Note on the Maxnet Dynamics

John P.F.Sum and Peter K.S.Tam

Department of Electronic Engineering,

Hong Kong Polytechnic University,

Hung Hom, Kowloon.

April 17, 1996

Abstract

A simple method is presented to derive the complete solution of the Maxnet network dynamics. Besides, the exact response time of the network is deduced.

1 Introduction

Since the beginning of neural network research, the Winner-Take-All network has played a very important role in the design of learning algorithms, in particular, most of the unsupervised learning algorithms (Pao 1989) such as competitive learning, self organizing map and Adaptive Resonance Theory model.

Conventionally, an N-neurons winner-take-all (WTA) network is defined as following:

$$\lim_{t\to\infty} h_i(v_i(t)) = \begin{cases} 1 & \text{if } v_i(0) > v_j(0) \forall\ j \neq i \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Many researchers have attempted to design and realize the WTA. In (Lippman 1987), a discrete time algorithm called *Maxnet* is proposed.

Maxnet is a fully connected neural network. Each neuron's output is positively fed back to its input and negatively fed back to other neurons' inputs. Suppose there are N neurons. We denote their state variables and outputs by $v_i(t)$ and h_i, where $i = 1, 2, \ldots, N$, respectively. The dynamics of the Maxnet is given by

$$v_i(t+1) = h_i\left(v_i(t) - \epsilon \sum_{j\neq i} v_j(t)\right), \tag{2}$$

where

$$0 < \epsilon < \frac{1}{N-1}. \tag{3}$$

and $h_i(x)$ is a function of x defined as following. For all $i = 1, 2, \ldots, N$,

$$h_i(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \tag{4}$$

In compact form, equation (2) can be written as

$$\underline{v}(t+1) = \underline{h}\left(A\underline{v}(t)\right), \tag{5}$$

where $\underline{v}(t) = (v_1(t), v_2(t), \ldots, v_N(t))^T$, and

$$A = \begin{bmatrix} 1 & -\epsilon & \ldots & -\epsilon \\ -\epsilon & 1 & \ldots & -\epsilon \\ \ldots & \ldots & \ldots & \ldots \\ -\epsilon & -\epsilon & \ldots & 1 \end{bmatrix} \quad (6)$$

Without loss of generality, we introduce the index π_1, π_2, ..., π_N to represent the neurons with initial state in ascending order, i.e.

$$0 < v_{\pi_1} < v_{\pi_2} < \ldots < v_{\pi_N}.$$

We assume that there are no two neurons having the same initial state.

The convergence property of the Maxnet was first introduced by Lippman (1987). This may be stated as: the convergence of the Maxnet is assured if $0 < \epsilon < 1/N$. Recently, Floreen (1991) and Koutroumbas et al.(1994) studied intensively on it. In Koutroumbas et al.(1994), a general updating mode which allows partially parallel is also discussed. In both of their papers, worst case bounds on the response time of the network are derived. However, none of them has derived the complete solution of the dynamics for the network. This paper provides a complete solution and a geometrical interpretation of network dynamics which illuminates network behavior.

In this paper, we present the explicit solution of the network. This is based on an appropriate partitioning of the state space into regions in which the network behaves linearly. Similar techniques have been employed to obtain partial results for continuous-time WTA network such as Perfetti (1990) and Dempsey and McVey (1993). Furthermore, the piecewise linear approach was used to approximate the dynamical behavior of Hopfield network in some applications, such as in Gee et al. (1993). In addition, based on linear modeling techniques, interesting properties have been derived for the Grossberg shunting network (Kosko 1992)[1].

A formula for calculating the exact response time of the network is deduced in the remainder of this paper. In the next section, some recent results on the properties of the Maxnet will be stated. Then the explicit solution of the network dynamic, based on eigensubspace analysis, is deduced in Section 3. Based on this explicit solution, the exact response time is also derived. Section 4 presents the conclusions.

2 Preliminary

In (Floreen 1991) and (Koutroumbas et al. 1994), some properties of the Maxnet have been studied and proven. Here, we state them without proof.

Theorem 1 [2]

a. *If* $v_{\pi_i(0)} > v_{\pi_j(0)}$, *then* $h_{\pi_i(t)} \geq h_{\pi_j(t)}$ *for all* $t \geq 0$. *Equality holds only when* $h_{\pi_i(t)} = h_{\pi_j(t)} = 0$.

b. $h_{\pi_i(t)} \geq h_{\pi_i(t+1)}$.

Theorem 2 (Lemma 3 of (Koutroumbas et al. 1994)) *If* $v_{\pi_N}(0) = v_{\pi_j}(0)$ *for some* $j = 1, 2, \ldots, N-1$, *then* $\lim_{t\to\infty} h_{\pi_N}(t) = \lim_{t\to\infty} h_{\pi_j}(t) = 0$.

Theorem 3 (Theorem 1 of (Koutroumbas et al. 1994)) $\lim_{t\to\infty} h_{\pi_i}(t) = 0$ *for all* $i = 1, 2, \ldots, N-1$.

[1] pp.94-99.

[2] Lemma 1 of (Floreen 1991), Lemma 2 of (Koutroumbas et al. 1994).

Theorem 4 [3]

a. $\lim_{t\to\infty} h_{\pi_N}(t)$ *exists and* $\lim_{t\to\infty} h_{\pi_N}(t) \geq 0$. *Equality holds when condition in Theorem 2 is satisfied.*

b. *The limit of the network is attainable after a finite number of steps if*

$$v_{\pi_N}(0) \neq v_{\pi_j}(0) \;\; \forall j \neq N.$$

For simplicity, let us define the term *neuron settling time* and *network response time.*

Definition 1 **a.** *For all* $i = 1, 2, \ldots, N-1$, T_i *is called the* i*-th neuron settling time if* $h_{\pi_i}(t) = 0$ *for all* $t \geq T_i$.

b. T_{rt} *is called the network response time if for all* $i = 1, 2, \ldots, N-1$, $h_{\pi_i}(t) = 0$ *and* $h_{\pi_N}(t+1) = h_{\pi_N}(t)$ *for all* $t \geq T_{rt}$.

According to Theorem 4, we obtain the fact that (i) $T_{rt} < \infty$, (ii) $h_{\pi_i}(t+1) = h_{\pi_i}(t)$, for all $i = 1, 2, \ldots, N$ and for all $t \geq T_{rt}$. To visualize the overall picture given by the above theorems and definitions, a simple example is presented below.

Example: Suppose that a Maxnet consists of 3 neurons and their initial state variables are $v_1(0) = 7$, $v_2(0) = 5$ and $v_3(0) = 9$. Hence, $\pi_1 = 2$, $\pi_2 = 1$ and $\pi_3 = 3$. We set $\epsilon = 0.25$. Figure(1) shows the values of v_{π_1}, v_{π_2} and v_{π_3} as functions of t. From the graphs, we observe that the settling time of the neurons are 2 and 4 steps, and the network response time is 4 steps. Therefore, $T_1 = 2$, $T_2 = 4$ and $T_{rt} = 4$.

□

[3] Theorem 1 of (Floreen 1991), Theorem 3 of (Koutroumbas et al. 1994).

To proceed to our main result, we need the following corollary.

Corollary 1 **a.** *If* $v_{\pi_i}(0) \neq v_{\pi_j}(0)$ *for all* $i \neq j$, *then* $0 < T_1 < T_2 < \ldots < T_{N-1} = T_{rt} < \infty$.

b. *If* $v_{\pi_i}(0) = v_{\pi_{i+1}}(0)$ *for all* $i \neq N-1$, *then* $0 < T_1 < \ldots < T_i = T_{i+1} < \ldots < T_{N-1} = T_{rt} < \infty$.

(Proof of (a)) The proof is accomplished by method of contradiction. Suppose that $T_{i+1} < T_i$. Then we can establish that $h_{\pi_i}(t) = v_{\pi_i}(t) > v_{\pi_{i+1}}(t) = h_{\pi_{i+1}}(t) = 0$ at $T_{i+1} < t < T_i$. But this contradicts Theorem 1. As a result, $T_i < T_{i+1}$ for all $i = 1, 2, \ldots, N-2$. In addition to Theorem 4, we conclude that $0 < T_1 < T_2 < \ldots < T_{N-1} = T_{rt} < \infty$.

(Proof of b) Since $v_{\pi_{i+1}}(t+1) - v_{\pi_i}(t+1) = (1+\epsilon)^t (v_{\pi_{i+1}}(0) - v_{\pi_i}(0))$ for all $t > 0$, $v_{\pi_i}(T_i) = 0$ implies that $v_{\pi_{i+1}}(T_i) = 0$. By definition, $T_i = T_{i+1}$. Hence the proof is completed.

$\square$

In the next section, we proceed to derive the solution of the Maxnet using the above corollary.

3 Network Dynamics

The dynamics of the network during $0 < t < T_1$ is given by

$$\underline{v}_N(t+1) = A_N \underline{v}_N(t), \tag{7}$$

where $\underline{v}_N(t) = (v_{\pi_1}(t), v_{\pi_2}(t), \ldots, v_{\pi_N}(t))^T$ and A_N is an $N \times N$ matrix with diagonal elements equal to 1 and off diagonal elements equal to $-\epsilon$. Then, during $0 < t < T_1$, the network may be regarded as a linear discrete-time time invariant system. The solution of this system can then be obtained by evaluating the eigenvalues and the eigensubspace of A_N.

Lemma 1 *The eigenvalues of A_N are $(1-(N-1)\epsilon)$ and $(1+\epsilon)$. The corresponding eigen-subspace of $(1-(N-1)\epsilon)$ and $(1+\epsilon)$ are M_N and M_N^- respectively, where*

$$M_N = span\left\{\underline{e}_{1N} = \left(\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}, \ldots, \frac{1}{\sqrt{N}}\right)^T\right\}$$

and

$$M_N^- = \left\{\underline{v} \in R^N | \underline{v}^T \underline{e}_{1N} = 0\right\}.$$

(Proof) Let $x_i = \frac{1}{\sqrt{N}}$ for all i, i.e. $\underline{x} \in M_N$. Then

$$\begin{aligned} (A_N\underline{x})_i &= (1)(\frac{1}{\sqrt{N}}) - (N-1)\epsilon(\frac{1}{\sqrt{N}}) \\ &= (1-(N-1)\epsilon)x_i. \end{aligned} \tag{8}$$

So that, $(1-(N-1)\epsilon)$ is an eigenvalue for A_N.

Next consider $\underline{w} = \underline{v} - (\underline{v}^T\underline{e}_{1N})\underline{e}_{1N}$, $\underline{w} \in M_N{}^-$,

$$\begin{aligned} (A_N\underline{w})_i &= w_i - \epsilon\sum_{j\neq i} w_j \\ &= \left(v_i - \sum_{j=1}^{N}\frac{v_j}{N}\right) - \epsilon\sum_{j\neq i}\left(v_j - \sum_{k=1}^{N}\frac{v_k}{N}\right) \\ &= (1+\epsilon)\left(v_i - \sum_{j=1}^{N}\frac{v_j}{N}\right), \end{aligned} \tag{9}$$

i.e. $(A_N\underline{w})_i = (1+\epsilon)w_i$ or $A_N\underline{w} = (1+\epsilon)\underline{w}$. Hence the other eigenvalue is $(1+\epsilon)$. Moreover, we can replace $\underline{w}$ by

$$\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0, \ldots, 0\right)^T,$$
$$\left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}, \ldots, 0\right)^T,$$
$$\ldots$$
$$\left(\frac{1}{\sqrt{2}}, 0, 0, \ldots, \frac{-1}{\sqrt{2}}\right)^T.$$

Therefore, it can conclude that $(1-(N-1)\epsilon)$ and $(1+\epsilon)$ are the only eigenvalues of A_N because $\dim(M_N)+\dim(M_N^-)=N$.

□

With the aid of Lemma 1, the solution of the dynamics equation (7) can be written as

$$\begin{aligned} \underline{v}_N(t+1) &= (1-(N-1)\epsilon)(\underline{v}_N^T(t)\underline{e}_{1N})\underline{e}_{1N} \\ &+ (1+\epsilon)\left(\underline{v}_N(t)-(\underline{v}_N^T(t)\underline{e}_{1N})\underline{e}_{1N}\right). \end{aligned} \tag{10}$$

That is to say, for all $i=1,2,\ldots,N$,

$$\begin{aligned} v_{\pi_i}(t) &= (1-(N-1)\epsilon)^t\left(\sum_{j=1}^{N}\frac{v_{\pi_j}(0)}{N}\right) \\ &+ (1+\epsilon)^t\left(v_{\pi_i}(0)-\sum_{j=1}^{N}\frac{v_{\pi_j}(0)}{N}\right), \end{aligned} \tag{11}$$

for all $t<T_1$. It is the exact solution of Maxnet in the time interval $0\le t\le T_1$. Furthermore, the settling time of π_1^{st} neuron is given by

$$T_1=\left\lceil\frac{\log\left\{\frac{\sum_{j=1}^{N}\left(v_{\pi_j}(0)-v_{\pi_1}(0)\right)}{\sum_{j=1}^{N}v_{\pi_j}(0)}\right\}}{\log\left\{\frac{(1-(N-1)\epsilon)}{1+\epsilon}\right\}}\right\rceil. \tag{12}$$

Once v_{π_1} reaches zero, the corresponding output will also be zero. After T_1, the network dynamics can be modeled in a lower dimensional space. There are two cases to be considered: (i) $v_{\pi_2}(T_1)=0$ and (ii) $v_{\pi_2}(T_1)>0$. For case (i), we can simply skip the time interval $T_1\le t<T_2$ and proceed to consider the dynamics of the network in the time interval $T_2\le t<T_3$. In case of (ii) we can denote that

$$\underline{v}_{N-1}(t)=(v_{\pi_2}(t),v_{\pi_3}(t),\ldots,v_{\pi_N}(t))^T,$$

and consider the dynamics as

$$\underline{v}_{N-1}(t+1) = A_{N-1}\underline{v}_{N-1}(t).$$

Since A_{N-1} is defined in the same way as A_N except that the dimension is $N-1$, we can follow the same principle applied to the derivation of equation (11) and Lemma 1 to deduce that

$$\begin{aligned} v_{\pi_i}(t) &= (1-(N-2)\epsilon)^{(t-T_1)}\left\langle \underline{v}_{N-1}(T_1)\right\rangle \\ &+ (1+\epsilon)^{(t-T_1)}\left(v_{\pi_i}(T_1) - \left\langle \underline{v}_{N-1}(T_1)\right\rangle\right), \end{aligned} \tag{13}$$

for all $i = 2, 3, \ldots, N$ and

$$v_{\pi_1}(t) = 0,$$

for all $T_1 \le t < T_2$. Here

$$\left\langle \underline{v}_{N-1}(T_1)\right\rangle = \sum_{j=2}^{N} \frac{v_{\pi_j}(T_1)}{N-1}.$$

Repeating the same procedure, we can derive the general solution of the Maxnet for all time $t \ge 0$. Denoting that

$$\underline{v}_{N-k}(t) = (v_{\pi_{k+1}}(t), v_{\pi_{k+2}}(t), \ldots, v_{\pi_N}(t))^T,$$

the general solution of the network at time $T_k \le t < T_{k+1}$ is given by

$$\begin{aligned} v_{\pi_i}(t) &= (1-(N-k-1)\epsilon)^{(t-T_k)}\left\langle \underline{v}_{N-k}(T_k)\right\rangle \\ &+ (1+\epsilon)^{(t-T_k)}\left(v_{\pi_i}(T_k) - \left\langle \underline{v}_{N-k}(T_k)\right\rangle\right), \end{aligned} \tag{14}$$

for all $i = k+1, k+2, \ldots, N$ and

$$v_{\pi_i}(t) = 0, \tag{15}$$

for all $i = 1, 2, \ldots, k$, where

$$\langle \underline{v}_{N-k}(T_k) \rangle = \sum_{j=k+1}^{N} \frac{v_{\pi_j}(T_k)}{N-k}.$$

Besides, the settling time for π_2^{nd}, π_3^{rd}, ..., π_{N-1}^{th} neurons can be obtained recursively by

$$T_{k+1} = \begin{cases} T_k + \left\lceil \dfrac{\log\left\{ \frac{\sum_{j=k+1}^{N} \left(v_{\pi_j}(T_k) - v_{\pi_{k+1}}(T_k) \right)}{\sum_{j=k+1}^{N} v_{\pi_j}(T_k)} \right\}}{\log\left\{ \frac{(1-(N-k-1)\epsilon)}{1+\epsilon} \right\}} \right\rceil & \text{if } v_{\pi_{k+1}}(T_k) > 0 \\ T_k & \text{if } v_{\pi_{k+1}}(T_k) = 0. \end{cases} \tag{16}$$

Since $v_{\pi_N}(t+1) = v_{\pi_N}(t)$ whenever $t \geq T_{N-1}$, the network response time is given by

$$T_{rt} = \sum_{i=1}^{N-1} \left\lceil \frac{\log\left\{ \frac{\sum_{k=i}^{N} \left(v_{\pi_k}(T_{i-1}) - v_{\pi_i}(T_{i-1}) \right)}{\sum_{k=i}^{N} v_{\pi_k}(T_{i-1})} \right\}}{\log\left\{ \frac{(1-(N-i)\epsilon)}{1+\epsilon} \right\}} \right\rceil, \tag{17}$$

where $T_0 = 0$ and $\lceil x \rceil$ is the smallest integer which is just greater than x.

4 Geometrical Interpretation

In (Koutroumbas et al. 1994), Koutroumbas et.al. presented a brief geometrical interpretation for two dynamical properties of the Maxnet: (i) once the initial state $\underline{v}_N$ is located on the hyperplane that bisects the angles between the reference axis, the limit vector will be the null vector, and (ii) otherwise, the limit point will be on the axis that corresponds to the node π_N. Essentially, these properties can be easily visualized from equations (14), (15) and Lemma 1.

To simplify the discussion, we describe the case of two neurons but the interpretation can be extended to N neurons. From Lemma 1, it is observed that the component of $\underline{v}_2$ which is parallel to $\underline{e}_{12}$ will decrease at a rate $(1-\epsilon)$, while the component perpendicular to $\underline{e}_{12}$ will increase at a rate of $(1+\epsilon)$. Figure(2) shows three situations, indicated by x, y and z. x_1 and y_1 are the components of x and y which are parallel to e. Whereas, x_2 and y_2 are the components perpendicular to e. The lengths of the arrows indicate the corresponding rates of change.

Consider x, i.e. in region A, the magnitude of the change of x along e is $-\epsilon x_1$ which is larger than ϵx_2. The resultant change of x is pointing towards the axis v_2. Similarly, the resultant change of x will point towards the v_1 axis if x is located in the other A region. In region B, $y_1 \geq y_2$. Equality holds only when y lies on the axis v_1. Therefore, the change of y along the direction of e is also larger than that along the direction perpendicular to e. The resultant change of y is again towards one of the axes. Consider z, which is on the line of e; its resultant change is pointing towards $(0,0)$.

In summary, if $v_1(0) > v_2(0)$ ($v_2(0) > v_1(0)$), then the limit point will be on the v_1 (v_2) axis. If $v_1(0) = v_2(0)$, then the limit point will be $(0,0)$.

5 Conclusion

In this paper, we have derived the complete solution of the Maxnet. This solution provides an alternative approach to understand the properties of Maxnet. Besides, the exact response time is also deduced as long as $v_{\pi_N}(0) \neq v_{\pi_{N-1}}(0)$. Since our derivation of the solution is based on the method of eigensubspace analysis, the geometrical interpretation of the network dynamics can be described vigorously. Such a technique can be readily adapted to the analysis of similar WTA networks such as Imax (Yen and Chang 1992), Gemnet (Yang et al. 1995) and Selectron (Yen et al. 1994).

Acknowledgement

We would like to thank the anonymous referee for the valuable comments.

References

Dempsey G.L. and McVey E.S. 1993. Circuit Implementation of a Peak Detector Neural Network, *IEEE Transactions on Circuits and Systems-II*, Vol.40, No.9, September, pp.585-591.

Floreen P. 1991. The Convergence of Hamming Neural Networks, *IEEE Transactions on Neural Networks*, Vol.2, No.4, July, pp.449-457.

Gee A.H. et al. 1993. An analytical framework for optimizing neural networks, *Neural Networks*, vol. 6. no. 1, 79-97.

Kosko B. 1992. *Neural Networks And Fuzzy Systems.* Prentice-Hall, Inc.

Koutroumbas K. and Kalouptsidis N. 1994. Qualitative Analysis of the Parallel and Asynchronous Modes of the Hamming Network, *IEEE Transactions on Neural Networks*, Vol.15, No.3, May, pp.380-391.

Lippman R. 1987. An introduction to computing with neural nets, *IEEE ASSP Magazine*, Vol. 4, pp.4-22.

Pao Y. 1989. *Adaptive Pattern Recognition and Neural Networks.* Addison-Wesley.

Perfetti R. 1990. Winner-take-all circuit for neurocomputing applications, *IEE Proceedings Part G*, Vol.137, No.5, 353-359.

Yang J. et al. 1995. A General Mean-Based Iterative Winner-Take-All Neural Network, *IEEE Transactions on Neural Networks*, Vol.6, No.1, 14-24.

Yen J.-C. and Chang S. 1992. Improved Winner-Take-All Neural Network, *Electronics Letters*, Vol.28, No.7 662-664.

Yen J.-C. et al. 1994. A New Winners-Take-All Architecture in Artificial Neural Networks, *IEEE Transactions on Neural Networks*, Vol.5, No.5, 838-843.

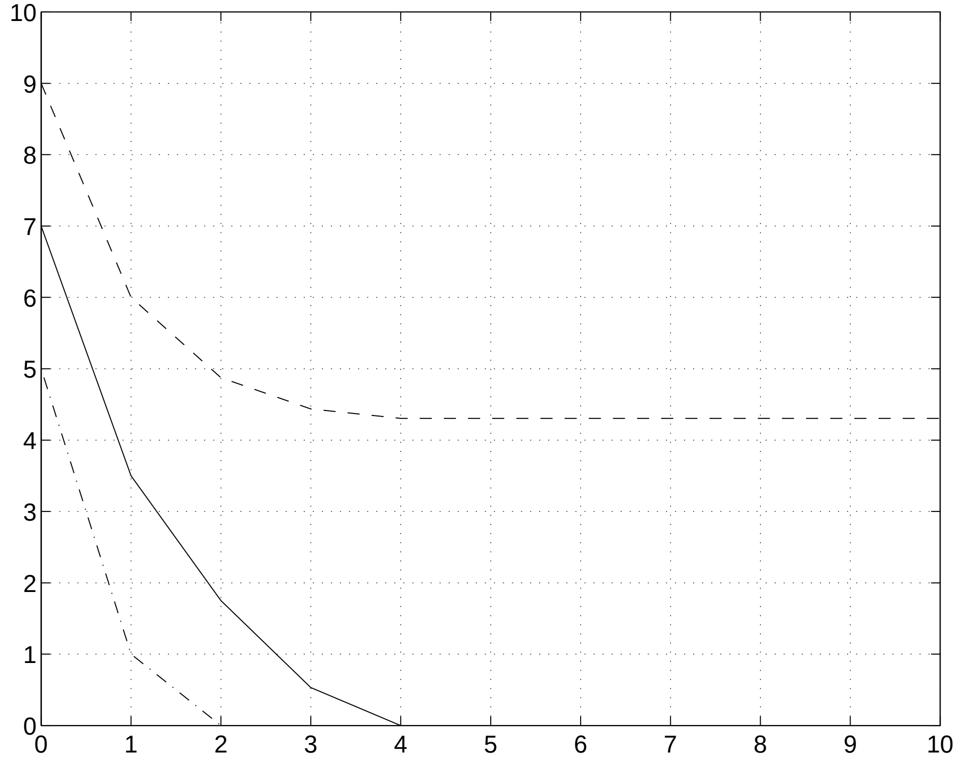

Figure 1: Changes of v_1 (solid line), v_2 (dot-dash line) and v_3 (dash line) against time.

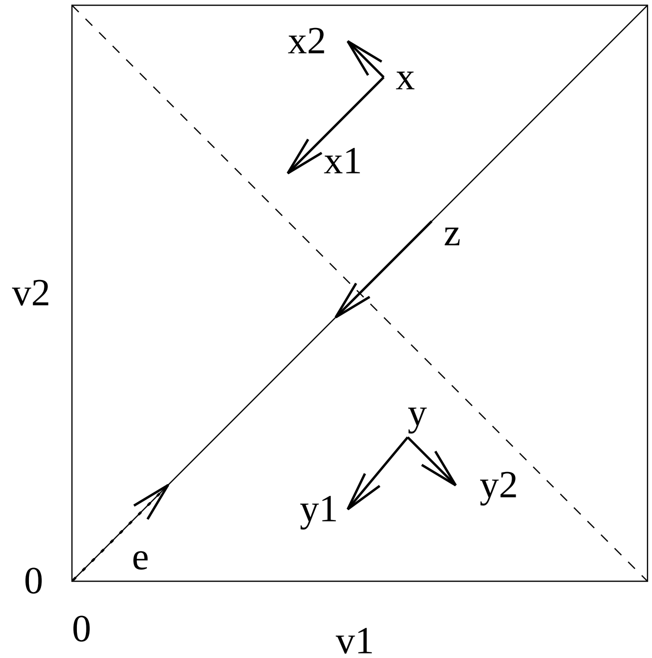

Figure 2: The geometrical interpretation of the dynamics of the Maxnet. x, y, z are corresponding to three initial conditions which are located in three regions, A, B and the line along e.