Brain State in a Convex Body

Martin Bohner and Stefen Hui

Abstract—**We study a generalization of the brain-state-in-a-box (BSB) model for a class of nonlinear discrete dynamical systems where we allow the states of the system to lie in an arbitrary convex body. The states of the classical BSB model are restricted to lie in a hypercube. Characterizations of equilibrium points of the system are given using the support function of a convex body. Also, sufficient conditions for a point to be a stable equilibrium point are investigated. Finally we study the system in polytopes. The results in this special case are more precise and have simpler forms than the corresponding results for general convex bodies. The general results give one approach of allowing pixels in image reconstruction to assume more than two values.**

I. Introduction

The brain-state-in-a-box (BSB) neural model was proposed by Anderson and coworkers in 1977 (see [1]). It can be described by the equation

$$x_{k+1} = g(x_k + Wx_k)$$

where x_0 is an element of the closed n-dimensional unit hypercube, x_k is the state of the system at time k, W some weight matrix, and the function g ensures that the states of the system are constrained to be in the unit hypercube. The BSB model has been investigated by many researchers, among them Anderson *et al.* ([1], [5, chapter 4]), Golden [2], Greenberg [3], Grossberg [4], Hui and Żak [7], and Hui *et al.* [5, chapter 11].

One of the applications of the BSB model is to store patterns in such a way that when presented with a new pattern p, the system responds by finding the stored pattern most closely resembles p. This problem is known as the associative memory problem (see [5] and [6]). One can study the equilibrium points of the system: the points e such that $g(e+We) = e$. Of greater interest is the set of all stable equilibrium points, namely, those points s where there exists an entire neighborhood around s with $g(x + Wx) = s$ for all x in that neighborhood. We can consider the stable equilibrium points of the system described above as the stored patterns. The neighborhood of attraction then contains the noisy versions of the stored pattern q which should be identified with q. It is useful to choose the extreme points of the hypercube to be the equilibrium points of the system. Hui and Żak ([5, chapter 11], [7]) were able to give conditions on the matrix W so that this occurs.

Of course the number of stored patterns is restricted to be 2^n for some natural number n in this case. The BSB model also only allows the coordinates to assume two values. For example, if one thinks of each coordinate as the value of a pixel in a two-dimensional image, then the BSB model only allows each pixel to be on or off with no possibility of a gray scale. In the present paper we introduce a generalization of the BSB model which can be used to address these problems.

We fix an arbitrary closed, convex, bounded, and nonempty set S and consider the system described by

$$x_{k+1} = g(f(x_k))$$

where f is any continuous function and g maps from S to itself. The precise descriptions are given in Section II. Of course we are interested in the equilibrium points of the system, and to find them, it is useful to introduce the support function of S and to look at some properties of convex sets and convex functions. Using the support function of S, we will give a characterization of the set of all equilibrium points which yields a necessary and sufficient condition for the statement

$$\text{all vertices of } S \text{ are equilibrium points.} \tag{i}$$

Moreover, if the support function of S is differentiable at certain points, then it is even easier to check whether a point is an equilibrium point of the system. This is explored in Section III. In Section IV we look at stability of the equilibrium points. First, a sufficient condition for a point to be a stable equilibrium point is given. We can simplify this condition if S is a polytope and we give conditions which imply

$$\text{all vertices of } S \text{ are stable equilibrium points.} \tag{ii}$$

For the remainder of Section IV, our system is governed by a linear function $f(x) = x + Wx$, where W is a weight matrix. We give conditions on W for (i) and (ii) which are numerically very easy to check. On the other hand, if we would like to have a finite number of fixed points be the equilibrium points, we can choose S to be the convex hull of those points. This may be one approach to reducing the number of spurious equilibrium points. Also, a pool of matrixes W which will work for (i) and (ii) is given in this section. We can choose from this pool the matrixes that are the best for a particular application. In particular, the results of Hui and Żak [7] for $S = [-1, 1]^n$ will be easy consequences of our general theory. In Section V, we indicate how the results can be applied to the gray scale problem and give a numerical example.

II. Definitions and Background Results

Definition 1: Let $\mathcal{H}$ be a Hilbert space over $\mathbb{R}$ with dim $\mathcal{H} = n \in \mathbb{N}$. A closed, convex, bounded, nonempty, and n-dimensional subset of $\mathcal{H}$ is called a convex body. ■

Let S denote a convex body. Let f be a continuous function and let $x_k = f(x_{k-1})$. We are interested in the restriction of

Manuscript received October 22, 1992; revised November 23, 1994.

M. Bohner is with Abteilung Mathematik V, University of Ulm, Helmholtzstrasse 18, D-89069 Ulm, Germany.

S. Hui is with the Department of Mathematical Sciences, San Diego State University, San Diego, CA 92182 USA.

IEEE Log Number 9409355.

1045–9227/95$04.00 © 1995 IEEE

the system to S. Since we want all the points to stay in S, we need to send the points which fall out of S back to S. To do that, we need the following lemma.

Lemma 1: For each $y \in \mathcal{H}$ there exists a unique $g(y) \in S$ such that

$$\|y - g(y)\| = \inf_{s \in S} \|y - s\|.$$

Furthermore $g : \mathcal{H} \to S$ is continuous.

Proof: A proof can be found, for example, in [11, p. 27]. ∎

With the "nearest-point-map" g of the above lemma we can define our system.

Definition 2: Let $x_0 \in S$. Define

$$x_{k+1} := g(f(x_k)) \qquad \forall\, k \in \mathbb{N} \cup \{0\}$$

where $f : S \to \mathcal{H}$ is continuous and $g: \mathcal{H} \to S$ is the nearest-point-map. ∎

That is, if $f(x_k) \notin S$ for some k, we take it back to the unique point in S which minimizes the distance to $f(x_k)$.

Definition 3: Let $T(x) := g(f(x))$ for $x \in S$ and let $x^* \in S$.

1) If $T(x^*) = x^*$, then x^* is called an equilibrium point of the system. With Equi(S) we denote the set of all such points.
2) Let $\Delta(x^*, \delta) := \{s \in \mathcal{H} |\ \|x^* - s\| < \delta\}$. If there exists $\delta > 0$ so that $T(S \cap \Delta(x^*, \delta)) = \{x^*\}$, then x^* is called a stable equilibrium point of the system. The set of all stable equilibrium points is referred to as Equi$^*(S)$. ∎

In other words, an equilibrium point is stable if there exists a neighborhood of the point so that all the points in that neighborhood are sent to the equilibrium point in one step. Observe also that Equi$(S) \neq \emptyset$ by a consequence of Brouwer's fixed point theorem.

Before we can give conditions for a point of S to be an equilibrium point, we need some properties of convex sets. To begin with, we define the support function of a convex body.

Definition 4: The function $h : \mathcal{H} \to \mathbb{R}$ defined by

$$h(u) := \sup_{s \in S} \langle s, u \rangle, \qquad u \in \mathcal{H}$$

is called the support function of S. For each $u \in \mathcal{H}$, let $H_u := \{x \in \mathcal{H} | \langle x, u \rangle \le h(u)\}$. ∎

Of course, H_u is just the half-space containing S determined by the hyperplane that is orthogonal to u and tangent to S. Clearly, if $s_0 \in S$ with $h(u) = \langle s_0, u \rangle$, then $s_0 \in S \cap \partial H_u$ and $\partial H_u = \{x \in \mathcal{H} | \langle x, u \rangle = h(u)\}$ is a hyperplane which supports S at s_0.

Some properties of the support function are collected in the following lemma.

Lemma 2: Let S be a convex body and h its support function. Then

1) h is real-valued,
2) $h(u) = \max_{s \in \partial S} \langle s, u \rangle \ \forall\, u \in \mathcal{H}$, and
3) h is subadditive, positively homogenous, and convex on $\mathcal{H}$.

Proof: The boundedness of S together with the Cauchy–Schwarz inequality imply 1). Everything else can be verified easily. ∎

To become familiar with the support function, we will give two easy examples on how to compute it. We use $\overline{X}$, ∂X, and $\mathring{X}$ to denote the closure, boundary, and interior of a set X, respectively.

Example 1 (Support Function):

1) Let $S = [-1, 1]^n$ be the closed n-dimensional unit hypercube in $\mathbb{R}^n$. We can calculate the support function h of S as follows

$$\begin{aligned} h(u) &= \sup_{s \in S} \langle s, u \rangle = \sup_{-1 \le s_i \le 1} \sum_{i=1}^{n} s_i u_i = \sum_{i=1}^{n} \mathrm{sgn}(u_i) u_i \\ &= \sum_{i=1}^{n} |u_i| \ \ \forall\, u = (u_i)_{1 \le i \le n} \in \mathbb{R}^n. \end{aligned}$$

2) Let $S = \overline{\Delta(x_0, k)} = \{s \in \mathbb{R}^n |\ \|x_0 - s\| \le k\}$ be the closed n-dimensional ball of radius k in $\mathbb{R}^n$ around x_0. For $s \in S$, we can find $a \in \overline{\Delta(0, k)}$ with $s = x_0 + a$. Thus we have

$$\begin{aligned} h(u) &= \sup_{s \in S} \langle s, u \rangle = \langle x_0, u \rangle + \sup_{a \in \overline{\Delta(0,k)}} \langle a, u \rangle \\ &= \langle x_0, u \rangle + \left\langle \frac{ku}{\|u\|}, u \right\rangle \\ &= \langle x_0, u \rangle + k\|u\| \qquad \forall\, u \in \mathbb{R}^n \end{aligned}$$

where we applied the equality part of the Cauchy–Schwarz inequality. ∎

Since $\langle x - y, u \rangle = h(u) - h(u) = 0$ for all $x, y \in \partial H_u$, the vector u is normal to the hyperplane ∂H_u. Now the geometric meaning of the following definition, where we denote $\{x + a | a \in A\}$ for $x \in \mathcal{H}$ and $A \subset \mathcal{H}$ by $x + A$, is clear.

Definition 5: Let $x \in \partial S$.

1) $N(x) := x + \{u \in \mathcal{H} | x \in S \cap \partial H_u\}$ is called the normal cone of S at x.
2) $N^*(x) := x + \{u \in \mathcal{H} | \{x\} = S \cap \partial H_u\}$ is called the absolute normal cone of S at x.
3) x is said to be a vertex of S, if all affine subspaces containing $N(x)$ have dimension n. The set of all vertices of S is denoted by Vert(S). ∎

We illustrate the above definitions with an example.

Example 2: Let $S \subset \mathbb{R}^2$ be the region as depicted in Fig. 1. Let $l_a, l_b, l_u, \tilde{l}_u, l_v$, and $\tilde{l}_v$ be line segments and R_u, R_v be open sectors as shown. We have

$$\begin{aligned} &N(a) = N^*(a) = l_a, \quad N(b) = l_b, \quad N^*(b) = \emptyset, \\ &N(u) = l_u \cup \tilde{l}_u \cup R_u, \quad N^*(u) = l_u \cup R_u, \\ &N(v) = l_v \cup \tilde{l}_v \cup R_v, \quad \text{and} \quad N^*(v) = R_v. \end{aligned}$$

Furthermore, we have Vert$(S) = \{u, v, w\}$. ∎

Observe that $N(x) = x + \{u \in \mathcal{H} | \langle x, u \rangle = h(u)\}$ and that $N(x)$ is the collection of the outward normal vectors to the supporting hyperplanes at x. Moreover, it is easy to verify that $N(x)$ is convex for each $x \in \partial S$. In the next section, the normal cone $N(x)$ will be used to give necessary and

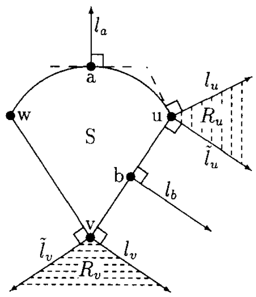


Fig. 1. Illustration of Definition 5.

sufficient conditions for Vert(S) $\subset$ Equi(S). In Section IV, we will then use the absolute normal cone $N^*(x)$ to derive sufficient conditions for Vert(S) $\subset$ Equi*(S). For both results we need the following two theorems which are well-known results about convex sets and convex functions.

Theorem 1 (Separating and Supporting Properties):

1) If S is convex and $x_0 \in \partial S$, then there exists at least one hyperplane supporting S at x_0.
2) If A and B are convex with $\overset{\circ}{A} \neq \emptyset$ and $\overset{\circ}{A} \cap B = \emptyset$, then there exists a hyperplane separating A and B.
3) If S is a convex body and $x_0 \notin S$, then there exists a hyperplane which strictly separates $\{x_0\}$ and S.

Proof: For a proof see, for example, [8, pp. 36, 38, and 41]. ■

Theorem 2: Let $h : \mathcal{H} \to \mathbb{R}$ be convex. Then

1) h is continuous on $\mathcal{H}$,
2) $\delta h(x, y) := \lim_{\varepsilon \to 0^+}(h(x + \varepsilon y) - h(x))/\varepsilon$ exists $\forall\, x, y \in \mathcal{H}$,
3) $\delta h(x, y) = -\delta h(x, -y) \Leftrightarrow (\partial h/\partial y)(x)$ exists, and
4) if $(\partial h/\partial y)(x)$ exists for all $y \in \mathcal{H}$, then h is differentiable at x with

$$\frac{\partial h}{\partial y}(x) = \langle \nabla h(x), y \rangle.$$

Proof: Again we refer the reader to [9, pp. 93, 101]. ■

III. Equilibrium Points

Now we can return to the model described in Section I. The goal is now to give a necessary and sufficient condition for a point to belong to Equi(S). The corollaries of the following two theorems will give necessary and sufficient conditions for Vert(S) $\subset$ Equi(S). They can be considered as the main results of this paper. Before studying the proofs of Theorems 3 and 4, the reader may also first have a look at Example 3 where the stability results concerning the classical BSB model (see [5, chapter 11] and [7]) are derived as easy consequences of our new general theory.

Theorem 3: Let $x \in S$ and $s \in \partial S$. Then

$$f(x) \in N(s) \Leftrightarrow T(x) = s.$$

Proof: First, suppose $f(x) \in N(s)$, i.e., $\langle s, f(x) - s\rangle = h(f(x) - s)$ and let $t \in S$. Then with the aid of the Cauchy–Schwarz inequality we find that

$$\begin{aligned} \|f(x) - t\|\,\|f(x) - s\| &\geq \langle f(x) - t, f(x) - s\rangle \\ &= \langle f(x), f(x) - s\rangle - \langle t, f(x) - s\rangle \\ &\geq \langle f(x), f(x) - s\rangle - h(f(x) - s) \\ &= \langle f(x) - s, f(x) - s\rangle \\ &= \|f(x) - s\|^2. \end{aligned}$$

Now we have either $f(x) \in S$ which implies $0 \geq \|f(x) - s\|^2$, i.e., $s = f(x) = g(f(x)) = T(x)$, or $f(x) \notin S$, and then

$$\inf_{t \in S} \|f(x) - t\| \geq \|f(x) - s\|$$

so that again (see Lemma 1) $s = g(f(x)) = T(x)$ holds.

Now suppose $T(x) = s$. Since $h(0) = 0$ we can assume without loss of generality that $f(x) \notin S$. We define a new convex set

$$B := \overline{\Delta(f(x), f(x) - s)}.$$

Note that B is the ball around $f(x)$ which touches S at the point $s \in \partial S$. We have $\overset{\circ}{B} \cap S = \emptyset$ since

$$\|f(x) - s\| = \min_{t \in S} \|f(x) - t\|.$$

Thus we can separate those two convex sets by a hyperplane [see Theorem 1-2)], i.e., there exists $u \in \mathcal{H}\backslash\{0\}$ such that

$$\langle t, u\rangle \leq h(u) \leq \langle b, u\rangle \quad \forall\, t \in S, \quad \forall\, b \in B.$$

Of course, we have $s \in B \cap S$, which yields $\langle s, u\rangle = h(u)$. Defining now $u^* := -u$ we see that

$$\langle b, u^*\rangle \leq h(u^*) = \langle x, u^*\rangle \qquad \forall\, b \in B.$$

Thus the hyperplane $\{a \in \mathcal{H} |\ \langle a, u^*\rangle = h(u^*)\}$ supports B at $s \in \partial B$. Looking at Example 1-2), where we computed the support function of a ball, we conclude that

$$\langle s, u^*\rangle = h(u^*) = \langle f(x), u^*\rangle + \|f(x) - s\|\,\|u^*\|.$$

We have then

$$\langle s - f(x), u^*\rangle = \|f(x) - s\|\,\|u^*\|$$

and by the equality part of the Cauchy–Schwarz inequality, there exists $\alpha > 0$ so that $u^* = \alpha(s - f(x))$. Remembering how u^* was defined and applying Lemma 2-3), we arrive at

$$\begin{aligned} \alpha h(f(x) - s) &= h(\alpha(f(x) - s)) \\ &= h(-u^*) = h(u) = \langle s, u\rangle \\ &= \langle s, \alpha(f(x) - s)\rangle = \alpha\langle s, f(x) - s\rangle \end{aligned}$$

which yields $\langle s, f(x) - s\rangle = h(f(x) - s)$ and $f(x) \in N(s)$. ■

Corollary 1: For $x_0 \in \partial S$ we have

$$f(x_0) \in N(x_0) \Leftrightarrow x_0 \in \text{Equi}(S).$$

Furthermore, the condition

$$f(x_0) \in N(x_0) \qquad \forall\, x_0 \in \text{Vert}(S) \qquad (A)$$

is necessary and sufficient for Vert(S) $\subset$ Equi(S).

Proof: Let $s = x = x_0$ in Theorem 3. ■

Theorem 4: Let $u_0 \in \mathcal{H}$. Then h is differentiable at u_0 if and only if there exists $x_0 \in \partial S$ such that $x_0 + u_0 \in N^*(x_0)$ and in this case we have $x_0 = \nabla h(u_0)$.

Proof: Let us assume that h is differentiable at u_0. Let $x^* \in S \cap \partial H_{u_0}$. Our goal is to show that $x^* = \nabla h(u_0)$. For arbitrary $u \in \mathcal{H}$ and $\varepsilon > 0$ we have

$$\frac{h(u_0 + \varepsilon u) - h(u_0)}{\varepsilon} = \frac{h(u_0 + \varepsilon u) - \langle x^*, u_0 \rangle}{\varepsilon} \geq \frac{\langle x^*, u_0 + \varepsilon u \rangle - \langle x^*, u_0 \rangle}{\varepsilon} = \langle x^*, u \rangle.$$

Letting ε tend to zero from above, we find

$$\delta h(u_0, u) \geq \langle x^*, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Thus we have for all $u \in \mathcal{H}$

$$\langle x^*, u \rangle = -\langle x^*, -u \rangle \geq -\delta h(u_0, -u) = \delta h(u_0, u) \geq \langle x^*, u \rangle.$$

Note that the last equality is a consequence of Theorem 2.3) since h is differentiable at u_0. Now it follows by Theorem 2.4) that

$$\langle \nabla h(u_0), u \rangle = \frac{\partial h}{\partial u}(u_0) = \delta h(u_0, u) = \langle x^*, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Since the above is true for all $u \in \mathcal{H}$, we have $\|\nabla h(u_0) - x^*\| = 0$ which shows that $S \cap \partial H_{u_0} = \{\nabla h(u_0)\}$ holds.

Conversely suppose $x_0 \in \partial S$ with $S \cap \partial H_{u_0} = \{x_0\}$. To compute $\delta h(u_0, u)$ [which exists by Theorem 2-2)] for $u \in \mathcal{H}$ we begin with

$$\frac{h(u_0 + \varepsilon u) - h(u_0)}{\varepsilon} \geq \frac{\langle x_0, u_0 + \varepsilon u \rangle - \langle x_0, u_0 \rangle}{\varepsilon} = \langle x_0, u \rangle \qquad \forall\, \varepsilon > 0.$$

So we have $\delta h(u_0, u) \geq \langle x_0, u \rangle \;\; \forall\, u \in \mathcal{H}$. Now we turn our attention to the opposite inequality. For each $\varepsilon > 0$ there exists $x_0(\varepsilon) \in S$ with

$$h(u_0 + \varepsilon u) = \langle x_0(\varepsilon), u_0 + \varepsilon u \rangle.$$

Since S is bounded, the sequence $\{x_0(1/n)\}_{n \in \mathbb{N}}$ is bounded also. Therefore, by the Banach–Alaoglu Theorem for Hilbert spaces (see, for example, [10, p. 77]), it contains a weakly convergent subsequence $\{x_0(1/n_k)\}$, say

$$\lim_{k \to \infty} \left\langle x_0\left(\frac{1}{n_k}\right), u \right\rangle = \langle x^*, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Also, $x^* \in S$, since closed convex sets are weakly closed (see [10, p. 81]). By the definition of $\{x_0(1/n_k)\}$, we have

$$\frac{h\left(u_0 + \dfrac{u}{n_k}\right) - h(u_0)}{\dfrac{1}{n_k}} \leq \frac{\left\langle x_0\left(\dfrac{1}{n_k}\right), u_0 + \dfrac{u}{n_k} \right\rangle - \left\langle x_0\left(\dfrac{1}{n_k}\right), u_0 \right\rangle}{\dfrac{1}{n_k}} = \left\langle x_0\left(\frac{1}{n_k}\right), u \right\rangle.$$

Therefore, letting $k \to \infty$ in the inequality, it follows that

$$\delta h(u_0, u) \leq \langle x^*, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

But since h is continuous [Theorem 2-1)], we can write

$$h\left(u_0 + \frac{u}{n_k}\right) = \left\langle x_0\left(\frac{1}{n_k}\right), u_0 \right\rangle + \frac{1}{n_k}\left\langle x_0\left(\frac{1}{n_k}\right), u \right\rangle$$

and let $k \to \infty$ to arrive at

$$h(u_0) = \langle x^*, u_0 \rangle + 0\langle x^*, u \rangle = \langle x^*, u_0 \rangle.$$

By assumption, x_0 is the only element in S which satisfies the above equality, therefore $x^* = x_0$ and

$$\delta h(u_0, u) \leq \langle x_0, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Combining the above inequalities, we conclude that

$$-\delta h(u_0, -u) = -\langle x_0, -u \rangle = \langle x_0, u \rangle = \delta h(u_0, u) \qquad \forall\, u \in \mathcal{H}.$$

Applying finally parts 3) and 4) of Theorem 2, we see that the (two-sided) directional derivative $(\partial h / \partial u)(u_0)$ exists for all $u \in \mathcal{H}$. Thus h is differentiable at u_0 and

$$\langle x_0, u \rangle = \frac{\partial h}{\partial u}(u_0) = \langle \nabla h(u_0), u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Therefore $\nabla h(x_0) = x_0$. and the proof is complete. ■

Theorem 5: Let $x \in S$ and $s \in \partial S$. Then the following are equivalent:

1) $T(x) = s$ and h is differentiable at $f(x) - s$,
2) h is differentiable at $f(x) - s$ with $\nabla h(f(x) - s) = s$, and
3) $f(x) \in N^*(s)$.

Proof: Suppose 1) holds. Then $f(x) \in N(s)$ by Theorem 3 and

$$\nabla h(f(x) - s) + f(x) - s \in N^*(\nabla h(f(x) - s))$$

by Theorem 4 which yields

$$s \in S \cap \partial H_{f(x)-s} = \{\nabla h(f(x) - s)\}.$$

Therefore 1) implies 2).

Using Theorem 4, we see that 2) implies

$$f(x) = \nabla h(f(x) - s) + f(x) - s \in N^*(\nabla h(f(x) - s)) = N^*(s)$$

and 3) holds.

From 3) it follows that

$$s + (f(x) - s) = f(x) \in N^*(s) \subset N(s)$$

holds which implies 1) by Theorem 3 and Theorem 4. ■

Corollary 2: Let $x_0 \in \partial S$. Then the following are equivalent:

1) $x_0 \in \mathrm{Equi}(S)$ and h is differentiable at $f(x_0) - x_0$,
2) h is differentiable at $f(x_0) - x_0$ with $\nabla h(f(x_0) - x_0) = x_0$, and
3) $f(x_0) \in N^*(x_0)$.

Proof: Theorem 5 with $s = x = x_0$. ■

To see how applicable the condition given in Corollary 2 is, we will now give two examples. The first deals with the n-dimensional hypercube and the second with the n-dimensional unit ball. The first example contains a derivation of a well-known result from the study of the BSB model using the techniques presented above.

Example 3: Let $S = [-1,1]^n$ and define

$$E := \{e = (e_i)_{1\le i\le n} \mid |e_i| = 1 \quad \forall\, i \in \{1,\cdots,n\}\}.$$

The set E is the collection of all extreme points of S. Let us assume that $E \subset \text{Equi}(S)$ for $f(x) = x + Wx + b$, where $f_i(e) \ne e_i \;\forall\, e \in E$. Since $h(u) = \sum_{i=1}^n |u_i|$ by Example 1-1), we have

$$\nabla h(u) = (\text{sgn} u_i)_{1\le i\le n}, \quad u = (u_i)_{1\le i\le n}, \quad u_i \ne 0 \quad \forall\, i \in \{1,\cdots,n\}.$$

We have $E \subset \text{Equi}(S)$ and so the following holds by Corollary 2 for each $e = (e_i)_{1\le i\le n} \in E$

$$e = \nabla h(f(e) - e) = (\text{sgn}(b_i + (We)_i))_{1\le i\le n} = \left(\text{sgn}\left(w_{ii}e_i + b_i + \sum_{j=1, j\ne i}^n w_{ij}e_j\right)\right)_{1\le i\le n}.$$

By a suitable choice of the vector e we see that a necessary condition of the required equation is given by

$$w_{ii} > -b_i + \sum_{j=1,j\ne i}^n |w_{ij}| \quad \text{and}$$
$$w_{ii} > b_i + \sum_{j=1,j\ne i}^n |w_{ij}| \qquad \forall\, i \in \{1,\cdots,n\}$$

that is

$$w_{ii} > |b_i| + \sum_{j=1,j\ne i}^n |w_{ij}| \qquad \forall\, i \in \{1,\cdots,n\}. \qquad (S)$$

We will show later that this condition is also sufficient for the stability of the vertices.

Observe that a matrix which satisfies condition (S) is necessarily strongly row diagonal dominant, that is

$$w_{ii} > \sum_{j=1,j\ne i}^n |w_{ij}| \qquad \forall\, i \in \{1,\cdots,n\}$$

holds. For some properties of strongly row diagonal dominant matrices see for example [7]. ■

Example 4: Let $S = \Delta(0,1)$. By Example 1-2), we know that $h(u) = \|u\|$ for $u \in \mathcal{H}$. Thus h is differentiable whenever $u \in \mathcal{H}\backslash\{0\}$ and we can calculate the partial derivatives

$$\frac{\partial h}{\partial u_\nu}(u) = \frac{1}{2\|u\|}2u_\nu = \frac{u_\nu}{\|u\|} \qquad \forall\, u \in \mathcal{H}\backslash\{0\}.$$

Therefore

$$\nabla h(u) = \frac{u}{\|u\|} \qquad \forall\, u \in \mathcal{H}\backslash\{0\}.$$

We conclude that h is differentiable at $f(x_0) - x_0$ provided $f(x_0) \ne x_0$ and in this case we have by Corollary 2 that

$$x_0 \in \text{Equi}(S) \Leftrightarrow \nabla h(f(x_0) - x_0) = x_0 \Leftrightarrow x_0 = \frac{f(x_0) - x_0}{\|f(x_0) - x_0\|} \Leftrightarrow f(x_0) = x_0(\|f(x_0) - x_0\| + 1).$$

So we have: $\text{Equi}(S) = \partial S \Leftrightarrow \exists \alpha \ge 1$ with $f(x) = \alpha x$. ■

IV. Stable Equilibrium Points

We already described conditions under which $\text{Vert}(S) \subset \text{Equi}(S)$ is true. It would be even more pleasant if we have $\text{Vert}(S) \subset \text{Equi}^*(S)$. The next goal is to give a sufficient condition for this desired situation.

Theorem 6: Let $x_0 \in \partial S$. Then

$$f(x_0) \in \overset{\circ}{N}(x_0) \Rightarrow x_0 \in \text{Equi}^*(S).$$

Proof: By assumption there exists $\varepsilon > 0$ so that $\Delta(f(x_0),\varepsilon) \subset N(x_0)$. Recall that f is assumed to be continuous (see Definition 2). Hence, corresponding to the above ε there exists $\delta > 0$ so that

$$f(\Delta(x_0,\delta)) \subset \Delta(f(x_0),\varepsilon).$$

Let $x \in S \cap \Delta(x_0,\delta)$. Then

$$\|f(x) - f(x_0)\| < \varepsilon.$$

Consequently

$$f(x) \in \Delta(f(x_0),\varepsilon) \subset N(x_0)$$

and thus $T(x) = x_0$ by Theorem 3. Since $x \in S \cap \Delta(x_0,\delta)$ is arbitrary, we have in fact $T(S \cap \Delta(x_0,\delta)) = \{x_0\}$, which implies by Definition 3.2) that x_0 is in $\text{Equi}^*(S)$. ■

For the remainder of this section we assume that S is a polytope, i.e., the convex hull of finitely many points, and that $V = \{x_j\}_{1\le j\le m}$ is a minimal generating set, or minimal representation, of S. In this special case the assumption in the above theorem is easier to check. We need the following lemma which is a little technical.

Lemma 3: $N^*(x)$ is open for all $x \in V$.

Proof: Recall that $N^*(x) = x + \{u \in \mathcal{H} \mid S \cap \partial H_u = \{x\}\}$. Suppose $x = x_i$ for some $i \in \{1,\cdots,m\}$ and $u_0 \in N^*(x_i) - x_i$. We must show the existence of an $\varepsilon > 0$ so that $\Delta(u_0,\varepsilon) \subset N^*(x_i) - x_i$. To do so, we need to define the following

$$\delta_j := h(u_0) - \langle x_j, u_0\rangle, \qquad 1 \le j \le m,\; j \ne i$$
$$\rho := \frac{\min_{1\le j\le m, j\ne i}\delta_j}{2\max_{1\le j\le m}\|x_j\|}$$
$$\varepsilon^* := \min_{1\le j\le m, j\ne i}\delta_j - \rho \max_{1\le j\le m}\|x_j\| = \frac{1}{2}\min_{1\le j\le m, j\ne i}\delta_j.$$

Note that the above quantities are strictly positive. Since h is continuous by Theorem 2-1), there exists $\delta = \delta(\varepsilon^*) > 0$ so that

$$|h(z) - h(u_0)| < \varepsilon^* \qquad \forall\, z \in \Delta(u_0,\delta).$$

Let $\varepsilon := \min\{\delta, \rho\}$. Clearly $\varepsilon > 0$. We claim that this ε does the required job. Let $u \in \Delta(0,1)$. Then

$$\|(u_0 + \varepsilon u) - u_0\| = \varepsilon\|u\| < \varepsilon \le \delta$$

and by the definition of δ

$$|h(u_0 + \varepsilon u) - h(u_0)| < \varepsilon^*.$$

Now we compute for each $j \in \{1, \cdots, m\} \backslash \{i\}$

$$\begin{aligned} h(u_0 + \varepsilon u) &> h(u_0) - \varepsilon^* = h(u_0) - \min_{1 \leq j \leq m, j \neq i} \delta_j \\ &\quad + \rho \max_{1 \leq j \leq m} \|x_j\| \\ &\geq h(u_0) - \delta_j + \rho\|x_j\| = \langle x_j, u_0 \rangle + \rho\|x_j\| \\ &\geq \langle x_j, u_0 \rangle + \varepsilon\|x_j\| \\ &\geq \langle x_j, u_0 \rangle + \varepsilon\langle x_j, u \rangle \\ &= \langle x_j, u_0 + \varepsilon u \rangle. \end{aligned}$$

Note that $\|u\| \leq 1$ and so the last inequality is just the Cauchy–Schwarz inequality. Now take an arbitrary $x \in S \backslash \{x_i\}$. Then there exists $\{\alpha_j\}_{1 \leq j \leq m} \subset [0, 1]$ with $\sum_{j=1}^m \alpha_j = 1$, and $i^* \in \{1, \cdots, m\} \backslash \{i\}$ with $\alpha_{i^*} > 0$ so that $x = \sum_{j=1}^m \alpha_j x_j$. We have therefore by the above estimate that

$$\begin{aligned} \langle x, u_0 + \varepsilon u \rangle &= \sum_{j=1}^m \alpha_j \langle x_j, u_0 + \varepsilon u \rangle < \sum_{j=1}^m \alpha_j h(u_0 + \varepsilon u) \\ &= h(u_0 + \varepsilon u). \end{aligned}$$

But since equality has to hold for at least one element of S, this element must be x_i itself and we have immediately

$$S \cap \partial H_{u_0 + \varepsilon u} = \{x_i\} \qquad \forall\, u \in \Delta(0, 1).$$

Therefore

$$S \cap \partial H_u = \{x_i\} \qquad \forall\, u \in \Delta(u_0, \varepsilon).$$

This shows that $\Delta(u_0, \varepsilon) \subset N^*(x_i) - x_i$ and $N^*(x_i)$ is open. ■

It is not hard to show that for polytopes S with minimal representation $V = \{x_j\}_{j \leq 1 \leq m}$ we have Vert$(S) = V$. Using the above lemma, we can now give immediately the following corollary (compare also Corollary 1).

Corollary 3: Let $V = \{x_j\}_{1 \leq j \leq m}$ be a minimal representation of the polytope S. Then

$$f(x_i) \in N^*(x_i) \Rightarrow x_i \in \text{Equi}^*(S).$$

Also, a sufficient condition for Vert$(S) \subset$ Equi$^*(S)$ is

$$f(x_i) \in N^*(x_i) \qquad \forall\, i \in \{1, \cdots, m\}. \qquad (A_p^*)$$

Proof: This is clear by Theorem 6 and Lemma 3. ■

Corollary 4: If h is differentiable at $f(x_i) - x_i$ for all $i \in \{1, \cdots, m\}$, then

$$\begin{aligned} \nabla h(f(x_i) - x_i) = x_i \qquad &\forall\, i \in \{1, \cdots, m\} \\ &\Rightarrow \text{Vert}(S) \subset \text{Equi}^*(S). \end{aligned}$$

Proof: This is Corollary 2 with Corollary 3. ■

A demonstration of the practicality of condition (A_p^*) follows.

Example 5: Let $S = [-1, 1]^n, f(x) = x + Wx + b$. Suppose that W and b satisfy condition (S) given in Example 3. Then Vert$(S) \subset$ Equi$^*(S)$ (This is a result of Hui and Żak from [7]). We have shown in Example 3 that condition (S) is necessary.

To show this assertion recall condition (S)

$$w_{ii} > |b_i| + \sum_{j=1, j \neq i}^n |w_{ij}| \qquad \forall\, i \in \{1, \cdots, n\}. \qquad (S)$$

We have $V = E$ (see Example 3). Now assume that $(f(e) - e)_{i^*} = 0$ for some $e \in$ Vert(S) and $i^* \in \{1, \cdots, m\}$. But then we have

$$w_{i^*i^*} = \frac{b_{i^*}}{e_{i^*}} + \sum_{j=1, j \neq i^*}^n \frac{w_{i^*j}}{e_{i^*}} \leq |b_{i^*}| + \sum_{j=1, j \neq i^*}^n |w_{i^*j}|$$

contradicting condition (S). Thus the support function h is differentiable at any $e \in$ Vert(S) with (see Example 3)

$$\begin{aligned} &\nabla h(f(e) - e) \\ &\quad = \nabla h(b + We) \\ &\quad = \left(\text{sgn} \left(w_{ii} e_i + b_i + \sum_{j=1, j \neq i}^n w_{ij} e_j \right) \right)_{1 \leq i \leq n.} \end{aligned}$$

We now claim that the last expression is equal to e. Note first that by the triangle inequality

$$\begin{aligned} -|b_i| - \sum_{j=1, j \neq i}^n |w_{ij}| &\leq b_i + \sum_{j=1, j \neq i}^n w_{ij} e_j \\ &\leq |b_i| + \sum_{j=1, j \neq i}^n |w_{ij}|. \end{aligned}$$

Now, if $e_i = 1$, we have by condition (S) and the left part of the above inequality

$$\text{sgn} \left(w_{ii} e_i + b_i + \sum_{j=1, j \neq i}^n w_{ij} e_j \right) = 1 = e_i$$

and if $e_i = -1$, we multiply condition (S) by (-1) and use the right part of the above inequality to obtain

$$\text{sgn} \left(w_{ii} e_i + b_i + \sum_{j=1, j \neq i}^n w_{ij} e_j \right) = -1 = e_i.$$

This proves the claim and we have

$$\nabla h(f(e) - e) = e \qquad \forall\, e \in \text{Vert}(S)$$

and therefore Vert$(S) \subset$ Equi$^*(S)$ by Corollary 4. ■

Our last goal is now to give explicit conditions on the matrix W such that Vert$(S) \subset$ Equi(S) or Vert$(S) \subset$ Equi$^*(S)$ is true if S is a polytope and if $f(x) = x + Wx$. To find such conditions, we first need to compute the support function of S in the case when S is a polytope.

Lemma 4: Let $V = \{x_i\}_{1 \leq i \leq m}$ be a minimal representation of the polytope S. Then

$$h(u) = \max_{1 \leq i \leq m} \langle x_i, u \rangle \qquad \forall\, u \in \mathcal{H}.$$

Proof: A simple calculation shows that for $u \in \mathcal{H}$

$$\begin{aligned} h(u) &= \sup_{s\in S}\langle s, u\rangle \\ &= \sup\left\{\left\langle \sum_{i=1}^{m}\alpha_i x_i, u\right\rangle \middle| \sum_{i=1}^{m}\alpha_i = 1,\ 0 \le \alpha_i \le 1\right\} \\ &= \sup\left\{\sum_{i=1}^{m}\alpha_i\langle x_i, u\rangle \middle| \sum_{i=1}^{m}\alpha_i = 1, 0 \le \alpha_i \le 1\right\} \\ &\le \max_{1\le i\le m}\langle x_i, u\rangle \\ &\quad \sup\left\{\sum_{i=1}^{m}\alpha_i \middle| \sum_{i=1}^{m}\alpha_i = 1, 0 \le \alpha_i \le 1\right\} \\ &= \max_{1\le i\le m}\langle x_i, u\rangle \le h(u). \end{aligned}$$

Thus it follows that $h(u) = \max_{1\le i\le m}\langle x_i, u\rangle$. ∎

With Lemma 4 we can now rewrite the conditions given in the last section. This is done in the following:

Corollary 5: Consider the conditions:

$$\langle x_j, f(x_j) - x_j\rangle = \max_{i\in\{1,\cdots,m\}}\langle x_i, f(x_j) - x_j\rangle \quad \forall\, j \in \{1,\cdots,m\}. \qquad (A)$$

$$\langle x_j, f(x_j) - x_j\rangle > \max_{i\in\{1,\cdots,m\}\setminus\{j\}}\langle x_i, f(x_j) - x_j\rangle \quad \forall\, j \in \{1,\cdots,m\}. \qquad (A_p^*)$$

Then condition (A) is equivalent to $\mathrm{Vert}(S) \subset \mathrm{Equi}(S)$ and condition (A_p^*) implies $\mathrm{Vert}(S) \subset \mathrm{Equi}^*(S)$.

Proof: Lemma 4 with Corollary 1 and Corollary 3. ∎

Finally, let us consider linear functions of the form $f(x) = x + Wx$ where W is a linear operator on $\mathcal{H}$. In this case we have immediately from the above corollary:

Corollary 6: Assume $f(x) = x + Wx$. Then conditions (A) and (A_p^*) have the form

$$\langle x_j, Wx_j\rangle = \max_{i\in\{1,\cdots,m\}}\langle x_i, Wx_j\rangle \quad \forall\, j \in \{1,\cdots,m\}. \ (A)$$

$$\langle x_j, Wx_j\rangle > \max_{i\in\{1,\cdots,m\}\setminus\{j\}}\langle x_i, Wx_j\rangle \quad \forall\, j \in \{1,\cdots,m\}. \qquad (A_p^*)$$

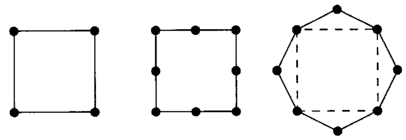


Fig. 2. Illustration of our approach.

Proof: Corollary 5. ∎

V. Application to Associative Memory

We now propose an approach to attack the gray scale problem mentioned in the introduction. We have to assume that at least one pixel (or coordinate) attains the maximum or minimum possible values. This assumption ensures that all desired equilibrium points are on the boundary of a convex set.

To illustrate our approach, consider the case of only two pixels. If the pixels can only be on or off, we have four possible values: (1, 1), $(1,-1), (-1,-1)$, and $(-1,1)$ [see Fig. 2(a)]. Suppose we now desire an intermediate value zero. Then, with our assumption, there are eight possible values: (0, 1), (1, 1), (1, 0), $(1,-1), (0,-1), (-1,-1), (-1,0)$, and $(-1,1)$ [see Fig. 2(b)]. These are not all vertices and so we perturb the nonvertices to obtain a regular polytope with eight sides [see Fig. 2(c)]. Our theory now applies, and we can find suitable W such that the vertices are stable equilibrium points.

The same idea applies when there are more pixels with more intermediate values. We next give a numerical example to illustrate our idea.

Example 6: Let

$$x_1 = \begin{pmatrix} 0 \\ 1.1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1.1 \\ 0 \end{pmatrix},$$
$$x_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 0 \\ -1.1 \end{pmatrix}, \quad x_5 = \begin{pmatrix} -1 \\ -1 \end{pmatrix},$$
$$x_7 = \begin{pmatrix} -1.1 \\ 0 \end{pmatrix}, \quad x_8 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

be the vertices of a regular polytope and let

$$X = \begin{pmatrix} 0 & 1 & 1.1 & 1 & 0 & -1 & -1.1 & -1 \\ 1.1 & 1 & 0 & -1 & -1.1 & -1 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \boxed{14.52} & 14.3 & 1.21 & -12.1 & -14.52 & -14.3 & -1.21 & 12.1 \\ 14.3 & \boxed{26} & 14.3 & 0 & -14.3 & -26 & -14.3 & 0 \\ 1.21 & 14.3 & \boxed{14.52} & 12.1 & -1.21 & -14.3 & -14.52 & -12.1 \\ -12.1 & 0 & 12.1 & \boxed{22} & 12.1 & 0 & -12.1 & -22 \\ -14.52 & -14.3 & -1.21 & 12.1 & \boxed{14.52} & 14.3 & 1.21 & -12.1 \\ -14.3 & -26 & -14.3 & 0 & 14.3 & \boxed{26} & 14.3 & 0 \\ -1.21 & -14.3 & -14.52 & -12.1 & 1.21 & 14.3 & \boxed{14.52} & 12.1 \\ 12.1 & 0 & -12.1 & -22 & -12.1 & 0 & 12.1 & \boxed{22} \end{pmatrix}$$

be the matrix whose columns are $x_1, \cdots, x_8$. With

$$W = \begin{pmatrix} 12 & 1 \\ 1 & 12 \end{pmatrix}$$

we can compute $X^T W X$ to be the matrix shown at the bottom of the preceding page. We see that condition

$$\langle x_j, Wx_j \rangle > \max_{i \in \{1,\cdots,m\} \setminus \{j\}} \langle x_i, Wx_j \rangle \ \forall\, j \in \{1, \cdots, m\} \qquad (A_p^*)$$

of Corollary 6 is satisfied. Therefore, all vertices are stable equilibrium points of our system. ■

VI. Conclusion

We studied the BSB model on general convex bodies. We gave necessary and sufficient conditions for vertices to be equilibrium points and sufficient conditions for vertices to be stable equilibrium points in the generalized system. These results can be used in the study of associative memory problems as shown in Section V, where we proposed an approach to allow a gray scale in the pixels. The results here also contain as special cases the main results in [3] and [7].

References

[1] J. Anderson, J. Silverstein, S. Ritz, and R. Jones, "Distinctive features, categorical perception, and probability learning: Some applications of a neural model," *Psychological Rev.*, vol. 84, pp. 413–451, 1977.

[2] R. M. Golden, "The brain-state-in-a-box neural model is a gradient decent algorithm," *J. Math Psych.*, vol. 30, pp. 73–80, 1986.

[3] H. J. Greenberg, "Equilibria of the brain-state-in-a-box (BSB) neural model," *Neural Networks*, vol. 1, pp. 323–324, 1988.

[4] S. Grossberg, "Nonlinear neural networks: Principles, mechanisms, and architectures," *Neural Networks*, vol. 1, pp. 17–61, 1988.

[5] M. Hassoun, *Associative Neural Memories*. New York: Oxford University Press, 1993.

[6] A. Hertz, J. Krogh, and R. Palmer, *Introduction to the Theory of Neural Computation*. Redwood City, CA: Addison-Wesley, 1991.

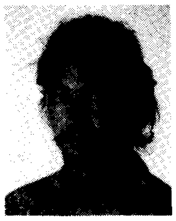
[7] Hui, S. and Żak, S., "Dynamical analysis of the BSB neural model," *IEEE Trans. Neural Networks*, vol. 3, pp. 86–94, 1992.

[8] S. Lay, *Convex Sets and their Applications*. New York: Wiley, 1982.

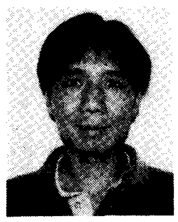
[9] W. Roberts and D. Varberg, *Convex Functions*. New York: Academic, 1973.

[10] J. Weidmann, *Linear Operators in Hilbert Spaces*. New York: Springer-Verlag, 1980.

[11] N. Young, *An Introduction to Hilbert Space*. Cambridge: Cambridge University Press, 1988.



Martin Bohner received the Diplom degree in business mathematics (mathematics, economics, computer science, statistics, and operations research) from the University of Ulm, Germany, in 1993. He was an exchange student at San Diego State University for one year, where he received the M.S. degree in applied mathematics with a thesis on a generalization of the BSB Model in 1992. Currently he is working on a Ph.D. degree in mathematics in the field of linear Hamiltonian Difference Systems at the University of Ulm.



Stefen Hui received the B.A. degree from the University of California at Berkeley and his M.S. and Ph.D. degrees from the University of Washington in Seattle, both in mathematics.

He worked as a scientist at the Naval Ocean Systems Center in San Diego and as a Research Assistant Professor at Purdue University in West Lafayette, Indiana, before joining the Department of Mathematical Sciences at San Diego State University in 1988 where he is currently a Professor of Mathematics. His research interests are in the areas of neural computation, control theory, signal processing, and function theory.