k)$ is badly chosen the neuron weights can become trapped in disorganised configurations. This effect which is highly undesirable was first suggested by Ritter and Schulten [3] to be a result of the existence of *metastable* stationary states in disorganised configurations. Tolat [4] has analysed this situation using a system of energy functions. He finds that if certain conditions are satisfied by the neighbourhood function then these disorganised metastable states can be avoided. Erwin *et al* [5] have also analysed the problem of metastable states using a set of ordinary differential equations. They proved that for the one dimensional case when the distribution of the input signal is uniform that if $h(k,v)$ is *convex* then the only stationary states are in organised configurations. They also give extensive simulation results which back up their theoretical work, and show how the existence of disorganised metastable states, even if they do not trap the weights in disorganised configurations, considerably slow down the organisation phase of the neuron weights.

In this work metastable states are also analysed using the Ordinary Differential Equation (ODE) method, a technique for analysing the convergence properties of stochastic processes [6], [7]. This method has already been applied to the analysis of the final convergence phase of the KNN in [8], [9]. However here we are interested in how the ODE method can be applied to the KNN algorithm for the analysis of self-organisation. In section 2 a brief description of the ODE applied to the KNN algorithm is described and in section 3 an example of a particular KNN is described. By comparing trajectories of the neuron weights obtained from numerical solutions of the ODEs to the trajectories obtained from simulation, we arrive at an understanding as to how the ODE method should be applied to the KNN to analyse self-organisation.

2 The ODE Method and the KNN

With each neuron weight there is an associated ODE given by the average of the update $h(k,v)(X_k(t)-\omega(t))$ over all possible winner neurons v. The average is normally performed for one configuration of the neuron weights [5]. The resultant ODE is of the form,

$$\frac{dx_k}{d\tau} = \sum_{v=1}^{N} h(v,k) \int_{\Omega_v} (x_k - \omega) d\mu(\omega) \qquad \forall\, k \tag{3}$$

where $\mu(\omega)$ is the probability distribution of ω, and Ω_v the Voronoi Tessellation cell of the winner neuron v. Stable stationary points $\mathbf{x}_\infty$ of these sets of ODEs are possible stationary points [6] of the stochastic process $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}$ represented by the neuron weights. A working definition of a metastable state is now given as,

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

Definition 1 *A 'metastable' state in configuration $\mathcal{C}_k$ is a stationary point of the ODE's relative to $\mathcal{C}_k$, which lies in $\mathcal{C}_k$, and is considered a disorganised state if $\mathcal{C}_k$ is not an organised configuration.*

3 An Example

We now analyse metastable states in a one dimensional KNN with a uniform probability distribution for the input and $N = 3$. The neighbourhood function is given by $h(0) = 1$, $h(1) = a$ and $h(k) = 0\ \forall k > 1$. Given this KNN there are 6 possible configurations of the neurons $\mathcal{C}_1, \ldots, \mathcal{C}_6$.

$$\begin{aligned}
\mathcal{C}_1 &= \{(x_1, x_2, x_3) : 0 \le x_1 \le x_2 \le x_3 \le 1\} \\
\mathcal{C}_2 &= \{(x_1, x_2, x_3) : 0 \le x_1 \le x_3 \le x_2 \le 1\} \\
\mathcal{C}_3 &= \{(x_1, x_2, x_3) : 0 \le x_2 \le x_1 \le x_3 \le 1\} \\
\mathcal{C}_4 &= \{(x_1, x_2, x_3) : 0 \le x_3 \le x_2 \le x_1 \le 1\} \\
\mathcal{C}_5 &= \{(x_1, x_2, x_3) : 0 \le x_3 \le x_1 \le x_2 \le 1\} \\
\mathcal{C}_6 &= \{(x_1, x_2, x_3) : 0 \le x_2 \le x_3 \le x_1 \le 1\}
\end{aligned} \tag{4}$$

Some known facts about this simple setup include, $\mathcal{C}_1$ and $\mathcal{C}_4$ are *absorbing*[2], that is if $\mathbf{x}(t_1) \in \mathcal{C}_1$ then $\mathbf{x}(t) \in \mathcal{C}_1\ \forall t > t_1$. Also, if $\mathbf{x}(t_1) \in \mathcal{C}_2$ then $\mathbf{x}(t) \in \mathcal{C}_2 \bigcup \mathcal{C}_1\ \forall t \ge t_1$. Similar relations exist for the pairs of configurations $(\mathcal{C}_3, \mathcal{C}_1)$, $(\mathcal{C}_4, \mathcal{C}_5)$, $(\mathcal{C}_4, \mathcal{C}_6)$. Writing the ODEs for configuration $\mathcal{C}_2$ gives,

$$\begin{aligned}
\frac{dx_1}{dt} &= -0.5[((x_1 + x_3)/2 - x_1)^2 - x_1^2 + a((1 - x_1)^2 - ((x_2 + x_3)/2 - x_1)^2)] \\
\frac{dx_2}{dt} &= -0.5[(a - 1)(x_1 + 2x_3 - 7x_2)(x_1 - x_2)/4 + (1 - x_2)^2 - ax_2^2] \\
\frac{dx_3}{dt} &= -0.5[(a - 1)((x_2 + x_3)/2 - x_3)^2 - ((x_1 + x_2)/2 - x_3)^2) + a(1 - x_3)^2]
\end{aligned} \tag{5}$$

Figure 1 shows a plot of $(x_{1\infty}, x_{2\infty}, x_{3\infty})$ the stationary state of the set of equations (5) for each value of a varied between 0 and 0.5. From the figure it is seen that for $a < 0.25$ that the stationary state $0 < x_{1\infty} < x_{3\infty} < x_{2\infty} < 1$ is a disorganised metastable state. For $a > 0.25$ the stationary state is the organised state $0 < x_{1\infty} < x_{2\infty} < x_{3\infty} < 1$.

If we write the ODEs for each of the 6 possible configurations of the three neurons and plot the solution to the equations (solved numerically) then the left diagram of figure 2 shows resultant trajectories for $a = 0.5$ projected onto the x_1, x_2 plane (i.e all the phase plots which follow are projections of three dimensional trajectories onto the x_1, x_2 plane).

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

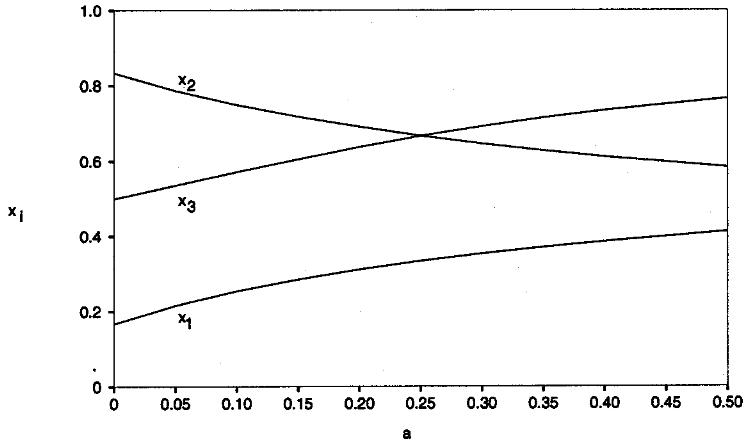


Figure 1: Solution of the ode for three neuron weights x_1, x_2, x_3 plotted against a.

There are two stable stationary states, one in each of C_1 and C_4. Next non-organised metastable states are introduced by letting $a = 0.15$, the right diagram of figure 2 shows a plot of x_2 v's x_1 for the solution of equations (5). There are 6 six stable stationary points one in each of the possible configurations.

Next the KNN was simulated using the parameters described above and $a = 0.5$. Starting from a given initial condition the algorithm was run for several thousand iterations, with the value of each neuron weight recorded at each iteration. The weights were initialised to the same values and the algorithm run again with a different input sequence (i.e. though still uniformly distributed). After several such iterations an ensemble average of the trajectories was obtained. This complete process was carried out for several different initial conditions. The result is a set of trajectories shown in figure 3.

Comparing this phaseplot with the left diagram of figure 2 it is possible to see how theory and practise produce a similar phaseplot.

However if the simulations are run for the case of $a = 0.15$ and a gain function $\alpha(t) = 300/(t+800)$ then the phase plot of the left diagram of figure 4 is obtained. In this phase plot the direction of the heavier trajectories is from r to p from t to p from s to q and v to q. Comparing this to the phase plot predicted by the ODE in figure 2 there is a distinct difference in that these dark trajectories are not predicted by the ODEs. The same procedure was tried again but this time with a different gain function $\alpha(t) = 75000/(t + 80000)$, which is initially larger and decreases less quickly with time than the previous case. The result is shown in the right diagram of figure 4.

The noticeable difference between the two diagrams in this figure is the difference between the heavier trajectories between the metastable states. In the former the trajectories are much closer while in the latter they have spread out

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

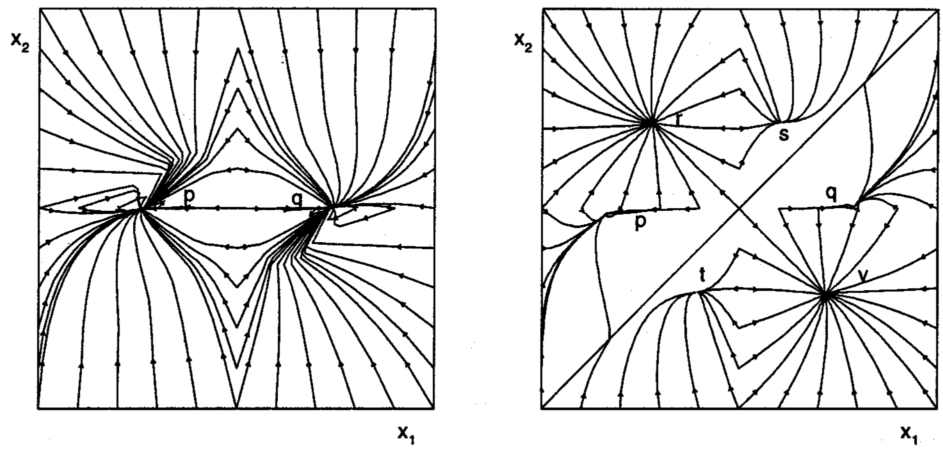


Figure 2: The left diagram shows a solution of the ODEs for $a = 0.5$, the right diagram is for $a = 0.15$. Metastable states are indicated by p, q, r, s, t, v.

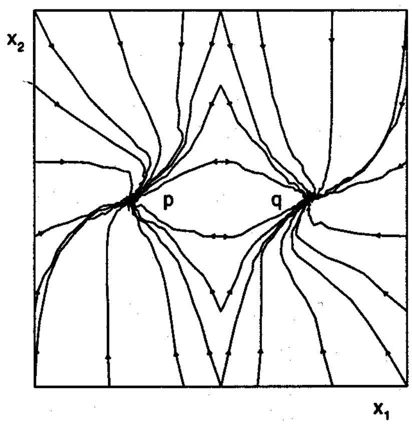


Figure 3: Phase plots of the averages of the neuron weights x_1, x_2 for $a = 0.5$. Metastable states of the ODEs are indicated by p, q.

ESANN'1995 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 19-20-21 April 1995, D-Facto public., ISBN 2-9600049-3-0, pp. 1-8

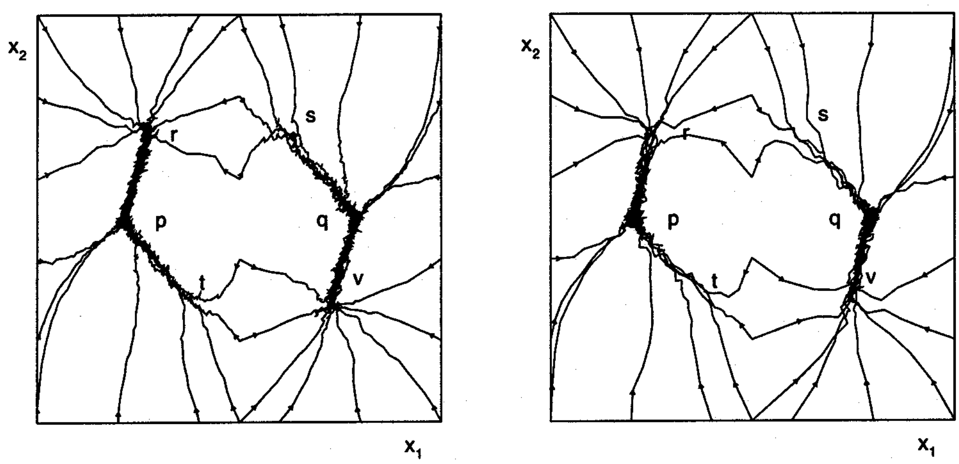
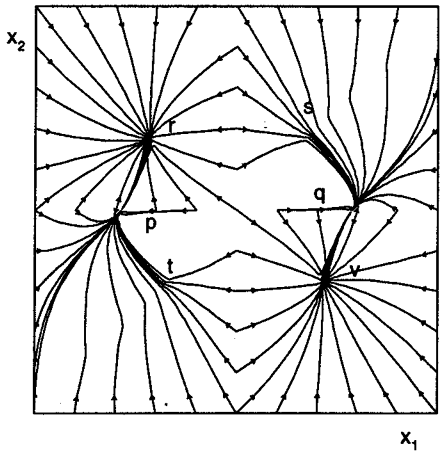


Figure 4: The left diagram is the average of the trajectories for $a = 0.15$, $\alpha(t) = 300/(t + 800)$. The right diagram for $a = 0.5$, $\alpha(t) = 75000/(80000 + t)$. Metastable states are indicated by p, q, r, s, t, v.

somewhat. They are both however different from the phase plot predicted by the ODE. So where is the problem, or, is there a problem? Consider the case $\mathbf{x}(0) \in \mathcal{C}_2$. It is noticeable from the phase plots determined from simulations that initially the trajectories go towards the point where the ODEs predict a metastable state in $\mathcal{C}_2$ (i.e. state v in figure 2). By running simulations however it is seen that, sometimes, for large enough t then $\mathbf{x}(t) \in \mathcal{C}_1$. As the gain function is made initially smaller and decreases more quickly then the longer the process spends at the disorganised metastable state and the longer it takes for $\mathbf{x}(t)$ to enter $\mathcal{C}_1$, if it does at all. Of course if the process is started several times from the same initial condition then it will enter $\mathcal{C}_1$ at a different t each time. Therefore when the ensemble average is taken of the trajectories there is an average taken over $\mathbf{x}(t) \in \mathcal{C}_2$ and $\mathbf{x}(t) \in \mathcal{C}_1$. Hence the dark trajectory from s to q in figure 4, where $\mathbf{x}(t)$ is caught between the two stationary points. By increasing the initial value of the gain function and allowing it to decrease more slowly then the result is that $\mathbf{x}(t)$ can escape from $\mathcal{C}_2$ to $\mathcal{C}_1$ (i.e. absorbing) much sooner. Thus the process which moves between the two stationary points is closer to the state q and thus the trajectory around s lightens, or the average trajectory moves closer to q.

Here we see how the ODE method is limited as it is a first order method, that is it predicts averages of the process. It does not however take into consideration the second order statistics of the process. How is it possible to apply the ODE so that this dark trajectory between two metastable states is predicted. Instead of considering the ODE for each configuration perform the average as in the simulation and average over all possible configurations that $\mathbf{x}(t)$ can be in for



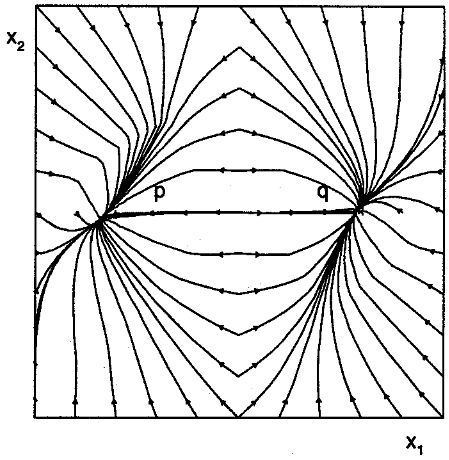


Figure 5: Left diagram for the solution of the ODEs with $\beta = 0.3$. The right diagram is for $\beta = 0.9$. Metastable states are indicated by p, q, r, s, t, v.

$t > 0$. For the case considered with $\mathbf{x}(t) \in \mathcal{C}_1 \bigcup \mathcal{C}_2$ the ODE is then given by,

$$\frac{d\mathbf{x}}{dt} = (1-\beta)\frac{d\mathbf{x}^1}{dt} + \beta\frac{d\mathbf{x}^2}{dt} \qquad (6)$$

where $0 < \beta < 1$ is a function dependent on the gain function, $\frac{d\mathbf{x}^1}{dt}$ is the ODE for $\mathcal{C}_1$, and $\frac{d\mathbf{x}^2}{dt}$ is the ODE for $\mathcal{C}_2$. The resultant ODE is solved and the phase plot is shown in the left diagram of figure 5 for $\beta = 0.3$ and in the right diagram of figure 5 for $\beta = 0.9$. In these examples the dark trajectory has been shown between the points r and p etc.

From this simple analysis it would seem that in the general case of the KNN the ODEs should be written in the form,

$$\frac{d\mathbf{x}}{dt} = \sum_{i=1}^{K} \beta_i \frac{d\mathbf{x}^i}{dt} \qquad (7)$$

where $\sum_{i=1}^{K} \beta_i = 1$ and $\frac{d\mathbf{x}^i}{dt}$ is the ODE for configuration i. The values β_i are also dependent on the gain function. From this simple analysis we see that a more accurate global picture of the average behaviour of the neuron weights is thus obtained by averaging over all the possible neuron configurations. An analysis of self-organisation in the KNN using the ODE method should now be possible.

4 Conclusion

Metastable states in a simple KNN have been examined using a set of related ODEs. In analysing the metastable states it has been found that a satisfactory explanation, based on intuition, of the complete process of self-organisation can only be obtained if the ODEs are formed from an appropriate weighting of the ODEs for each configuration. This weighting of the configuration is dependent on both the value and time derivative of the gain function. Using this set of global ODEs it should be possible to analyse the self-organising properties of the KNN.

Acknowledgements

This work was financially supported by the Swiss National Science Foundation, grant no. 20-36562,92.

References

[1] T. Kohonen. *Self-Organisation and Associative Memory.* Springer-Verlag, New York, Berlin, 1984.

[2] M. Cottrell and J.C. Fort. Étude d'un processus d'auto-organisation. *Ann. Inst. Henri Poincaré*, 23(1):1–20, Jan. 1987.

[3] H. Ritter and K. Schulten. Kohonens self-organising maps : exploring their computational capabilites. In *Proceedings of IEE Int. Conf Neural Networks*, pages 109–116, 1988.

[4] V.V. Tolat. An analysis of Kohonen's self-organising maps using a system of energy functions. *Biol. Cybern.*, 64:155–164, 1992.

[5] E. Erwin, K. Obermayer, and K. Schulten. Self-organising maps: Stationary states, metastability and convergence rates. *Biol. Cybern.*, 67:35–45, 1992.

[6] L. Ljung. Analysis of recursive stochastic algorithms. *IEEE Trans. on Automatic Control*, AC-22(4):551–575, Aug. 1977.

[7] H. J. Kushner and D. S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, volume 26 of *Applied Mathematical Sciences.* Springer-Verlag, New York, Heidelberg, Berlin, 1978.

[8] C. Bouton and G. Pagès. Convergence in distribution of the one-dimensional Kohonen algorithm when the stimuli are not uniform. *Advances in Applied Probability*, 26(1), March 1994.

[9] J.-C. Fort and G. Pagès. Sur la convergence p.s. de l'algorithme de Kohonen généralisé, 1993. Note aux Comptes Rendus de l'Académie des Sciences de Paris, t. 317, Série 1.