Fixed points of complex-valued bidirectional associative memory

Xiaosheng Zhu, Wei Wu

School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Keywords:
Complex-valued neural networks
Bidirectional associative memory
Fixed point

ABSTRACT

A discrete complex-valued bidirectional associative memory (CVBAM) is considered. The input and output of CVBAM are bipolar complex numbers, of which the two components take the value of either −1 or 1. We prove that the stored patterns are all fixed points of CVBAM provided certain sufficient conditions are satisfied. It is shown that each fixed point belongs to a fixed point group which consists of four fixed points. Supporting numerical results are provided.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

With the development of signal and image processing involving complex-valued data, there has been an increasing demand for complex-valued neural networks (CVNN) [1], of which the input, output, activation function and connection weights are complex numbers. The property and performance of CVNN have been widely explored. The back-propagation algorithm and its improved version for the complex-valued multi-layer perceptrons (MLP) are proposed in [2–4]. The approximation ability of complex-valued MLP is investigated in [5]. The complex-valued radial basis function (RBF) network together with its learning algorithm is proposed in [6,7]. Complex-valued MLP and RBF are successfully applied to several problems such as classification, signal processing [8] and image recognition [9]. Apart from the investigation on the complex-valued feed forward networks, the complex-valued recurrent neural networks also have been studied. A complex-valued multi-state Hopfield neural network (HNN) has been presented in [10,11], which can be used to reconstruct noisy gray-scale images. The stability of complex-valued multi-state bidirectional associative memory (BAM) and its learning algorithm is studied in [12]. The most significant feature of CVNN is that they are more suitable to process the phase information of complex-valued signals than real-valued neural networks [13], and complex-valued recurrent neural networks can more efficiently deal with gray-scale images [10–12].

The first requirement for the BAM is that each stored pattern should be perfectly recalled when it is supplied to the network as an input, i.e., each stored pattern should be a fixed point of the network. In [14], the fixed points of real-valued BAM are discussed, and a sufficient condition is given to ensure all the stored patterns to be fixed points of the network. In this paper, we investigate the property of fixed points of CVBAM, and specify certain sufficient condition for the fixed points. It is also shown that each fixed point belongs to a fixed point group which consists of four fixed points.

The rest of the paper is organized as follows. In Section 2, we introduce the CVBAM model and its bidirectional associative memory procedure. In Section 3, we discuss the fixed point property of CVBAM. In Section 4, a few numerical experiments are carried out to support our theoretical results.
2. Complex-valued BAM

We use $\mathbb{C}^m_{[-1,1]}$ to denote a m dimensional bipolar complex-valued vector space, which is defined by

$$\mathbb{C}^m_{[-1,1]} = \{(a_1, a_2, \ldots, a_m) \in \mathbb{C}^m | \operatorname{Re}(a_p), \operatorname{Im}(a_p) \in (-1, 1), p = 1, 2, \ldots, m\}. \quad (1)$$

Let $\{(a'_1, b'_1), (a'_2, b'_2), \ldots, (a'_m, b'_m)\}$ be a given set of sample patterns, where $a'_i = (a'_1, a'_2, \ldots, a'_m) \in \mathbb{C}^m_{[-1,1]}$, $b'_i = (b'_1, b'_2, \ldots, b'_m) \in \mathbb{C}^m_{[-1,1]}$. By the Hebbian rule we define the weight matrix of the network as follows. The connection weight from neuron $p$ in $L_b$ layer to neuron $q$ in $L_a$ layer is

$$s_{qp} = \sum_{j=1}^{L} b'_j (a'_p)^*, \quad (2)$$

and the connection weight from neuron $q$ in $L_b$ layer to neuron $p$ in $L_a$ layer is

$$u_{pq} = \sum_{j=1}^{L} a'_j (b'_q)^*. \quad (3)$$

From (2) and (3), it is easy to know that $s_{qp} = u_{qp}^*$. The complex activation function of the network is defined as

$$f(\xi) = \text{sgn}(\operatorname{Re}(\xi)) + \text{sgn}(\operatorname{Im}(\xi))i, \quad (4)$$

where $\text{sgn}(\cdot)$ is a sign function

$$\text{sgn}(\alpha) = \begin{cases} 1, & \text{for } \alpha \geq 0 \\ -1, & \text{for } \alpha < 0 \end{cases}. \quad (5)$$

Let the network initial state in $L_a$ layer is $a(0) = (a_1(0), a_2(0), \ldots, a_m(0))^T \in \mathbb{C}^m_{[-1,1]}$, and set $t = 0$. Then, the memory procedure goes as follows.

Step 1. Update the net input $y_q$ and the output $b_q(t)$ of the $q$th neuron in $L_a$ layer,

$$\begin{cases} y_q = \sum_{p=1}^{m} s_{qp}a_p(t), \\ b_q(t) = f(y_q), \quad q = 1, 2, \ldots, n. \end{cases} \quad (6)$$

Step 2. Update the net input $x_p$ and the output $a_p(t + 1)$ of the $p$th neuron in $L_a$ layer,

$$\begin{cases} x_p = \sum_{q=1}^{n} u_{pq}b_q(t), \\ a_p(t + 1) = f(x_p), \quad p = 1, 2, \ldots, m. \end{cases} \quad (7)$$

Step 3. If $a(t + 1) = a(t)$, stop the iteration, and take $b(t)$ as the association of $a(0)$; else set $t = t + 1$, and turn to step 1.

3. Fixed point of CVBAM

We first introduce the definition of the fixed point of the network.

**Definition 1.** If a given pattern $(a, b)$ with its net inputs $x_p = \sum_{q=1}^{n} u_{pq}b_q$, $y_q = \sum_{p=1}^{m} s_{qp}a_p$ satisfies

$$\begin{cases} a_p = f(x_p), \quad p = 1, 2, \ldots, m, \\ b_q = f(y_q), \quad q = 1, 2, \ldots, n, \end{cases} \quad (8)$$

then $(a, b)$ is called a fixed point of the network.

In what follows, we first present two lemmas. Their proofs are straightforward and omitted. Then, we use them to prove our main theorem on the fixed points.

**Lemma 1.** For the network presented in Section 2, if a given pattern $(a, b)$ with its net inputs satisfies

$$\begin{cases} \operatorname{Re}(a_p)\operatorname{Re}(x_p) > 0, \\ \operatorname{Im}(a_p)\operatorname{Im}(x_p) > 0, \quad p = 1, 2, \ldots, m, \\ \operatorname{Re}(b_q)\operatorname{Re}(y_q) > 0, \\ \operatorname{Im}(b_q)\operatorname{Im}(y_q) > 0, \quad q = 1, 2, \ldots, n, \end{cases} \quad (9)$$

then it is a fixed point of the network.
Lemma 2. For complex numbers \( u \) and \( v \), we have
\[
\begin{align*}
\text{Re}(u)\text{Re}(v) &= \frac{1}{4}(uv^* + u^*v + uv + u^*v^*), \\
\text{Im}(u)\text{Im}(v) &= \frac{1}{4}(uv^* + u^*v - uv - u^*v^*).
\end{align*}
\] (10)

Theorem 1. For the stored patterns \( \{(a_j^k, b_j^k)\}_{j=1}^l \), write
\[
\begin{align*}
K_A &= \max_{1 \leq k \leq j} \left\{ \frac{\sqrt{2}}{2} \sum_{j=1}^l \sum_{p=1}^m |(a_p^k)_j|^2 \right\}, \\
K_B &= \max_{1 \leq k \leq j} \left\{ \frac{\sqrt{2}}{2} \sum_{j=1}^l \sum_{p=1}^m |(b_p^k)_j|^2 \right\}.
\end{align*}
\] (11) (12)

If \( K_A < m \) and \( K_B < n \), then each stored pattern is a fixed point of the network.

Proof. For any pattern \( \{(a_j^k, b_j^k)\}, k = 1, 2, \ldots, J \), denote its net input in \( L_\lambda \) layer by
\[
x_p^k = \sum_{q=1}^n u_{pq} b_q^k = \sum_{q=1}^n \sum_{j=1}^l a_p^j (b_q^j)^* b_q^k
\]
\[
= \sum_{q=1}^n \left[ a_p^k (b_q^k)^* b_q^k + \sum_{j \neq k} a_p^j (b_q^j)^* b_q^k \right]
\]
\[
= 2n a_p^k + \sum_{j \neq k} a_p^j \sum_{q=1}^n (b_q^j)^* b_q^k
\]
\[
= 2n a_p^k + C,
\] (13)

where \( C = \sum_{j \neq k} a_p^j \sum_{q=1}^n (b_q^j)^* b_q^k \). Then we have
\[
(a_p^k)^* x_p^k = 2n (a_p^k)^* a_p^k + (a_p^k)^* C = 4n + (a_p^k)^* C,
\] (14)
\[
(a_p^k)^* x_p^k = 4n + (a_p^k)^* C,
\] (15)
\[
a_p^k x_p^k = 2n (a_p^k)^2 + a_p^k C,
\] (16)
\[
(a_p^k)^* x_p^k = 2n (a_p^k)^* + (a_p^k)^* C^*.
\] (17)

For \( C = \sum_{j \neq k} a_p^j \sum_{q=1}^n (b_q^j)^* b_q^k \), we have
\[
|C| = \sum_{j \neq k} |a_p^j| \sum_{q=1}^n (b_q^j)^* b_q^k \leq \sum_{j \neq k} \sum_{q=1}^n |a_p^j| (b_q^j)^* b_q^k \leq 2K_B.
\] (18)

Using Lemma 2 and (14)-(18), we have
\[
\begin{align*}
\text{Re}(a_p^k)\text{Re}(x_p^k) &= \frac{1}{4} \left[ (a_p^k)^* x_p^k + (a_p^k)^* x_p^k + a_p^k x_p^k + (a_p^k)^* x_p^k \right] \\
&= \frac{1}{4} \left[ 18n + 2n \left( (a_p^k)^2 + (a_p^k C)^2 \right) + \left( (a_p^k)^* C + (a_p^k)^* C^* \right) + (a_p^k C + (a_p^k)^* C^*) \right] \\
&= \frac{1}{4} \left( 8n + \left[ (a_p^k)^* C + (a_p^k C)^* \right] + \left[ (a_p^k)^* C + (a_p^k)^* C^* \right] \right) \\
&= \left\{ 2n - \text{Re}(a_p^k) \text{Re}(C) \right\} \\
&\geq \left\{ 2n - |\text{Re}(a_p^k)| |\text{Re}(C)| \right\}
\end{align*}
\]
Similarly, we have
\[
\text{Im}(\alpha_p^k) \text{Im}(\alpha_q^k) = \frac{1}{4} \left[ (\alpha_p^k)^* x_p^k + \alpha_p^k (x_p^k)^* - (\alpha_q^k)^* x_q^k - \alpha_q^k x_q^k \right]
\]
\[
= \left\{ 2n - \text{Im}(\alpha_p^k) \text{Im}(C) \right\}
\]
\[
\geq 2\{n - K_b\}
\]
\[> 0.\]

We can also prove that
\[
\text{Re}(b_q^k) \text{Re}(y_q^k) \geq 2\{m - K_a\} > 0,
\]
\[
\text{Im}(b_q^k) \text{Im}(y_q^k) \geq 2\{m - K_a\} > 0.
\]

Thus, \((\alpha^k, b^k)\) is a fixed point by using Lemma 1. We thus complete the proof. □

Next, we turn to investigate the group of fixed points. For \(\xi \in \mathbb{C}_m^{n-1,1}\) and \(\eta \in \mathbb{C}_m^{n-1,1}\), write
\[
G_1(\xi, \eta) = \{ (\xi, \eta), -(\xi, \eta), (i\xi, i\eta), -(i\xi, i\eta) \},
\]
\[
G_2(\xi) = \{ \xi, -\xi, i\xi, -i\xi \}.
\]

We note that \(G_1(\xi, \eta)\) and \(G_2(\xi)\) are closed, respectively, i.e., for \(u, v \in G_1(\xi, \eta), z \in G_2(\xi)\), we have \(G_1(u, v) = G_1(\xi, \eta), G_2(z) = G_2(\xi)\).

**Theorem 2.** If a pattern \((a, b)\) is a fixed point of the network and its net inputs satisfy
\[
\begin{align*}
\text{Re}(x_p) \text{Im}(x_p) & \neq 0, \quad p = 1, 2, \ldots, m, \\
\text{Re}(y_q) \text{Im}(y_q) & \neq 0, \quad q = 1, 2, \ldots, n,
\end{align*}
\]
then any pattern in \(G_1(a, b)\) is a fixed point and \(G_1(a, b)\) is called a fixed point group.

**Proof.** Since \((a, b)\) is a fixed point, it follows from (21) that
\[
\begin{align*}
\text{Re}(a_p) \text{Re}(x_p) & > 0, \quad \text{Im}(a_p) \text{Im}(x_p) > 0, \quad p = 1, 2, \ldots, m, \\
\text{Re}(b_q) \text{Re}(y_q) & > 0, \quad \text{Im}(b_q) \text{Im}(y_q) > 0, \quad q = 1, 2, \ldots, n.
\end{align*}
\]

Write \((\bar{a}, \bar{b}) = (ia, ib)\) and let \(\bar{x}_p, \bar{y}_q\) be the corresponding net inputs, then
\[
\bar{x}_p = \sum_{q=1}^{n} w_{pq} \bar{b}_q = \sum_{q=1}^{n} w_{pq}(ib_q) = i \sum_{q=1}^{n} w_{pq} b_q = ix_p,
\]
\[
\bar{y}_q = \sum_{p=1}^{m} s_{qp} \bar{a}_p = \sum_{p=1}^{m} s_{qp}(ia_p) = i \sum_{p=1}^{m} s_{qp} a_p = iy_q.
\]

In terms of Lemma 2, we have
\[
\text{Re}(\bar{a}_p) \text{Re}(\bar{x}_p) = \frac{1}{4} (\bar{a}_p^* \bar{x}_p + \bar{a}_p \bar{x}_p^* + \bar{a}_p \bar{x}_p + \bar{a}_p^* \bar{x}_p^*)
\]
\[
= \frac{1}{4} (a_p^* x_p + a_p x_p^* - a_p x_p - a_p^* x_p^*)
\]
\[
= \text{Im}(a_p) \text{Im}(x_p) > 0,
\]
\[
\text{Im}(\bar{a}_p) \text{Im}(\bar{x}_p) = \frac{1}{4} (\bar{a}_p^* \bar{x}_p + \bar{a}_p \bar{x}_p^* - \bar{a}_p \bar{x}_p - \bar{a}_p^* \bar{x}_p^*)
\]
\[
= \frac{1}{4} (a_p^* x_p + a_p x_p^* - a_p x_p - a_p^* x_p^*)
\]
\[
= \text{Re}(a_p) \text{Re}(x_p) > 0.
\]

Similarly, we have
\[
\text{Re}(\bar{b}_q) \text{Re}(\bar{y}_q) = \text{Im}(b_q) \text{Im}(y_q) > 0,
\]
\[
\text{Im}(\bar{b}_q) \text{Im}(\bar{y}_q) = \text{Re}(b_q) \text{Re}(y_q) > 0.
\]
By Lemma 1, \((ia, ib)\) is a fixed point. Since \((-a, -b) = (i \cdot ia, i \cdot ib)\), \((-a, -b)\) is also a fixed point. Finally, \((-ia, -ib)\) is a fixed point as well. This completes the proof. □

**Theorem 3.** If the stored patterns \(\{(a', b')\}_j\) satisfy \(K_A < m\) and \(K_B < n\), then \(\{G_1(a', b')\}_j\) are fixed point groups of the network.

**Proof.** By Theorem 1, \(\{(a', b')\}_j\) are fixed points and
\[
\text{Re}(a'_p)\text{Re}(x'_p) > 0, \quad \text{Im}(a'_p)\text{Im}(x'_p) > 0, \quad p = 1, 2, \ldots, m,
\]
\[
\text{Re}(b'_q)\text{Re}(y'_q) > 0, \quad \text{Im}(b'_q)\text{Im}(y'_q) > 0, \quad q = 1, 2, \ldots, n.
\]
Consequently,
\[
\text{Re}(x'_p)\text{Im}(x'_p) \neq 0, \quad \text{Re}(y'_q)\text{Im}(y'_q) \neq 0.
\]
Using Theorem 2, we see that \(\{G_1(a', b')\}_j\) are fixed point groups. □

Note that if the patterns \(\{(a', b')\}_j\) satisfy \(a' = b'\), CVBAM degenerates into a synchronous complex-valued Hopfield neural network (CVHNN) with non-negative diagonal elements in the weight matrix. So our theoretical results about fixed points are valid in the case of synchronous CVHNN. Furthermore, due to the one-layer structure of CVHNN the definition of the fixed point in CVHNN is actually independent of the network updating mode, no matter synchronous or asynchronous. So the results about fixed points also hold in the case of asynchronous CVHNN. Let \(a' \in \mathbb{C}^m_{\{(-1, 1\}_j=1}\} \) be the stored patterns in CVHNN, then we have the following corollaries.

**Corollary 1.** If the stored patterns \(\{a'\}_j\) satisfy \(K_A < m\), then each stored pattern is a fixed point of CVHNN.

**Corollary 2.** If the pattern \(a\) is a fixed point of CVHNN and its net inputs satisfy
\[
\text{Re}(x_p)\text{Im}(x_p)\neq 0, \quad p = 1, 2, \ldots, m,
\]
then \(G_2(a)\) is a fixed point group.

**Corollary 3.** If the stored patterns \(\{a'\}_j\) satisfy \(K_A < m\), then \(\{G_2(a')\}_j\) are fixed point groups of CVHNN.

### 4. Numerical examples

#### 4.1. A simple example

First we consider a network with 4 neurons in \(L_A\) layer and 3 neurons in \(L_B\) layer. We embed three patterns showed in Table 1 into the network. The following connection weight matrix \(\{w_{pq}\}\) is learned by using the Hebbian rule:

\[
\{w_{pq}\} = \begin{pmatrix}
2i & 2i & 4 + 2i \\
-4 + 2i & 2i & -2i \\
4 + 2i & -2i & -4 - 2i \\
-2i & 4 - 2i & 4 + 2i
\end{pmatrix}.
\]

For this network, we have \(K_A = 2.83 < 4\) and \(K_B = 2.83 < 3\). So the conditions in Theorem 1 are satisfied. A direct check shows that all the three patterns are fixed points, and that \(\{G_1(a', b')\}_j\) are fixed point groups, just as predicted by Theorem 1.

#### 4.2. More examples

Note that the result in Theorem 1 is a sufficient condition for the fixed points. We shall verify it pro and con. On one hand, if the stored patterns satisfy \(K_A < m\) and \(K_B < n\), then by Theorem 1 all the patterns must be fixed points of the network. On the other hand, if there exist one or more stored patterns which are not fixed points, then by Theorem 1, either \(K_A < m\) or \(K_B < n\) is not satisfied.

### Table 1

<table>
<thead>
<tr>
<th>(j)</th>
<th>(a')</th>
<th>(b')</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((-1 - i, -1 + i, -1 - i + i)^T)</td>
<td>((-1 + i, -1 + i, 1 + i)^T)</td>
</tr>
<tr>
<td>2</td>
<td>((-1 + i, -1 + i, -1 + i + i)^T)</td>
<td>((1 - i, 1 + i, 1 + i)^T)</td>
</tr>
<tr>
<td>3</td>
<td>((-1 + i, 1 - i, 1 + i)^T)</td>
<td>((-1 + i, 1 + i, 1 + i)^T)</td>
</tr>
</tbody>
</table>
We use two networks with different number of neurons and stored patterns, one with $m = 30$, $n = 20$ and $J = 4$, and the other with $m = 50$, $n = 60$ and $J = 7$. The experimental results are showed in Tables 2 and 3, respectively. We see that if the stored patterns satisfy $K_A < m$ and $K_B < n$, then they are all fixed points, represented by the symbol “√”. And if there exist one or more stored patterns which are not fixed points, represented by the symbol “×”, then either $K_A < m$ or $K_B < n$ is not satisfied.

Similar numerical results have also been obtained to support our Theorems 2 and 3 by using the two networks. But, to save the space, they are not included in the paper.

References