On the Decision Boundaries of Hyperbolic Neurons

Tohru Nitta and Sven Buchholz

Abstract— **In this paper, the basic properties, especially decision boundary, of the hyperbolic neurons used in the hyperbolic neural networks are investigated. And also, a non-split hyperbolic sigmoid activation function is proposed.**

I. Introduction

Hyperbolic neural network as has been proposed in 2000 [1] is one possible extension of usual real-valued neural networks to two dimensions. It is based on hyperbolic numbers which are a counterpart to complex numbers. Complex–valued neural networks have found many applications in recent years [3]. Deriving a general theory of decision boundaries in both the complex and hyperbolic case will be beneficial for further applications. This is the main motivation for our study. As a first step towards this goal, this paper studies basic properties of the *hyperbolic neuron* as being used in the hyperbolic neural network.

The main results can be summarized as follows. Weight parameters of a hyperbolic neuron are coupled by a certain constraint, and learning proceeds under this restriction. This is analogous to complex neurons [8]. The decision boundary of a hyperbolic neuron consists of two hypersurfaces which can intersect orthogonally or be parallel depending on the values of the hyperbolic-valued weight vectors. Note that intersection is always orthogonally for complex neurons. A non-split hyperbolic sigmoid activation function is proposed, which is analytic and bounded in the hyperbolic plane unlike the non-split (fully) complex-valued sigmoid activation function [4]. Using this function is a significant extension of the model in [1].

As the result of this study, it was learned that one of the advantages of hyperbolic neurons is that the angle between the decision boundary for the real-part and that for the unipotent part can be easily controlled by changing the weight parameters. And also, the proposed analytic and bounded non-split hyperbolic sigmoid activation function offers the advantage of uncritical computation, since operations for avoiding poles (which can have fatal impacts on learning) are not needed unlike in the complex-valued case.

II. Hyperbolic Numbers

Hyperbolic numbers, blood relatives of the popular complex numbers, are numbers of the form

$$w = a + ub \tag{1}$$

Tohru Nitta is with the Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba-shi, Ibaraki 305-8568, Japan (email: tohru-nitta@aist.go.jp).

Sven Buchholz is with Institut fuer Informatik, Universitaet Kiel, Olshausenstrasse 40, D-24098 Kiel, Germany (email: sbh@ks.informatik.uni-kiel.de).

where $a, b \in \boldsymbol{R}$ and u is called *unipotent* which has the algebraic property that $u \neq \pm 1$ but $u^2 = 1$ [9]. It follows that multiplication in $\boldsymbol{H}$ is defined by $(a+ub)(c+ud) = (ac+bd)+u(ad+bc)$ where $\boldsymbol{H}$ denotes the set of hyperbolic numbers. A hyperbolic number can be identified with a point in the plane $\boldsymbol{R}^2$ (called *hyperbolic number plane*). The hyperbolic modulus of $w = a+ub \in \boldsymbol{H}$ is defined by $|w|_h \stackrel{\text{def}}{=} \sqrt{|a^2-b^2|}$. If $|a| = |b|$ then w has zero modulus. Hence $\boldsymbol{H}$ contains divisors of zero. $\boldsymbol{H}$ is isomorphic to the matrix algebra generated by

$$\left\{ \left[\begin{array}{cc} a & b \\ b & a \end{array} \right] \middle| \, a, b \in \boldsymbol{R} \right\}. \tag{2}$$

This representation allows for a geometric interpretation of hyperbolic numbers. Another equivalent representation

$$\left\{ \left[\begin{array}{cc} a+b & 0 \\ 0 & a-b \end{array} \right] \middle| \, a, b \in \boldsymbol{R} \right\} \tag{3}$$

gives rise to a direct formulation of the hyperbolic exponential function. The matrix exponential of a diagonal matrix is given by the matrix that results from applying the exponential to each entry in the diagonal. Thus we get from Eq. (3)

$$\exp(a+ub) = \left[\begin{array}{cc} \exp(a+b) & 0 \\ 0 & \exp(a-b) \end{array} \right]. \tag{4}$$

Using hyperbolic trigonometric functions and extracting the common factor $\exp(a)$ gives

$$\exp(a+ub) = \exp(a)\,(\cosh(b) + u\sinh(b)). \tag{5}$$

III. The Hyperbolic Neuron and Its Properties

A. The Hyperbolic Neuron Model

This section briefly describes the hyperbolic neuron model proposed in [1]. The input and output signals, and weights and threshold values of a hyperbolic neuron are all hyperbolic numbers, and the activation function $f_H^{(1)}$ of a hyperbolic neuron is defined to be

$$f_H^{(1)}(z) = f_R(x) + u f_R(y), \tag{6}$$

where $z = x + uy$ is a hyperbolic-valued net input to the hyperbolic neuron, and $f_R(s) = 1/(1+\exp(-s))$ for $s \in \boldsymbol{R}$. That is, the real and unipotent parts of the hyperbolic-valued output of a hyperbolic neuron mean the sigmoid functions of the real part x and unipotent part y of the net input z to neuron, respectively.

Although Eq. (6) is used as an activation function of the hyperbolic neuron in the above, various types of activation functions other than Eq. (6) can be considered naturally.

978-1-4244-1821-3/08/$25.00©2008 IEEE

B. Non-Split Hyperbolic Activation Function

In this section, a non-split activation function is proposed, which is defined as

$$f_H^{(2)}(z) \stackrel{\text{def}}{=} \frac{1}{1+e^{-x}\cdot(\cosh y - u\sinh y)} \qquad (7)$$

for any $z = x + uy \in \boldsymbol{H}$ where $\cosh y - u \sinh y$ lies in the hyperbolic quadrant $H - I$ of the hyperbolic number plane. Eq. (7) can be rewritten as

$$f_H^{(2)}(z) = \frac{1}{1+e^{-z}}, \qquad (8)$$

using the formula $\cosh\phi + u\sinh\phi = e^{u\phi}$. Note that the hyperbolic exponential e^{-z} is analytical. Eq. (7) is bounded because for any $z = x + uy \in \boldsymbol{H}$,

$$\begin{aligned} |f_H^{(2)}(z)|_h^2 &= \frac{1}{1+e^{-x+y}+e^{-(x+y)}+e^{-2x}} \\ &\le 1. \end{aligned} \qquad (9)$$

Thus, Eq. (7) has no poles unlike the non-split (fully) complex-valued activation function [4]:

$$f_C(z) = \frac{1}{1+e^{-z}}, \qquad z = x + iy \qquad (10)$$

where $i = \sqrt{-1}$. As is well known ([2], [5], [7]), one can only choose either the regularity or the boundedness for an activation function of complex-valued neurons. This is due to Liouville's theorem, which says that if a function G is regular at all $z \in \boldsymbol{C}$ and bounded, then G is a constant function where $\boldsymbol{C}$ denotes the set of complex numbers. However, there is no such conflict in the case of hyperbolic neurons.

C. Weight Parameters of a Hyperbolic Neuron

Next, we examine the basic structures of weights of a hyperbolic neuron. Consider a hyperbolic neuron with N-inputs, weights $h_k = v_k + uw_k \in \boldsymbol{H}$ $(1 \le k \le N)$, and a threshold value $\theta = c + ud \in \boldsymbol{H}$. Then, for N input signals $x_k + uy_k \in \boldsymbol{H}$ $(1 \le k \le N)$, the hyperbolic neuron generates

$$X + uY = f_H(S + uT), \qquad (11)$$

as an output where f_H is a hyperbolic-valued activation function, and

$$\begin{aligned} S+uT &= \sum_{k=1}^{N}(v_k+uw_k)(x_k+uy_k) + (c+ud) \\ &= \left(\sum_{k=1}^{N}(v_k x_k + w_k y_k)\right) + c \\ &\qquad + u\left(\sum_{k=1}^{N}(w_k x_k + v_k y_k) + d\right) \\ &= \left([\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + c\right) \\ &\qquad + u\left([\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + d\right) \end{aligned} \qquad (12)$$

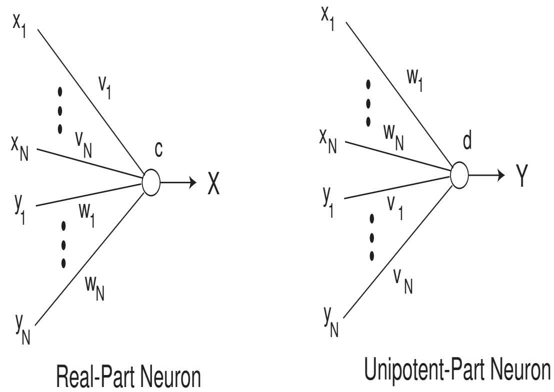

Fig. 1. Two real-valued neurons which are equivalent to a hyperbolic neuron.

where $\boldsymbol{v} = (v_1 \cdots v_N)^t$, $\boldsymbol{w} = (w_1 \cdots w_N)^t$, $\boldsymbol{x} = (x_1 \cdots x_N)^t$ and $\boldsymbol{y} = (y_1 \cdots y_N)^t$. Hence, in the case of the split-type activation function such as Eq. (6), the hyperbolic neuron with N-inputs is equivalent to two real-valued neurons with $2N$-inputs in Fig. 1. We shall refer to a real-valued neuron corresponding to the real part X of an output of a hyperbolic neuron as a *Real-Part Neuron*, and a real-valued neuron corresponding to the unipotent part Y as an *Unipotent-Part Neuron*. It should be noted that there are the following restrictions on a set of weight parameters of the hyperbolic neuron.

(Weight for the real part x_k of an input signal to *Real-Part Neuron*)
$=$ (Weight for the unipotent part y_k of an input signal to *Unipotent-Part Neuron*), (13)
(Weight for the unipotent part y_k of an input signal to *Real-Part Neuron*)
$=$ (Weight for the real part x_k of an input signal to *Unipotent-Part Neuron*). (14)

Learning is carried out under these restrictions.

Moreover, note here that

$$\begin{aligned} \begin{bmatrix} S \\ T \end{bmatrix} &= \left[\begin{array}{cc|c|cc} v_1 & w_1 & \cdots & v_N & w_N \\ w_1 & v_1 & \cdots & w_N & v_N \end{array}\right] \left[\begin{array}{c} x_1 \\ y_1 \\ \hline \vdots \\ \hline x_N \\ y_N \end{array}\right] \\ &\qquad + \begin{bmatrix} c \\ d \end{bmatrix} \\ &= |h_1|_h \begin{bmatrix} \cosh\phi_1 & \sinh\phi_1 \\ \sinh\phi_1 & \cosh\phi_1 \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \cdots \\ &\qquad + |h_N|_h \begin{bmatrix} \cosh\phi_N & \sinh\phi_N \\ \sinh\phi_N & \cosh\phi_N \end{bmatrix}\begin{bmatrix} x_N \\ y_N \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}, \end{aligned}$$

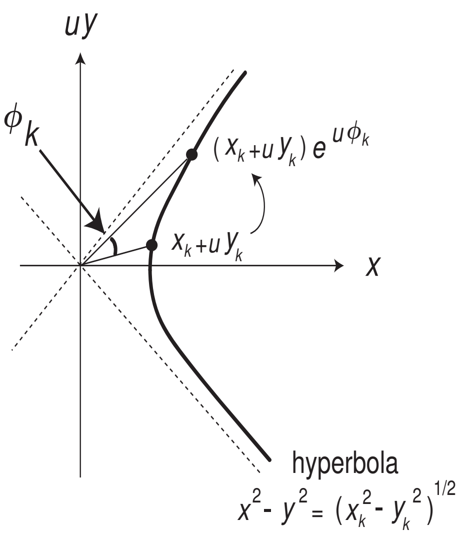

Fig. 2. Counterclockwise hyperbolic rotation by the angle ϕ_k about the origin in the hyperbolic number plane.

(15)

where $\phi_k = \tanh^{-1}(w_k/v_k) \quad (1 \le k \le N)$. In Eq. (15), $|h_k|_h$ means reduction or magnification of the distance between a point (x_k, y_k) and the origin in the hyperbolic number plane, $\begin{bmatrix} \cosh\phi_k & \sinh\phi_k \\ \sinh\phi_k & \cosh\phi_k \end{bmatrix}$ (2-dimensional orthogonal matrix) the counterclockwise *hyperbolic rotation* by the angle ϕ_k about the origin, and $[c \quad d]^t$ translation. Here, the point (x_k, y_k) in the hyperbolic number plane moves on the hyperbola $x^2 - y^2 = \sqrt{x_k^2 - y_k^2}$ by the counterclockwise hyperbolic rotation by the angle ϕ_k about the origin (Fig. 2).

IV. Decision Boundaries in Hyperbolic Neurons

The decision boundary is a border by which pattern classifiers classify input patterns, and generally consists of hypersurfaces. Decision boundaries of neural networks of real-valued neurons and complex-valued neurons were examined [6], [8]. This section mathematically analyzes decision boundaries of hyperbolic neurons.

A. The Case of the Split Hyperbolic Step Activation Function

First, we investigate the decision boundary of the hyperbolic neuron with the following activation function:

$$f_H^{(3)}(z) = 1_R(x) + u1_R(y), \quad z = x + uy \tag{16}$$

where 1_R is a real-valued step function defined on $\boldsymbol{R}$, that is, $1_R(s) = 1$ (if $s \ge 0$), $1_R(s) = 0$ (otherwise) for any $s \in \boldsymbol{R}$. Consider the hyperbolic neuron with N-inputs described in Section III-C whose activation function is Eq. (16) (i.e., $f_H = f_H^{(3)}$). For N input signals (hyperbolic numbers) $\boldsymbol{z} = [z_1 \cdots z_N]^t = \boldsymbol{x} + u\boldsymbol{y}$ where $\boldsymbol{x} = [x_1 \cdots x_N]^t$ and $\boldsymbol{y} = [y_1 \cdots y_N]^t$, the hyperbolic neuron generates

$$\begin{aligned} X + uY \;=\; & 1_R\Big([\boldsymbol{v}^t \;\; \boldsymbol{w}^t]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} + c\Big) \\ & + u \cdot 1_R\Big([\boldsymbol{w}^t \;\; \boldsymbol{v}^t]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} + d\Big) \end{aligned} \tag{17}$$

as an output. Thus, the decision boundary for the real part of an output of the hyperbolic neuron with N-inputs is

$$S(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} + c = 0. \tag{18}$$

That is, input signals $(\boldsymbol{x}^t, \boldsymbol{y}^t)^t \in \boldsymbol{R}^{2n}$ are classified into two decision regions $\{(\boldsymbol{x}^t, \boldsymbol{y}^t)^t \in \boldsymbol{R}^{2n} | S(\boldsymbol{x}, \boldsymbol{y}) \ge 0\}$ and $\{(\boldsymbol{x}^t, \boldsymbol{y}^t)^t \in \boldsymbol{R}^{2n} | S(\boldsymbol{x}, \boldsymbol{y}) < 0\}$ by the hypersurface given by Eq. (18). Similarly, Eq. (19) is the decision boundary for the unipotent part.

$$T(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} + d = 0. \tag{19}$$

The normal vectors $H^R(\boldsymbol{x}, \boldsymbol{y})$ and $H^U(\boldsymbol{x}, \boldsymbol{y})$ of the decision boundaries (Eqs. (18), (19)) are given by

$$H^R(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{v}^t \quad \boldsymbol{w}^t]^t, \tag{20}$$

$$H^U(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{w}^t \quad \boldsymbol{v}^t]^t, \tag{21}$$

respectively. Here, since

$$\frac{H^R(\boldsymbol{x}, \boldsymbol{y})^t \cdot H^U(\boldsymbol{x}, \boldsymbol{y})}{|H^R(\boldsymbol{x}, \boldsymbol{y})| \cdot |H^U(\boldsymbol{x}, \boldsymbol{y})|} = \frac{2\boldsymbol{v}^t \cdot \boldsymbol{w}}{|\boldsymbol{v}|^2 + |\boldsymbol{w}|^2}, \tag{22}$$

we can find that the angle between the decision boundary for the real part of the hyperbolic neuron and that for the unipotent depends on the weight values, whereas the decision boundary for the real part of a complex-valued neuron and that for the imaginary part always inersects orthogonally [8]. It can be easily shown from Eq. (22) that the following proposition holds true.

Proposition 1: *(i) The decision boundary for the real part and that for the unipotent part of a hyperbolic neuron with Eq. (16) as an activation function inersects orthogonally if and only if*

$$\boldsymbol{v}^t \cdot \boldsymbol{w} = 0. \tag{23}$$

(ii) The decision boundary for the real part and that for the unipotent part of a hyperbolic neuron with Eq. (16) as an activation function are parallel if and only if

$$\boldsymbol{v}^t \cdot \boldsymbol{w} = \pm 1. \tag{24}$$

□

B. The Case of the Split Hyperbolic Sigmoid Activation Function

Next, we investigate the decision boundary of the hyperbolic neuron with the split hyperbolic sigmoid activation function (Eq. (6)). Consider the hyperbolic neuron with N-inputs described in Section III-C whose activation function is Eq. (6) (i.e., $f_H = f_H^{(1)}$). For N input signals (hyperbolic numbers) $\boldsymbol{z} = [z_1 \cdots z_N]^t = \boldsymbol{x} + u\boldsymbol{y}$ where $\boldsymbol{x} = [x_1 \cdots x_N]^t$ and $\boldsymbol{y} = [y_1 \cdots y_N]^t$, the hyperbolic neuron generates

$$\begin{aligned} X + uY &= f_R\Big([\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + c\Big) \\ &\quad + u f_R\Big([\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + d\Big) \end{aligned} \tag{25}$$

as an output. Here, for any two constants $C^R, C^U \in (0,1)$, let

$$X(\boldsymbol{x},\boldsymbol{y}) = f_R\left([\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + c\right) = C^R, \tag{26}$$

$$Y(\boldsymbol{x},\boldsymbol{y}) = f_R\left([\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + d\right) = C^U. \tag{27}$$

Eq. (26) is the decision boundary for the real part of an output of the hyperbolic neuron with N-inputs, and Eq. (27) the decision boundary for the unipotent part. The normal vectors $H^R(\boldsymbol{x},\boldsymbol{y})$ and $H^U(\boldsymbol{x},\boldsymbol{y})$ of the decision boundaries (Eqs. (26), (27)) are given by

$$\begin{aligned} H^R(\boldsymbol{x},\boldsymbol{y}) &= \left(\frac{\partial X}{\partial x_1} \cdots \frac{\partial X}{\partial x_N}\ \frac{\partial X}{\partial y_1} \cdots \frac{\partial X}{\partial y_N}\right) \\ &= f_R'\Big([\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + c\Big) \cdot [\boldsymbol{v}^t \quad \boldsymbol{w}^t]^t, \end{aligned} \tag{28}$$

$$\begin{aligned} H^U(\boldsymbol{x},\boldsymbol{y}) &= \left(\frac{\partial Y}{\partial x_1} \cdots \frac{\partial Y}{\partial x_N}\ \frac{\partial Y}{\partial y_1} \cdots \frac{\partial Y}{\partial y_N}\right) \\ &= f_R'\Big([\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + d\Big) \cdot [\boldsymbol{w}^t \quad \boldsymbol{v}^t]^t, \end{aligned} \tag{29}$$

respectively. Here, since

$$\frac{H^R(\boldsymbol{x},\boldsymbol{y})^t \cdot H^U(\boldsymbol{x},\boldsymbol{y})}{|H^R(\boldsymbol{x},\boldsymbol{y})| \cdot |H^U(\boldsymbol{x},\boldsymbol{y})|} = \frac{2\boldsymbol{v}^t \cdot \boldsymbol{w}}{|\boldsymbol{v}|^2 + |\boldsymbol{w}|^2}, \tag{30}$$

we can find that the angle between the decision boundary for the real part and that for the unipotent part of the hyperbolic neuron whose activation function is Eq. (6), is the same as that of the hyperbolic neuron with Eq. (16). Thus, the same proposition as Proposition 1 holds true for this case.

C. The Case of the Non-Split Hyperbolic Sigmoidal Activation Function

Finally, we investigate the decision boundary of the hyperbolic neuron with the non-split hyperbolic sigmoid activation function (Eq. (7)). Consider the hyperbolic neuron with N-inputs described in Section III-C whose activation function is Eq. (7) (i.e., $f_H = f_H^{(2)}$). For N input signals (hyperbolic numbers) $\boldsymbol{z} = [z_1 \cdots z_N]^t = \boldsymbol{x} + u\boldsymbol{y}$ where $\boldsymbol{x} = [x_1 \cdots x_N]^t$ and $\boldsymbol{y} = [y_1 \cdots y_N]^t$, the hyperbolic neuron generates

$$\begin{aligned} X + uY &= f_H^{(2)}(S + uT) \\ &= \frac{2 + e^{T-S} + e^{-(T+S)}}{2(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})} \\ &\quad + u \cdot \frac{e^{T-S} - e^{-(T+S)}}{2(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})} \end{aligned} \tag{31}$$

as an output, where

$$S = [\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + c, \tag{32}$$

$$T = [\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} + d. \tag{33}$$

Here, for any two constants $C^R, C^U \in \boldsymbol{R}$, let

$$\begin{aligned} X(\boldsymbol{x},\boldsymbol{y}) &= \frac{2 + e^{T-S} + e^{-(T+S)}}{2(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})} \\ &= C^R, \end{aligned} \tag{34}$$

$$\begin{aligned} Y(\boldsymbol{x},\boldsymbol{y}) &= \frac{e^{T-S} - e^{-(T+S)}}{2(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})} \\ &= C^U. \end{aligned} \tag{35}$$

Eq. (34) is the decision boundary for the real part of an output of the hyperbolic neuron with the non-split activation function (Eq. (7)), and Eq. (35) the decision boundary for the unipotent part. The normal vectors $H^R(\boldsymbol{x},\boldsymbol{y})$ and $H^U(\boldsymbol{x},\boldsymbol{y})$ of the decision boundaries (Eqs. (34), (35)) are given by

$$H^R(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{\partial X}{\partial x_1} \cdots \frac{\partial X}{\partial x_N}\ \frac{\partial X}{\partial y_1} \cdots \frac{\partial X}{\partial y_N}\right), \tag{36}$$

where for any $1 \le k \le N$,

$$\begin{aligned} \frac{\partial X}{\partial x_k} &= \frac{(v_k - w_k) \cdot \cosh(T+S)}{e^{2S}(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})^2} \\ &\qquad \underline{+(v_k + w_k) \cdot \cosh(S-T) + 2v_k}, \end{aligned} \tag{37}$$

$$\begin{aligned} \frac{\partial X}{\partial y_k} &= \frac{-(v_k - w_k) \cdot \cosh(T+S)}{e^{2S}(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})^2} \\ &\qquad \underline{+(v_k + w_k) \cdot \cosh(S-T) + 2w_k}, \end{aligned} \tag{38}$$

and

$$H^U(\boldsymbol{x},\boldsymbol{y}) = \left(\frac{\partial Y}{\partial x_1} \cdots \frac{\partial Y}{\partial x_N}\ \frac{\partial Y}{\partial y_1} \cdots \frac{\partial Y}{\partial y_N}\right), \tag{39}$$

where for any $1 \le k \le N$,

$$\begin{aligned} \frac{\partial Y}{\partial x_k} &= \frac{-(v_k - w_k) \cdot \cosh(T+S)}{e^{2S}(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})^2} \\ &\qquad \underline{+(v_k + w_k) \cdot \cosh(S-T) + 2w_k}, \end{aligned} \tag{40}$$

$$\begin{aligned} \frac{\partial Y}{\partial y_k} &= \frac{(v_k - w_k) \cdot \cosh(T+S)}{e^{2S}(1 + e^{T-S} + e^{-(T+S)} + e^{-2S})^2} \\ &\qquad \underline{+(v_k + w_k) \cdot \cosh(S-T) + 2v_k}. \end{aligned} \tag{41}$$

So,

$$\begin{aligned} &\frac{H^R(\boldsymbol{x},\boldsymbol{y})^t \cdot H^U(\boldsymbol{x},\boldsymbol{y})}{|H^R(\boldsymbol{x},\boldsymbol{y})| \cdot |H^U(\boldsymbol{x},\boldsymbol{y})|} \\ &= \frac{(\cosh(T-S)+1)^2 \cdot \sum_{k=1}^N (v_k + w_k)^2}{(\cosh(T-S)+1)^2 \cdot \sum_{k=1}^N (v_k + w_k)^2} \\ &\qquad \frac{-(\cosh(T+S)+1)^2 \cdot \sum_{k=1}^N (v_k - w_k)^2}{+(\cosh(T+S)+1)^2 \cdot \sum_{k=1}^N (v_k - w_k)^2}. \end{aligned} \tag{42}$$

Thus, $H^R(\boldsymbol{x},\boldsymbol{y})$ and $H^U(\boldsymbol{x},\boldsymbol{y})$ intersect orthogonally if and only if

$$\begin{aligned}&(\cosh(T-S)+1)^2\cdot\sum_{k=1}^{N}(v_k+w_k)^2\\ =\ &(\cosh(T+S)+1)^2\cdot\sum_{k=1}^{N}(v_k-w_k)^2. \qquad (43)\end{aligned}$$

Therefore, the following proposition can be obtained.

Proposition 2: *Let V be the intersection of the decision boundary for the real part of the hyperbolic neuron with the non-split activation function $f_H^{(2)}$ (Eq. (7)) and that for the unipotent part. That is,*

$$V \stackrel{\text{def}}{=} \{(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{R}^{2N}|X(\boldsymbol{x},\boldsymbol{y})=C^R, Y(\boldsymbol{x},\boldsymbol{y})=C^U\}. \qquad (44)$$

Take any $(\boldsymbol{x}_0,\boldsymbol{y}_0)\in V$ and fix it. Then, the two decision boundaries $X(\boldsymbol{x},\boldsymbol{y})=C^R$ and $Y(\boldsymbol{x},\boldsymbol{y})=C^U$ intersect orthogonally at $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ if and only if

$$\begin{aligned}&(\cosh(T_0-S_0)+1)^2\cdot\sum_{k=1}^{N}(v_k+w_k)^2\\ =\ &(\cosh(T_0+S_0)+1)^2\cdot\sum_{k=1}^{N}(v_k-w_k)^2, \qquad (45)\end{aligned}$$

where

$$S_0 = [\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{y}_0\end{bmatrix}+c, \qquad (46)$$

$$T_0 = [\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{y}_0\end{bmatrix}+d. \qquad (47)$$

□

Example 1: *Consider the following case:*

$$X(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}, \qquad (48)$$

$$Y(\boldsymbol{x},\boldsymbol{y}) = 0. \qquad (49)$$

Then,

$$\begin{aligned}V &\stackrel{\text{def}}{=} \{(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{R}^{2N}\mid X(\boldsymbol{x},\boldsymbol{y})=\frac{1}{2},\ Y(\boldsymbol{x},\boldsymbol{y})=0\}\\ &= \{(\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{R}^{2N}\mid S=[\boldsymbol{v}^t \quad \boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}+c=0,\\ &\qquad T=[\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}+d=0\ \}, \qquad (50)\end{aligned}$$

and Eq. (45) implies

$$\boldsymbol{v}^t\cdot\boldsymbol{w}=0. \qquad (51)$$

Eq. (51) is the same as Eq. (23) which is a sufficient and necessary condition for the decision boundary for the real part and that for the unipotent part of the hyperbolic neuron with the split hyperbolic step activation function (Eq. (16)) to intersect orthogonally. □

D. The Case of the Non-Split Complex-Valued Sigmoid Activation Function

In this section, the decision boundary of the complex-valued neuron with the non-split (regular) complex-valued activation function (Eq. (10)) is investigated for comparison with that of the hyperbolic neuron with the non-split hyperbolic activation function (Eq. (7)).

Consider a complex-valued neuron with N-inputs, weights $v_k+iw_k\in\boldsymbol{C}$ $(1\le k\le N; i=\sqrt{-1})$, and a threshold value $\theta=c+id\in\boldsymbol{C}$. Then, for N input signals $x_k+iy_k\in\boldsymbol{C}$ $(1\le k\le N)$, the complex-valued neuron generates

$$\begin{aligned}X+iY &= f_C(U+iV)\\ &= \frac{e^U+\cos V}{e^U+e^{-U}+2\cos V}\\ &\qquad +i\left(\frac{\sin V}{e^U+e^{-U}+2\cos V}\right) \qquad (52)\end{aligned}$$

as an output where f_C is the non-split (regular) complex-valued sigmoid activation function (Eq. (10)), and

$$\begin{aligned}U+iV &= \left([\boldsymbol{v}^t \quad -\boldsymbol{w}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}+c\right)\\ &\qquad +i\left([\boldsymbol{w}^t \quad \boldsymbol{v}^t]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix}+d\right) \qquad (53)\end{aligned}$$

where $\boldsymbol{v}=(v_1\cdots v_N)^t$, $\boldsymbol{w}=(w_1\cdots w_N)^t$, $\boldsymbol{x}=(x_1\cdots x_N)^t$ and $\boldsymbol{y}=(y_1\cdots y_N)^t$.

Here, for any two constants $C^R, C^I\in\boldsymbol{R}$, let

$$X(\boldsymbol{x},\boldsymbol{y}) = \frac{e^U+\cos V}{e^U+e^{-U}+2\cos V}=C^R, \qquad (54)$$

$$Y(\boldsymbol{x},\boldsymbol{y}) = \frac{\sin V}{e^U+e^{-U}+2\cos V}=C^I. \qquad (55)$$

Eq. (54) is the decision boundary for the real part of an output of the complex-valued neuron with the non-split activation function, and Eq. (55) the decision boundary for the imaginary part. The normal vectors $H^R(\boldsymbol{x},\boldsymbol{y})$ and $H^I(\boldsymbol{x},\boldsymbol{y})$ of the decision boundaries (Eqs. (54), (55)) are given by

$$H^R(\boldsymbol{x},\boldsymbol{y})=\left(\frac{\partial X}{\partial x_1}\ \cdots\ \frac{\partial X}{\partial x_N}\ \frac{\partial X}{\partial y_1}\ \cdots\ \frac{\partial X}{\partial y_N}\right) \qquad (56)$$

where for any $1\le k\le N$,

$$\frac{\partial X}{\partial x_k} = \frac{2v_k+v_k(e^U+e^{-U})\cos V + w_k(e^U-e^{-U})\sin V}{(e^U+e^{-U}+2\cos V)^2}, \qquad (57)$$

$$\frac{\partial X}{\partial y_k} = \frac{-2w_k-w_k(e^U+e^{-U})\cos V + v_k(e^U-e^{-U})\sin V}{(e^U+e^{-U}+2\cos V)^2}, \qquad (58)$$

and

$$H^I(\boldsymbol{x},\boldsymbol{y})=\left(\frac{\partial Y}{\partial x_1}\ \cdots\ \frac{\partial Y}{\partial x_N}\ \frac{\partial Y}{\partial y_1}\ \cdots\ \frac{\partial Y}{\partial y_N}\right) \qquad (59)$$

TABLE I

THE XOR PROBLEM.

Input	Output
0 0	0
0 1	1
1 0	1
1 1	0

TABLE II

ENCODING OF THE XOR PROBLEM USED IN [8] FOR THE COMPLEX-VALUED NEURON.

Input	Output
$-1-i$	1
$-1+i$	0
$1-i$	$1+i$
$1+i$	i

where for any $1 \leq k \leq N$,

$$\frac{\partial Y}{\partial x_k} = \frac{2w_k + w_k(e^U + e^{-U})\cos V}{(e^U + e^{-U} + 2\cos V)^2} \qquad \frac{-v_k(e^U - e^{-U})\sin V}{}, \tag{60}$$

$$\frac{\partial Y}{\partial y_k} = \frac{2v_k + v_k(e^U + e^{-U})\cos V}{(e^U + e^{-U} + 2\cos V)^2} \qquad \frac{+w_k(e^U - e^{-U})\sin V}{}. \tag{61}$$

From Eqs. (57), (58), (60) and (61), for any $1 \leq k \leq N$,

$$\frac{\partial X}{\partial x_k} \cdot \frac{\partial Y}{\partial x_k} + \frac{\partial X}{\partial y_k} \cdot \frac{\partial Y}{\partial y_k} = 0. \tag{62}$$

Thus,

$$H^R(\boldsymbol{x}, \boldsymbol{y})^t \cdot H^I(\boldsymbol{x}, \boldsymbol{y}) = 0. \tag{63}$$

Therefore, the decision boundary for the real part of an output of the complex-valued neuron with the non-split activation function (Eq. (10)) and the decision boundary for the imaginary part always intersect orthogonally.

V. SIMULATIONS

The main focus of this paper lies in the theoretical study of decision boundaries of hyperbolic neurons. Therefore the following examples are mainly given for illustration purposes. In [8] it was shown how the XOR problem (Table I) can be solved by one single complex-valued neuron (using the encoding of Table II), which always has orthogonal decision boundaries. In Section III, we showed that the hyperbolic neuron can have orthogonal decision boundaries as well (Proposition 1: if the scalar product of the two parts of its weight vector equals zero). An encoding that allows to solve the XOR problem by a single hyperbolic neuron with one input and the activation function $f_H^{(3)}$ (Eq. (16)) is listed in Table III. Setting the weight $h_1 = 0 - u$ and the threshold $\theta = 0$ yields the decision boundary shown in Fig. 3, which solves the that way encoded XOR problem. Note that $v_1 w_1 = 0 \cdot -1 = 0$.

Fig. 4 shows another simulation result for a linearly non-separable problem in which the hyperbolic neuron with the

TABLE III

ENCODING OF THE XOR PROBLEM FOR THE HYPERBOLIC NEURON.

Input	Output
$-1-u$	$1+u$
$-1+u$	u
$1-u$	1
$1+u$	0

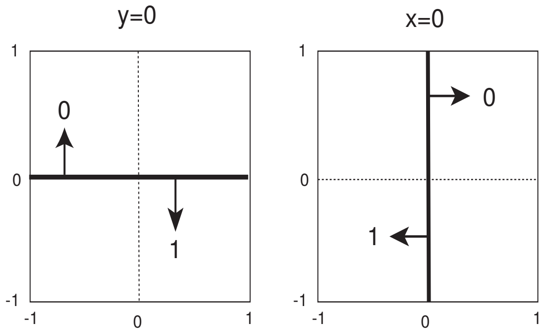

Fig. 3. Decision boundaries for the hyperbolic neuron solving the XOR problem. Decision boundary for the real-part (left) and that for the unipotent-part (right).

weight $h_1 = 1 + 2.4u$ and the threshold $\theta = 0$ yielded the decision boundary which successfully discriminated the four classes. Note that $2v_1 w_1/(v_1^2 + w_1^2) = 0.7071$, which means the angle between the two decision boundaries is $\pi/4$ radian. Thus, the angle between the decision boundaries can be controlled by changing the weight parameters.

VI. CONCLUSIONS

We investigated the basic properties (especially the decision boundary) of the hyperbolic neuron which is a hyperbolic counterpart of the complex-valued neuron, and proposed a non-split hyperbolic sigmoid activation function. The hyperbolic neural network with the non-split hyperbolic

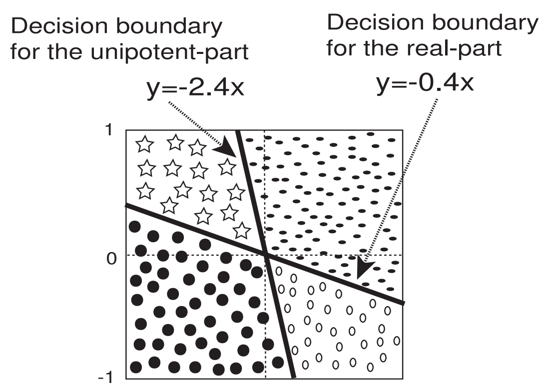

Fig. 4. Decision boundaries for the hyperbolic neuron solving a linearly non-separable problem.

sigmoid activation function seems to be promise because it is analytic and bounded in the hyperbolic plane. We expect that applications for the hyperbolic neural network will be found in fields dealing with non-Euclidean domains.

REFERENCES

[1] S. Buchholz and G. Sommer, “A hyperbolic multilayer perceptron,” *Proc. International Joint Conference on Neural Networks*, Como, Italy, 2000, vol. 2, pp. 129–133.

[2] G. M. Georgiou and C. Koutsougeras, “Complex domain backpropagation,” *IEEE Trans. Circuits and Systems–II: Analog and Digital Signal Processing*, vol. 39, no. 5, pp. 330–334, 1992.

[3] A. Hirose (ed.), *Complex-valued neural networks: theories and applications*, World Scientific Publishing, Singapore, 2003.

[4] M. S. Kim and C. C. Guest, “Modification of backpropagation networks for complex-valued signal processing in frequency domain,” *Proc. International Joint Conference on Neural Networks*, San Diego, June, 1990, vol. 3, pp. 27–31.

[5] T. Kim and T. Adali, “Approximation by fully complex multilayer perceptrons,” *Neural Computation*, vol. 15, no. 7, pp. 1641–1666, 2003.

[6] R. P. Lippmann, “An introduction to computing with neural nets,” *IEEE Acoustic, Speech and Signal Processing Magazine*, pp. 4–22, April 1987.

[7] T. Nitta, “An extension of the back-propagation algorithm to complex numbers,” *Neural Networks*, vol. 10, no. 8, pp. 1392–1415, 1997.

[8] T. Nitta, “Orthogonality of decision boundaries in complex-valued neural networks,” *Neural Computation*, vol. 16, no. 1, pp. 73–97, 2004.

[9] G. Sobczyk, “The hyperbolic number plane,” *The College Mathematics Journal*, vol. 26, no. 4, pp. 268–280, 1995.