Non-constant bounded holomorphic functions of hyperbolic numbers – Candidates for hyperbolic activation functions

* Eckhard Hitzer (University of Fukui)

Abstract— The Liouville theorem states that bounded holomorphic complex functions are necessarily constant. Holomorphic functions fulfill the so-called Cauchy-Riemann (CR) conditions. The CR conditions mean that a complex \( z \)-derivative is independent of the direction. Holomorphic functions are ideal for activation functions of complex neural networks, but the Liouville theorem makes them useless. Yet recently the use of hyperbolic numbers, lead to the construction of hyperbolic number neural networks.

We will describe the Cauchy-Riemann conditions for hyperbolic numbers and show that there exists a new interesting type of bounded holomorphic functions of hyperbolic numbers, which are not constant. We give examples of such functions. They therefore substantially expand the available candidates for holomorphic activation functions for hyperbolic number neural networks.

Keywords: Hyperbolic numbers, Liouville theorem, Cauchy-Riemann conditions, bounded holomorphic functions

1 Introduction

For the sake of mathematical clarity, we first carefully review the notion of holomorphic functions in the two number systems of complex and hyperbolic numbers.

The Liouville theorem states that bounded holomorphic complex functions \( f : \mathbb{C} \to \mathbb{C} \) are necessarily constant \[1\]. Holomorphic functions are functions that fulfill the so-called Cauchy-Riemann (CR) conditions. The CR conditions mean that a complex \( z \)-derivative

\[
\frac{df(z)}{dz}, \quad z = x + iy \in \mathbb{C}, \quad x, y \in \mathbb{R}, \quad ii = -1, \quad (1)
\]

is independent of the direction with respect to which the incremental ratio, that defines the derivative, is taken \[5\]. Holomorphic functions would be ideal for activation functions of complex neural networks, but the Liouville theorem means that careful measures need to be taken in order to avoid poles (where the function becomes infinite).

Yet recently the use of hyperbolic numbers

\[
z = x + hy, \quad h^2 = 1, \quad x, y \in \mathbb{R}, \quad h \notin \mathbb{R}. \quad (2)
\]

lead to the construction of hyperbolic number neural networks. We will describe the generalized Cauchy-Riemann conditions for hyperbolic numbers and show that there exist bounded holomorphic functions of hyperbolic numbers, which are not constant. We give a new example of such a function. They are therefore excellent candidates for holomorphic activation functions for hyperbolic number neural networks \[2, 3\]. In \[3\] it was shown, that hyperbolic number neural networks allow to control the angle of the decision boundaries (hyperplanes) of the real and the unipotent \( h \)-part of the output. But Buchholz argued in \[4\], p. 114, that

Contrary to the complex case, the hyperbolic logistic function is bounded. This is due to the absence of singularities. Thus, in general terms, this seems to be a suitable activation function. Concretely, the following facts, however, might be of disadvantage. The real and imaginary part have different squashing values. Both component functions do only significantly differ from zero around the line \( x = y (x > 0) \) and \(-x = y (x < 0)\).

Complex numbers are isomorphic to the Clifford geometric algebra \( Cl_{0,1} \) which is generated by a single vector \( e_1 \) of negative square \( e_1 = -1 \), with algebraic basis \( \{1, e_1\} \). The isomorphism \( \mathbb{C} \cong Cl_{0,1} \) is realized by mapping \( i \mapsto e_1 \).

Hyperbolic numbers are isomorphic to the Clifford geometric algebra \( Cl_{1,0} \) which is generated by a single vector \( e_1 \) of positive square \( e_1 = +1 \), with algebraic basis \( \{1, e_1\} \). The isomorphism between hyperbolic numbers and \( Cl_{1,0} \) is realized by mapping \( h \mapsto e_1 \).

2 Complex variable functions

We follow the treatment given in \[5\]. We assume a complex function given by an absolute convergent

Note that we slightly correct the two formulas of Buchholz, because we think it necessary to delete \( e_1 \) in Buchholz' original \( x = ye_1 (x > 0) \), etc.
power series.

\[ w = f(z) = f(x + iy) = u(x, y) + iv(x, y), \quad (3) \]

where \( u, v : \mathbb{R}^2 \to \mathbb{R} \) are real functions of the real variables \( x, y \). Since \( u, v \) are obtained in an algebraic way from the complex number \( z = x + iy \), they cannot be arbitrary functions but must satisfy certain conditions. There are several equivalent ways to obtain these conditions. Following Riemann, we state that a function \( w = f(z) = u(x, y) + iv(x, y) \) is a function of the complex variable \( z \) if its derivative is independent of the direction (in the complex plane) with respect to which the incremental ratio is taken. This requirement leads to two partial differential equations, named after Cauchy and Riemann (CR), which relate \( u \) and \( v \).

One method for obtaining these equations is the following. We consider the expression \( w = u(x, y) + iv(x, y) \) only as a function of \( z \), but not of \( \bar{z} \), i.e. the derivative with respect to \( \bar{z} \) shall be zero. First we perform the bijective substitution

\[ x = \frac{1}{2}(z + \bar{z}), \quad y = -i\frac{1}{2}(z - \bar{z}), \quad (4) \]

based on \( z = x + iy, \bar{z} = x - iy \). For computing the derivative \( w_z = \frac{dw}{dz} \), with the help of the chain rule we need the derivatives of \( x \) and \( y \) of (4).

\[ x_z = \frac{1}{2}, \quad y_z = \frac{1}{2}i. \quad (5) \]

Using the chain rule we obtain

\[ w_z = u_xx z_x + u_y y_z + i(v_xx z_x + v_y y_z) \]

\[ = \frac{1}{2}u_z + \frac{1}{2}i u_y + i(\frac{1}{2}v_z + \frac{1}{2}iv_y) \]

\[ = \frac{1}{2}[u_x - v_y + i(v_x + u_y)] = 0. \quad (6) \]

Requiring that both the real and the imaginary part of (6) vanish we obtain the Cauchy-Riemann conditions

\[ u_x = v_y, \quad u_y = -v_x. \quad (7) \]

Functions of a complex variable that fulfill the CR conditions are functions of \( x \) and \( y \), but they are only functions of \( z \), not of \( \bar{z} \).

It follows from (7), that both \( u \) and \( v \) fulfill the Laplace equation

\[ u_{xx} = v_{yy} = v_{xy} = -u_{yy} \Leftrightarrow u_{xx} + u_{yy} = 0, \quad (8) \]

and similarly

\[ v_{xx} + v_{yy} = 0. \quad (9) \]

The Laplace equation is a simple example of an elliptic partial differential equation. The general theory of solutions to the Laplace equation is known as potential theory. The solutions of the Laplace equation are called harmonic functions and are important in many fields of science, notably the fields of electromagnetism, astronomy, and fluid dynamics, because they can be used to accurately describe the behavior of electric, gravitational, and fluid potentials. In the study of heat conduction, the Laplace equation is the steady-state heat equation [3].

Liouville’s theorem [1] states, that any bounded holomorphic function \( f : \mathbb{C} \to \mathbb{C} \), which fulfills the CR conditions is constant. Therefore for complex neural networks it is not very meaningful to use holomorphic functions as activation functions. If they are used, special measures need to be taken to avoid poles in the complex plane. Instead separate componentwise (split) real scalar functions for the real part \( g_r : \mathbb{R} \to \mathbb{R}, u(x, y) \mapsto g_r(u(x, y)) \), and for the imaginary part \( g_i : \mathbb{R} \to \mathbb{R}, v(x, y) \mapsto g_i(v(x, y)) \), are usually adopted. Therefore a standard split activation function in the complex domain is given by

\[ g(u(x, y) + iv(x, y)) = g_r(u(x, y)) + ig_i(v(x, y)). \quad (10) \]

3 Hyperbolic numbers

Hyperbolic numbers are also known as split-complex numbers. They form a two-dimensional commutative algebra. The canonical hyperbolic system of numbers is defined [5] by

\[ z = x + hy, \quad h^2 = 1, \quad x, y \in \mathbb{R}, \quad h \notin \mathbb{R}. \quad (11) \]

The hyperbolic conjugate is defined as

\[ \bar{z} = x - hy. \quad (12) \]

Taking the hyperbolic conjugate corresponds in the isomorphic algebra \( Cl_{1,0} \) to taking the main involution (grade involution), which maps \( 1 \mapsto 1, e_1 \mapsto -e_1 \).

The hyperbolic invariant (corresponding to the Lorentz invariant in physics for \( y = ct \)), or modulus, is defined as

\[ z\bar{z} = (x + hy)(x - hy) = x^2 - y^2, \quad (13) \]

which is not positive definite.

Hyperbolic numbers are fundamentally different from complex numbers. Complex numbers and quaternions are division algebras, every non-zero element has a unique inverse. Hyperbolic numbers do not always have an inverse, but instead there are idempotents and divisors of zero.

We can define the following idempotent basis

\[ n_1 = \frac{1}{2}(1 + h), \quad n_2 = \frac{1}{2}(1 - h). \quad (14) \]
which fulfills
\[
\begin{align*}
n_1^2 &= \frac{1}{4}(1 + h)(1 + h) = \frac{1}{4}(2 + 2h) = n_1, \\
n_2^2 &= n_2, \quad n_1 + n_2 = 1, \\
n_1n_2 &= \frac{1}{4}(1 + h)(1 - h) = \frac{1}{4}(1 - 1) = 0, \\
\bar{n}_1 &= n_2, \quad \bar{n}_2 = n_1. 
\end{align*}
\]

(15)

The inverse basis transformation is simply
\[
1 = n_1 + n_2, \quad h = n_1 - n_2.
\]

(16)

Setting
\[
z = x + hy = \xi n_1 + \eta n_2, 
\]

(17)

we get the corresponding coordinate transformation
\[
x = \frac{1}{2}(\xi + \eta), \quad y = \frac{1}{2}(\xi - \eta), 
\]

(18)
as well as the inverse coordinate transformation
\[
\xi = x + y \in \mathbb{R}, \quad \eta = x - y \in \mathbb{R}. 
\]

(19)

The hyperbolic conjugate becomes, due to (15), in the idempotent basis
\[
\bar{z} = \xi n_1 + \eta n_2 = n_1 + \xi n_2. 
\]

(20)

In the idempotent basis, using (20) and (15), the hyperbolic invariant becomes multiplicative
\[
z\bar{z} = (\xi n_1 + \eta n_2)(\eta n_1 + \xi n_2) = \xi \eta (n_1 + n_2) = \xi \eta = x^2 - y^2. 
\]

(21)

In the following we consider the product and quotient of two hyperbolic numbers \(z, z^\prime \) both expressed in the idempotent basis \(\{n_1, n_2\}\)
\[
z z^\prime = (\xi n_1 + \eta n_2)(\xi^\prime n_1 + \eta^\prime n_2) = \xi \xi^\prime n_1 + \eta \eta^\prime n_2, 
\]

(22)

and
\[
\frac{z}{z^\prime} = \frac{\xi n_1 + \eta n_2}{\xi^\prime n_1 + \eta^\prime n_2} = \frac{z\bar{z}^\prime}{z^\prime \bar{z}} = \frac{(\xi n_1 + \eta n_2)(\eta^\prime n_1 + \xi^\prime n_2)}{(\xi^\prime n_1 + \eta^\prime n_2)(\eta n_1 + \xi n_2)} = \frac{(\xi \eta^\prime n_1 + \eta \xi^\prime n_2)(\eta^\prime n_1 + \xi^\prime n_2)}{\xi \eta} = \frac{\xi}{\xi^\prime n_1 + \eta^\prime n_2}. 
\]

(23)

Figure 1: The hyperbolic number plane \(\mathbb{H}\) with horizontal \(x\)-axis and vertical \(y\)-axis, showing: (a) Hyperbolas with modulus \(z\bar{z} = -1\) (green). (b) Straight lines with modulus \(z\bar{z} = 0\) \(\Leftrightarrow x^2 = y^2\) (red), i.e. divisors of zero. (c) Hyperbolas with modulus \(z\bar{z} = 1\) (blue).

4 Hyperbolic number functions

We assume a hyperbolic number function given by an absolute convergent power series
\[
w = f(z) = f(x + hy) = u(x, y) + hv(x, y), \\
h^2 = 1, \quad h \not\in \mathbb{R}. 
\]

(25)

where \(u, v : \mathbb{R}^2 \to \mathbb{R}\) are real functions of the real variables \(x, y\). An example of a hyperbolic number function is the exponential function
\[
e^z = e^{x + hy} = e^x e^{hy} = e^x (\cosh y + h \sinh y) = u(x, y) + hv(x, y), 
\]

(26)

with
\[
u(x, y) = e^x \cosh y, \quad v(x, y) = e^x \sinh y. 
\]

(27)

Since \(u, v\) are obtained in an algebraic way from the hyperbolic number \(z = x + hy\), they cannot be arbitrary functions but must satisfy certain conditions.
There are several equivalent ways to obtain these conditions. A function \( w = f(z) = u(x, y) + hv(x, y) \) is a function of the hyperbolic variable \( z \), if its derivative is independent of the direction (in the hyperbolic plane) with respect to which the incremental ratio is taken. This requirement leads to two partial differential equations, so called generalized Cauchy-Riemann (GCR) conditions, which relate \( u \) and \( v \).

To obtain the GCR conditions we consider the expression \( w = u(x, y) + hv(x, y) \) only as a function of \( z \), but not of \( \bar{z} = x - hy \), i.e. the derivative with respect to \( \bar{z} \) shall be zero. First we perform the bijective substitution

\[
x = \frac{1}{2}(z + \bar{z}), \quad y = h\frac{1}{2}(z - \bar{z}),
\]

based on \( z = x + hy, \bar{z} = x - hy \). For computing the derivative \( w_z = \frac{dw}{dz} \) with the help of the chain rule we need the derivatives of \( x \) and \( y \) of (28)

\[
x_z = \frac{1}{2}, \quad y_z = -\frac{1}{2} h.
\]

Using the chain rule we obtain

\[
w_z = u_xx x_z + u_y y_z + h(v_x x_z + v_y y_z) = \frac{1}{2} u_xx - \frac{1}{2} h u_y + h(\frac{1}{2} v_x - \frac{1}{2} h v_y) = \frac{1}{2}(u_x - v_y + h(u_x - u_y)) = 0.
\]

Requiring that both the real and the \( h \)-part of (30) vanish we obtain the GCR conditions

\[
u_x = v_y, \quad u_y = v_x.
\]

Functions of a hyperbolic variable that fulfill the GCR conditions are functions of \( x \) and \( y \), but they are only functions of \( z \), not of \( \bar{z} \). Such functions are called (hyperbolic) holomorphic functions.

It follows from (31), that \( u \) and \( v \) fulfill the wave equation

\[
u_{xx} = v_{yy} = u_{yy} \Rightarrow u_{xx} - u_{yy} = 0,
\]

and similarly

\[
v_{xx} - v_{yy} = 0.
\]

The wave equation is an important second-order linear partial differential equation for the description of waves – as they occur in physics – such as sound waves, light waves and water waves. It arises in fields like acoustics, electromagnetics, and fluid dynamics. The wave equation is the prototype of a hyperbolic partial differential equation [7].

Let us compute the partial derivatives \( u_x, u_y, v_x, v_y \) for the exponential function \( e^z \) of (26):

\[
u_x = e^z \cosh y, \quad u_y = e^z \sinh y, \quad v_x = e^z \sinh y = u_y, \quad v_y = e^z \cosh y = u_x.
\]

We clearly see that the partial derivatives (34) fulfill the GCR conditions (31) for the exponential function \( e^z \), as expected by its definition (26). The exponential function \( e^z \) is therefore a manifestly holomorphic hyperbolic function, but it is not bounded.

In the case of holomorphic hyperbolic functions the GCR conditions do not imply a Liouville type theorem like for holomorphic complex functions. This can most easily be demonstrated with a counter example

\[
f(z) = u(x, y) + hv(x, y), \quad u(x, y) = v(x, y) = \frac{1}{1 + e^{-x}e^{-y}}.
\]

The function \( u(x, y) \) is pictured in Fig. 2.

Let us verify that the function \( f \) of (35) fulfills the GCR conditions

\[
u_x = \frac{-1}{(1 + e^{-x}e^{-y})^2}(-e^{-x}e^{-y}) = \frac{e^{-x}e^{-y}}{(1 + e^{-x}e^{-y})^2},
\]

where we repeatedly applied the chain rule for differentiation. Similarly we obtain

\[
u_y = v_x = v_y = \frac{1 + e^{-x}e^{-y}}{(1 + e^{-x}e^{-y})}. \quad \text{(37)}
\]

The GCR conditions (31) are therefore clearly fulfilled, which means that the hyperbolic function \( f(z) \) of (35) is holomorphic. Since the exponential function \( e^{-z} \) has a range of \((0, \infty)\), the product \( e^{-x}e^{-y} \) also has values in the range of \((0, \infty)\). Therefore the function \( 1 + e^{-x}e^{-y} \) has values in \((1, \infty)\), and the components of the function \( f(z) \) of (35) have values

\[
0 < \frac{1}{1 + e^{-x}e^{-y}} < 1. \quad \text{(38)}
\]

We especially have

\[
\lim_{x,y \to -\infty} \frac{1}{1 + e^{-x}e^{-y}} = 0, \quad \text{(39)}
\]

and

\[
\lim_{x,y \to \infty} \frac{1}{1 + e^{-x}e^{-y}} = 1. \quad \text{(40)}
\]

The function (35) is representative for how to turn any real neural node activation function \( r(x) \) into holomorphic hyperbolic activation function via

\[
f(x) = r(x + y) (1 + h). \quad \text{(41)}
\]

We note that in \([3, 4]\) another holomorphic hyperbolic activation function was studied, namely

\[
f'(z) = \frac{1}{1 + e^{-z}}, \quad \text{(42)}
\]
The hyperbolic polar coordinate transformation \( [5] \) is an introduction. The split activation function used in but compare the quote from \([4]\), p. 114, given in the introduction. The split activation function used in \([8]\) can not be fulfilled.

\[ f''(x, y) = \frac{1}{1 + e^{-x} + h} \frac{1}{1 + e^{-y}} \] (43)
is clearly not holomorphic, because the real part \( u = 1/(1 + e^{-x}) \) depends only on \( x \) and not on \( y \), and the \( h \)-part \( v = 1/(1 + e^{-y}) \) depends only on \( y \) and not on \( x \), thus the GCR conditions \([31]\) can not be fulfilled.

5 Geometric interpretation of multiplication of hyperbolic numbers

In order to geometrically interpret the product of two complex numbers, it proves useful to introduce polar coordinates in the complex plane. Similarly, for the geometric interpretation of the product of two hyperbolic numbers, we first introduce hyperbolic polar coordinates for \( z = x + hy \) with radial coordinate

\[ \rho = \sqrt{|zz^*|} = \sqrt{|x^2 - y^2|}. \] (44)
The hyperbolic polar coordinate transformation \([5] \) is then given as

1. \( x^2 > y^2 \), \( x > 0 \):
\[ \theta = \text{artanh}(y/x), \quad z = \rho e^{h\theta}, \]
i.e. the quadrant in the hyperbolic plane of Fig. \([1]\) limited by the diagonal idempotent lines, and including the positive \( x \)-axis (to the right).

2. \( x^2 > y^2 \), \( x < 0 \):
\[ \theta = \text{artanh}(y/x), \quad z = \rho e^{h\theta}, \]
i.e. the quadrant in Fig. \([1]\) including the negative \( x \)-axis (to the left).

3. \( x^2 < y^2 \), \( y > 0 \):
\[ \theta = \text{artanh}(x/y), \quad z = h\rho e^{h\theta}, \]
i.e. the quadrant in Fig. \([1]\) including the positive \( y \)-axis (top).

4. \( x^2 < y^2 \), \( y < 0 \):
\[ \theta = \text{artanh}(x/y), \quad z = -h\rho e^{h\theta}, \]
i.e. the quadrant in Fig. \([1]\) including the negative \( y \)-axis (bottom).

The product of a constant hyperbolic number (assuming \( a_x^2 > a_y^2 \), \( a_x > 0 \))
\[ a = a_x + h a_y = \rho a e^{h\theta a}, \]
\[ \rho_a = \sqrt{a_x^2 - a_y^2}, \quad \theta_a = \text{artanh}(a_y/a_x), \] (45)
with a hyperbolic number \( z \) (assuming \( x^2 > y^2 \), \( x > 0 \)) in hyperbolic polar coordinates is
\[ az = \rho_a e^{h\theta a} \rho e^{h\theta} = \rho_a \rho e^{h(\theta + \theta_a)}. \] (46)
The geometric interpretation is a scaling of the modulus \( \rho \rightarrow \rho_a \rho \) and a hyperbolic rotation (movement along a hyperbola) \( \theta \rightarrow \theta + \theta_a \).

In the physics of Einstein’s special relativistic space-time \([11][12]\), the hyperbolic rotation \( \theta \rightarrow \theta + \theta_a \) corresponds to a Lorentz transformation from one inertial frame with constant velocity \( \tanh \theta \) to another inertial frame with constant velocity \( \tanh(\theta + \theta_a) \). Neural networks based on hyperbolic numbers (dimensionally extended to four-dimensional space-time) should therefore be ideal to compute with electromagnetic signals, including satellite transmission.

6 Conclusion

We have compared complex numbers and hyperbolic numbers, as well as complex functions and hyperbolic functions. We saw that according to Liouville’s theorem bounded complex holomorphic functions are necessarily constant, but non-constant bounded hyperbolic holomorphic functions exist. One such function has already being studied in \([3][4]\). We have studied a promising example of a hyperbolic holomorphic function
\[ f(z) = \frac{1 + h}{1 + e^{-x-y}}, \] (47)
in some detail. The distinct notions of idempotents and divisors of zero, special to hyperbolic numbers, were introduced. After further introducing hyperbolic polar coordinates, a geometric interpretation of the hyperbolic number multiplication was given.

Hyperbolic neural networks offer, compared to complex neural networks, therefore the advantage of suitable bounded non-constant hyperbolic holomorphic activation functions. It would certainly be of interest to study convergence, accuracy and decision boundaries of hyperbolic neural networks with the activation function (35), similar to [3, 4].

Acknowledgment

I want to acknowledge God [13]:

In the beginning was the Word and the Word was with God, and the Word was God. He was with God in the beginning. Through him all things were made; without him nothing was made that has been made. In him was life, and that life was the light of all mankind.

I want to thank my dear family, as well as T. Nitta and Y. Kuroe.

References

[1] K. Guerlebeck et al, Holomorphic Functions in the Plane and n-dimensional Space, Birkhauser, 2008, chp. 7.3.3.

[2] S. Buchholz, G. Sommer, A hyperbolic multilayer perceptron, Proceedings of the International Joint Conference on Neural Networks, Como, Italy, vol. 2, 129/133 (2000).

[3] T. Nitta, S. Buchholz, On the Decision Boundaries of Hyperbolic Neurons, Proceedings of the International Joint Conference on Neural Networks, IJCNN’08-HongKong, June 1-6, 2973/2979(2008).

[4] S. Buchholz, PhD Thesis, A Theory of Neural Computation with Clifford Algebras, University of Kiel, 2005.

[5] F. Catoni et al, The Mathematics of Minkowski Space-Time, Birkhauser, 2008.

[6] Laplace’s equation, Wikipedia, accessed 24 August 2011, http://en.wikipedia.org/wiki/Laplace’s_equation

[7] Wave equation, Wikipedia, accessed 24 August 2011, http://en.wikipedia.org/wiki/Wave_equation

[8] Online 3D function grapher, http://www.livephysics.com/ptools/online-3d-function-grapher.php?

[9] Split-complex number, Wikipedia, accessed 29 August 2011, http://en.wikipedia.org/wiki/Split-complex_number

[10] Notes of collaboration with H. Ishi, Feb. 2011, p. 15.

[11] C. Doran and A. Lasenby, Geometric Algebra for Physicists, Cambridge University Press, Cambridge (UK), 2003.

[12] E. Hitzer, Relativistic Physics as Application of Geometric Algebra, in K. Adhav (ed.), Proceedings of the International Conference on Relativity 2005 (ICR2005), University of Amravati, India, January 2005, 71/90(2005).

[13] The Bible, New International Version (NIV), The Gospel according to John, chapter 1, verses 1-4, http://www.biblegateway.com/

[14] Strong’s Bible lexicon entry G3056 for logos, available online at Blue Letter Bible. http://www.blueletterbible.org/lang/lexicon/lexicon.cfm?Strongs=G3056&t=KJV

2Greek term: logos. Note: A Greek philosopher named Heraclitus first used the term Logos around 600 B.C. to designate the divine reason or plan which coordinates a changing universe. [14]