The Hyperbolic Number Plane

Garret Sobczyk
Universidad de las Americas
email: sobczyk@udlapvms.pue.udlap.mx

INTRODUCTION. The complex numbers were grudgingly accepted by Renaissance mathematicians because of their utility in solving the cubic equation.[1] Whereas the complex numbers were discovered primarily for algebraic reasons, they take on geometric significance when they are used to name points in the plane. The complex number system is at the heart of *complex analysis* and has enjoyed more than 150 years of intensive development, finding applications in diverse areas of science and engineering.

At the beginning of the Twentieth Century, Albert Einstein developed his theory of special relativity, built upon Lorentzian geometry, yet at the end of the century almost all high school and undergraduate students are still taught only Euclidean geometry. At least part of the reason for this state of affairs has been the lack of a simple mathematical formalism in which the basic ideas can be expressed.

I argue that the hyperbolic numbers, blood relatives of the popular complex numbers, deserve to become a part of the undergraduate mathematics curriculum. They serve not only to put Lorentzian geometry on an equal mathematical footing with Euclidean geometry; their study also helps students develop algebraic skills and concepts necessary in higher mathematics. I have been teaching the hyperbolic number plane to my linear algebra and calculus students and have enjoyed an enthusiastic response.

THE HYPERBOLIC NUMBERS. The real number system can be extended in a new way. Whereas the algebraic equation $x^2 - 1 = 0$ has the real number solutions $x = \pm 1$, we assume the existence of a new number, the *unipotent* u, which has the algebraic property that $u \neq \pm 1$ but $u^2 = 1$. In terms of the *standard basis* $\{1, u\}$, any hyperbolic number $w \in \mathbb{H}$ can be written in the form $w = x + uy$ where x, y are real numbers. Thus the hyperbolic numbers $\mathbb{H} \equiv \mathbb{R}[u]$ are just the real numbers extended to include the unipotent u in the same way that the complex numbers $\mathbb{C} \equiv \mathbb{R}[i]$ are the real numbers extended to include the imaginary i.

[1] An historical account of this fascinating story is told by Tobias Dantzig in his book, *NUMBER: The Language of Science*, [**3**]. See also Struik's *A Concise History of Mathematics*, [**15**].

It follows that multiplication in $\mathbb{H}$ is defined by $(x+uy)(r+us) = (xr+ys)+u(xs+yr)$. Just as the complex numbers can be identified [**10**] with skew-symmetric 2×2 matrices with equal diagonal entries, $a+ib \leftrightarrow \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, the hyperbolic numbers correspond to the symmetric 2×2 matrices with equal diagonal entries: $x+uy \leftrightarrow \begin{pmatrix} x & y \\ y & x \end{pmatrix}$. This correspondence is an *isomorphism* because the operations of addition and multiplication in $\mathbb{H}$ correspond to the usual matrix operations. Further commutative and noncommutative extensions of the real number system and their relationship to matrix algebra are discussed in [**13**].

The complex numbers and the hyperbolic numbers are two-dimensional vector spaces over the real numbers, so each can identified with points in the plane $\mathbb{R}^2$. Using the standard basis $\{1, u\}$, we identify $w = x+uy$ with the point or vector (x, y), see *Figure 1*.

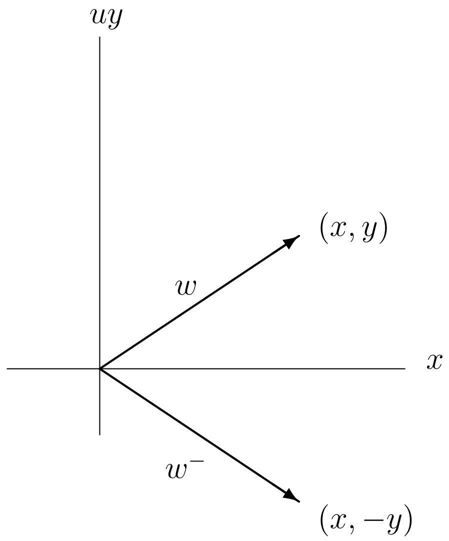

Figure 1. The hyperbolic number plane.

The hyperbolic number $w = x+uy$ names the corresponding point (x, y) in the coordinate plane. Also pictured is the conjugate $w^- = x - uy$.

The real numbers x and y are called the *real* and *unipotent* parts of the hyperbolic number w respectively. The *hyperbolic conjugate* w^- of w is defined by

$$w^- = x - uy.$$

The *hyperbolic modulus* of $w = x + uy$ is defined by

$$|w|_h \equiv \sqrt{|ww^-|} = \sqrt{|x^2 - y^2|} \tag{1}$$

and is considered the *hyperbolic distance* of the point w from the origin.

Note that the points $w \neq 0$ on the lines $y = \pm x$ are *isotropic* in the sense that they are nonzero vectors with $|w|_h = 0$. Thus the hyperbolic distance yields a geometry, *Lorentzian geometry*, on $\mathbb{R}^2$ quite unlike the usual Euclidean geometry of the complex plane where $|z| = 0$ only if $z = 0$. It is easy to verify that

$$w^{-1} = \frac{w^-}{ww^-} = \frac{x - uy}{x^2 - y^2}, \tag{2}$$

is the *multiplicative inverse* of w, whenever $|w|_h \neq 0$.

HYPERBOLIC POLAR FORM. Every nonzero complex number $z \in \mathbb{C}$ can be written in the *polar form*

$$z = r(\cos\theta + i\sin\theta) \equiv r\exp i\theta \tag{3}$$

for $0 \le \theta < 2\pi$, where $\theta = \tan^{-1}(y/x)$ is the angle that the vector z makes with the positive x-axis, and $r = |z| \equiv \sqrt{\bar{z}z}$ is the Euclidean distance of the point z to the origin. The set of all points in the complex number plane that satisfy the equation $|z| = r$ is a circle of radius $r \ge 0$ centered at the origin, see *Figure 2a.*

Figure 2. The r-circle and ρ-hyperbola.

Each time the parameter θ increases by 2π, the point $z = r\exp i\theta$ makes a complete counterclockwise revolution around the r-circle. In the case of the ρ-hyperbola, the points on the branches, given by $w = \pm\rho\exp u\phi$ in quadrants H-I and H-III, and $w = \pm\rho u\exp u\phi$ in the quadrants H-II and H-IV, respectively, are covered exactly once in the indicated directions as the parameter ϕ increases, $-\infty < \phi < \infty$.

Similarly the set of all points in the hyperbolic plane that satisfy the equation $|w|_h = \rho > 0$ is a *four branched* hyperbola of hyperbolic radius ρ. Such hyperbolic numbers $w = x + uy$ can be written

$$w = \pm\rho(\cosh\phi + u\sinh\phi) \equiv \pm\rho\exp u\phi \tag{4}$$

when w lies in the *hyperbolic quadrants H-I* or *H-III*, or

$$w = \pm\rho(\sinh\phi + u\cosh\phi) \equiv \pm\rho u\exp u\phi \tag{5}$$

when w lies in the hyperbolic quadrants *H-II* or *H-IV*, respectively. The *hyperbolic quadrants* of the hyperbolic plane are demarcated by the isotropic lines $|w|_h = 0$, which are the assymtotes of the ρ-hyperbolas $|w|_h = \rho > 0$. Each of the four hyperbolic branches is covered exactly once, in the indicated directions, as the parameter ϕ increases, $-\infty < \phi < \infty$. See *figure 2b*.

The *hyperbolic angle* ϕ is defined by $\phi \equiv \tanh^{-1}(y/x)$ in the quadrants *H-I* and *H-III*, or $\phi \equiv \tanh^{-1}(x/y)$ in *H-II* and *H-IV*, respectively. Just as the area of the sector of the unit circle with central angle θ is $\frac{1}{2}\theta$, the area of the unit hyperbolic sector determined by the ray from the origin to the point $\exp u\phi = \cosh\phi + u\sinh\phi$ (shaded in *Figure 2b*) is $\frac{1}{2}\phi$. The hyperbolic angle is discussed in [**16**], and more recently in [**14**].

The polar form of complex numbers provides the familiar geometric interpretation of complex number multiplication,

$$r_1\exp i\theta_1 \cdot r_2\exp i\theta_2 = r_1 r_2\exp[i(\theta_1 + \theta_2)]$$

Similarly the hyperbolic polar form gives a geometric interpretation of hyperbolic number multiplication, but because the hyperbolic plane is divided into four quadrants separated by the isotropic lines, we must keep track of the quadrants of the factors. For example if $w_1 = \rho_1\exp u\phi_1$ lies in quadrant *H-I* and $w_2 = \rho_2 u\exp u\phi_2$ lies in *H-II*, then

$$w_1 w_2 = \rho_1\rho_2 u\exp u\phi_1\exp u\phi_2 = \rho_1\rho_2 u\exp[u(\phi_1 + \phi_2)]$$

lies in quadrant *H-II* and is located as shown in *Figure 3*.

.

Figure 3

Shaded regions have equal area $\frac{1}{2}\phi_1$.

INNER AND OUTER PRODUCTS.

Let us now compare the multiplication of the hyperbolic numbers $w_1^- = x_1 - uy_1$ and $w_2 = x_2 + uy_2$ with the multiplication of the corresponding complex numbers $\overline{z_1} = x_1 - iy_1$ and $z_2 = x_2 + iy_2$. We get the *conjugate products*

$$\overline{z_1} z_2 = (x_1 - iy_1)(x_2 + iy_2) = (x_1x_2 + y_1y_2) + i(x_1y_2 - x_2y_1), \qquad (6)$$

and

$$w_1^- w_2 = (x_1 - uy_1)(x_2 + uy_2) = (x_1x_2 - y_1y_2) + u(x_1y_2 - x_2y_1). \qquad (7)$$

The *real* and *imaginary* parts of the conjugate product $\overline{z_1} z_2$ are called, respectively, the *inner* and *outer products* of the complex numbers z_1 and z_2. Likewise, the *real* and *unipotent* parts of the conjugate product $w_1^- w_2$ are called, respectively, the *hyperbolic inner* and *outer products* of the hyperbolic numbers w_1 and w_2. The vectors z_1 and z_2 in the complex number plane, and w_1 and w_2 in the hyperbolic number plane are said to be respectively *Euclidean orthogonal* or *hyperbolic orthogonal* if their respective inner products are zero.

From (6) and (7) it is seen that the components of the respective Euclidean and hyperbolic outer products are identical, and give the *directed area* of the parallelogram with w_1 and w_2 as adjacent edges. It follows that the concept of area is identical in Euclidean and Lorentzian geometry.

The conjugate products (6) and (7) are also nicely expressed in polar form. Letting $w_1 = \rho_1 \exp(u\phi_1) \in H\text{-}I$ and $w_2 = \rho_2 \exp(u\phi_2) \in H\text{-}I$, and $z_1 = r_1 \exp(i\theta_1)$ and $z_2 = r_2 \exp(i\theta_2)$, we find that

$$w_1^- w_2 = \rho_1\rho_2 \exp[u(\phi_2 - \phi_1)] = \rho_1\rho_2[\cosh(\phi_2 - \phi_1) + u\sinh(\phi_2 - \phi_1)]$$

where $\phi_2 - \phi_1$ is the hyperbolic angle between w_1 and w_2, and

$$\overline{z_1} z_2 = r_1 r_2 \exp[i(\theta_2 - \theta_1)] = r_1 r_2[\cos(\theta_2 - \theta_1) + i\sin(\theta_2 - \theta_1)]$$

where $\theta_2 - \theta_1$ is the Euclidean angle between z_1 and z_2. The special cases when $w_1 = su_\pm$ and/or $w_2 = tu_\pm$ for $s, t \in \mathbb{R}$ must be considered separately since no hyperbolic polar forms of these numbers exists. This happens when either or both of w_1 and w_2 lie on the isotropic lines $y = \pm x$.

Multiplication by $\exp(\phi u)$ is a linear transformation that sends the standard basis $\{1, u\}$ to $\{\exp(\phi u),\ u\exp(\phi u)\}$, and the product $\exp(\phi u)^- u \exp(\phi u) = u$ shows that the image basis is hyperbolic orthogonal like the standard basis. The mapping $w \to w\exp(\phi u)$ is naturally called the hyperbolic rotation through the hyperbolic angle ϕ. If $w = a + ub$ is any vector, the coordinates a', b' of w with respect to the rotated basis satisfy

$$w = a'\exp(\phi u) + b'u\exp(\phi u) = (a' + ub')\exp(\phi u),$$

so

$$a' + ub' = (a + ub)\exp(-\phi u). \tag{8}$$

It follows at once that $|a' + ub'|_h = |a + ub|_h$, that is, the hyperbolic distance of a point from the origin is independent of which hyperbolic orthogonal basis is used to coordinatize the plane. Figures 4 and 5 show the geometric pictures of multiplication by $\exp(i\theta)$ in the complex plane and multiplication by $\exp(\phi u)$ in the hyperbolic number plane, both as mappings sending each point to its image (alibi) and as a change of coordinates (alias). As a mapping the hyperbolic rotation $\exp(\phi u)$ moves each point $a+ub$ along the hyperbola $x^2 - y^2 = |a^2 - b^2|$, through the hyperbolic angle ϕ *(Figure 5a)*. Equivalently, as a change of coordinates, the new coordinates of any point are found by the usual parallelogram construction after rotation of the the axes of the original system through the hyperbolic angle ϕ, as in *Figure 5b*. (Note that the angles in Figure 5 are hyperbolic angles, not Euclidean angles!)

Figure 4. Euclidean rotation.
The shaded areas each equal $\frac{1}{2}\theta$.

Figure 5. Hyperbolic rotation.
The shaded areas each equal $\frac{1}{2}\phi$.

Notice that multiplying by i sends the basis $\{1, i\}$ into $\{i, -1\}$, a *counterclockwise* rotation about the origin taking the positive x-axis into the positive y-axis and the positive y-axis into the negative x-axis. Multiplying by u in the hyperbolic plane sends $\{1, u\}$ into $\{u, 1\}$, a *hyperbolic* rotation about the line $y = x$ that interchanges the positive x and y-axes.

THE IDEMPOTENT BASIS. Besides the standard basis $\{1, u\}$ in terms of which every hyperbolic number can be expressed as $w = x + uy$, the hyperbolic plane has another distinguished basis, associated with the two isotropic lines which separate the hyperbolic quadrants. The *idempotent basis* of the hyperbolic numbers is the pair $\{u_+, u_-\}$ where

$$u_+ = \frac{1}{2}(1 + u) \quad \text{and} \quad u_- = \frac{1}{2}(1 - u).$$

In terms of this basis the expression for $w = x + uy$ is $w = w_+u_+ + w_-u_-$, where

$$w_+ = x + y \quad \text{and} \quad w_- = x - y. \tag{9}$$

Conversely, given $w = w_+u_+ + w_-u_-$ we can recover the coordinates with respect to the standard basis by using the definitions of w_+ and w_-:

$$x = \frac{1}{2}(w_+ + w_-) \quad \text{and} \quad y = \frac{1}{2}(w_+ - w_-). \tag{10}$$

From the definitions of u_+ and u_- it is clear that $u_+ + u_- = 1$ and $u_+ - u_- = u$. We say that u_+ and u_- are *idempotents* because $u_+^2 = u_+$ and $u_-^2 = u_-$, and they are *mutually annihilating* because $u_+u_- = 0$. Therefore we have the *projective* properties

$$wu_+ = w_+u_+ \quad \text{and} \quad wu_- = w_-u_-. \tag{11}$$

(The decomposition $w = w_+u_+ + w_-u_-$ is the *spectral decomposition* of $w = x + uy$, that is, under the identification of hyperbolic numbers with symmetric 2×2 matrices, it corresponds to the spectral decomposition [**11**] of the symmetric matrix $\begin{pmatrix} x & y \\ y & x \end{pmatrix}$.)

Because of its special properties the idempotent basis is very nice for calculations. For example, from (9) we see that $w_+w_- = x^2 - y^2$. Then since for any hyperbolic numbers $v,\ w$ we have

$$vw = (v_+u_+ + v_-u_-)(w_+u_+ + w_-u_-) = (v_+w_+)u_+ + (v_-w_-)u_-,$$

it follows that

$$|vw|_h = \sqrt{|(v_+w_+)(v_-w_-)|} = \sqrt{|v_+v_-|}\sqrt{|w_+w_-|} = |v|_h|w|_h.$$

In particular, if one of the factors in a product is isotropic then the entire product is isotropic. Thus although strictly speaking the polar decomposition (4) is defined only for nonisotropic vectors, we can extend it to the isotropic lines by assigning the isotropic vectors $u_\pm$ to the exponentials of the hyperbolic arguments $\pm\infty$ respectively.

The binomial theorem takes a very simple form in the idempotent basis representation, so we can easily compute powers of hyperbolic numbers.

$$(w_+u_+ + w_-u_-)^k = (w_+)^k u_+^k + (w_-)^k u_-^k = (w_+)^k u_+ + (w_-)^k u_- \qquad (12)$$

This formula is valid for *all* real numbers $k \in \mathbb{R}$, and not just the positive integers. For example, for $k = -1$ we find that

$$1/w = w^{-1} = (1/w_+)u_+ + (1/w_-)u_-,$$

a valid formula for the inverse (2) of $w \in \mathbb{H}$, provided that $|w|_h \neq 0$.

Indeed, the validity of (12) allows us to extend the definitions of *all* of the elementary functions to the elementary functions in the hyperbolic number plane. If $f(x)$ is such a function, for $w = w_+u_+ + w_-u_-$ we define

$$f(w) \equiv f(w_+)u_+ + f(w_-)u_- \qquad (13)$$

provided that $f(w_+)$ and $f(w_-)$ are defined. It is natural to extend the hyperbolic numbers to the *complex hyperbolic numbers* or *unipodal numbers* by allowing $w_+, w_- \in \mathbb{C}$. The unipodal numbers have been studied in [**7**]. They will be used in the next section to find the solutions of the cubic equation.

Admitting the extension of the real number system to include the unipodal numbers raises questions as to the possibility and profitability of even further extensions? Shortly before his early death in 1879, William Kingdon Clifford *geometrically* extended the real number system to include the concept of direction, what is now called *Clifford algebra*. In [**8**], Clifford's *geometric algebra* is explored as a unified language for mathematics and physics.

THE CUBIC EQUATION. As mentioned in the introduction, the complex numbers were only grudgingly accepted because of their utility in solving the cubic equation. We demonstrate here the usefulness of complex hyperbolic numbers by finding a formula for the solutions of the venerated reduced cubic equation

$$x^3 + 3ax + b = 0. \qquad (14)$$

(The general cubic $Ay^3+3By^2+Cy+D=0,\ A\neq 0$, can be transformed to the reduced form (14) by dividing through by the leading coefficient and then making the linear substitution $y = x - B/A$.) The basic unipodal equation $w^n = r$ can easily be solved using the idempotent basis, with the help of equation (12). Writing $w = w_+u_+ + w_-u_-$, and $r = r_+u_+ + r_-u_-$ we get

$$w^n = w_+^n u_+ + w_-^n u_- = r_+u_+ + r_-u_-, \tag{15}$$

so $w_+^n = r_+$ and $w_-^n = r_-$. It follows that $w_+ = |r_+|^{\frac{1}{n}}\alpha^j$ and $w_- = |r_-|^{\frac{1}{n}}\alpha^k$ for some integers $0 \leq j,k \leq n-1$, where α is a primitive n^{th} root of unity. This proves the following theorem.

Theorem 1. For any positive interger n, the unipodal equation $w^n = r$ has n^2 solutions $w = \alpha^j r_+^{\frac{1}{n}}u_+ + \alpha^k r_-^{\frac{1}{n}}u_-$ for $j,k = 0,1,\ldots,n-1$, where $\alpha \equiv \exp(2\pi i/n)$.

The number of roots to the equation $w^n = r$ can be reduced by adding constraints. The following corollary follows immediately from the theorem, by noting that $w_+w_- = \rho \neq 0$ is equivalent to $w_- = \rho/w_+$.

Corollary. The unipodal equation $w^n = r$, subject to the constraint $w_+w_- = \rho$, for a nonzero complex number ρ, has the n solutions

$$w = \alpha^j r_+^{\frac{1}{n}}u_+ + \frac{\rho}{\alpha^j r_+^{\frac{1}{n}}}u_-,$$

for $j = 0,1,\ldots,n-1$, where $\alpha \equiv \exp(2\pi i/n)$, and $r_+^{\frac{1}{n}}$ denotes any n^{th} root of the complex number r_+.

We are now prepared to solve the reduced cubic equation (14).

Theorem 2. The reduced cubic equation $x^3 + 3ax + b = 0$, has the solutions, for $j = 0,1,2$,

$$x = \frac{1}{2}\left(\alpha^j\sqrt[3]{s+t} + \frac{\rho}{\alpha^j\sqrt[3]{s+t}}\right), \tag{16}$$

where $\alpha = \exp(2\pi i/3) = -\frac{1}{2} + \frac{1}{2}i\sqrt{3}$ is a primitive cube root of unity, and $\rho = -4a,\ s = -4b$, and $t = \sqrt{s^2-\rho^3} = 4\sqrt{b^2+4a^3}$.

Proof. The unipodal equation $w^3 = r$, where $r = s + ut$, is equivalent in the standard basis to $(x+yu)^3 = s+tu$, or $(x^3+3xy^2)+u(y^3+3x^2y) = s+ut$. Equating the complex scalar parts gives

$$x^3 + 3xy^2 - s = 0. \tag{17}$$

Making the additional constraint that $w_+w_- = x^2 - y^2 = \rho$, we can eliminate y^2 from (17), getting the equivalent equation

$$x^3 - \frac{3}{4}\rho x - \frac{1}{4}s = 0. \tag{18}$$

The constraint $w_+w_- = \rho$ further implies that

$$\rho^3 = (w_+w_-)^3 = w_+^3 w_-^3 = r_+r_- = s^2 - t^2,$$

which gives $t = \sqrt{s^2 - \rho^3}$. By letting $\rho = -4a$, and $s = -4b$, so $t = \sqrt{s^2 - \rho^3} = 4\sqrt{b^2 + 4a^3}$, the equation (18) becomes the reduced cubic equation $x^3 + 3ax + b = 0$. Since $r_+ = s + t$, the desired solution (16) is then obtained by taking the complex scalar part of the solution given in the corollary, using (10). Q.E.D.

Example. Find the solutions of the reduced cubic $x^3 - 6x + 4 = 0$.

Solution: Here $a = -2,\ b = 4$, so $\rho = 8,\ s = -16$, and we set $t = 16i$. Then $s + t = 16(-1 + i) = 2^{9/2}\exp(i3\pi/4)$, so we may take $\sqrt[3]{s + t} = 2^{3/2}\exp(i\pi/4)$. Thus

$$x = \frac{1}{2}[2^{3/2}\exp(i\pi/4)\alpha^j + \frac{8}{2^{3/2}\exp(i\pi/4)\alpha^j}]$$

$$= 2^{1/2}[\exp(i\pi/4)\alpha^j + \exp(-i\pi/4)\alpha^{-j}],$$

for $j = 0, 1, 2$. That is

$$i)\quad x = 2^{1/2}[\exp(i\pi/4) + \exp(-i\pi/4)] = 2^{1/2}[2\cos(\pi/4)] = 2,$$

$$ii)\quad x = 2^{1/2}[\exp(i\pi/4)\exp(2\pi i/3) + \exp(-i\pi/4)\exp(-2\pi i/3)]$$
$$= 2^{1/2}[\exp(11\pi i/12) + \exp(-11\pi i/12)]$$
$$= 2^{1/2}[2\cos(11\pi i/12)] = -(1 + \sqrt{3}) \simeq -2.7321,$$

or

$$iii)\quad x = 2^{1/2}[\exp(i\pi/4)\exp(-2\pi i/3) + \exp(-i\pi/4)\exp(2\pi i/3)]$$
$$= 2^{1/2}[\exp(-5\pi i/12) + \exp(5\pi i/12)]$$
$$= 2^{1/2}[2\cos(5\pi i/12)] = (-1 + \sqrt{3}) \simeq .7321.$$

A unipodal treatment of the quartic equation is found in [**7**].

SPECIAL RELATIVITY AND LORENTZIAN GEOMETRY. In 1905 Albert Einstein, at that time only 26 years of age, published his special theory of relativity, based on two postulates:

1. All coordinate systems (for measuring time and distance) moving with constant velocity relative to each other are equally legitimate for the formulation of the laws of physics.

2. Light propagates in every direction in a straight line and with the same speed c in every legitimate coordinate system.

The formulas for transforming space and time coordinates between systems in uniform relative motion had been found somewhat earlier by the Dutch physicist H. A. Lorentz, and were termed *Lorentz transformations* by H. Poincare. In 1907 Hermann Minkowski showed that Einstein's postulates imply a non-Euclidean geometry in four-dimensional space-time, called (flat) Minkowski spacetime. Whereas more elaborate mathematical formalisms have been developed [**6**], [**12**], and [**1**], most of the features of the geometry can be understood by studying a two-dimensional subspace [**12**] involving only one space dimension and time, which might naturally be called the Minkowski plane. But that term is already in use for another plane geometry that Minkowski developed for a quite different purpose, so this two dimensional model of spacetime is instead called the Lorentzian plane (even though Lorentz does not appear to have shared Minkowski's geometrical point of view). The hyperbolic numbers, which have also been called the "perplex numbers" [**4**], serve as coordinates in the Lorentzian plane in much the same way that the complex numbers serve as coordinates in the Euclidean plane.

An *event* X that occurs at time t and at the place x is specified by its coordinates t and x. If c is the speed of light, the product ct is a length, and the coordinates of the event X can be combined into the hyperbolic number $X = ct + ux$. By the *spacetime distance* between two events X_1 and X_2 we mean the hyperbolic modulus $|X_1 - X_2|_h$ which is the hyperbolic distance of the point $X_1 - X_2$ to the origin. If $X_1 = ct + x_1u$ and $X_2 = ct + x_2u$ are two events occuring at the same time t, then by (1),

$$|X_1 - X_2|_h = \sqrt{|0^2 - (x_1 - x_2)^2|} = |x_1 - x_2|, \tag{19}$$

which is exactly the Euclidean distance between the points x_1 and x_2 on the x-axis.

If the coordinates of an event in two-dimensional spacetime relative to one coordinate system are $X = ct + ux$, what are its coordinates $X' = ct' + ux'$ with respect to a second coordinate system that moves with uniform velocity $v < c$ with respect to the first system? We've already laid the groundwork for solving this problem.

Let $\phi = \tanh^{-1}(v/c)$, so $\sinh(\phi) = \frac{v}{c}\cosh(\phi)$ and the identity

$$1 = \cosh^2\phi - sin^2\phi = \cosh^2\phi(1 - v^2/c^2)$$

yields $\cosh\phi = \frac{1}{\sqrt{1-v^2/c^2}}$. Then since $X\exp(-\phi u) = X'$, by (8), we find

$$
\begin{aligned}
&(ct+ux)[\cosh(-\phi)+u\sinh(-\phi)] \\
&= (ct\cosh\phi - x\sinh\phi) + u(x\cosh\phi - ct\sinh\phi) \\
&= \frac{ct - xv/c}{\sqrt{1-v^2/c^2}} + u\frac{x - cvt/c}{\sqrt{1-v^2/c^2}} \\
&= ct' + ux'.
\end{aligned}
$$

This establishes the famous Lorentz equations

$$t' = \frac{t - xv/c^2}{\sqrt{1-v^2/c^2}}, \quad x' = \frac{x - vt}{\sqrt{1-v^2/c^2}}$$

relating the times and positions of an event as measured by two observers in relative motion [**2,p. 236**]. Diagrams such as Figure 6, showing two coordinate systems in spacetime in relative motion (using equal distance scales on the two axes), are called *Minkowski diagrams* [**5**]. In Figure 6 we have adopted the convention that the unipotent u lies along the horizontal x-axis, and 1 lies along the vertical time axis ct.

Figure 6: Minkowski diagram

The worldlines of two particles in relative uniform motion are given by $X(t) = ct$ and $X'(t') = ct'$. Each observer has a rest frame or hyperbolic orthogonal coordinate system in which he or she measures the relative time t and relative position x of an event $X = ct + xu$.

The *worldline* of a particle is just the graph of its location as a function of time, $X(t) = ct + ux(t)$. For example, in a coordinate system moving with the particle, so that the particle remains at $x(t) \equiv 0$ (the *rest frame* of the particle), the particle's worldline is just the straight line with equation $X = ct$, *Figure 6*.

The division of the Lorentz plane into quadrants has an important physical interpretation. The interval between two events, say the origin and X, is said to be *timelike* if there is a velocity $v < c$ such that in the coordinate system moving along the x-axis with velocity v the worldline of the origin passes through X. Thus to an observer in this coordinate system, the second event X occurs at the same place $x = 0$ but at a later time $t > 0$. Events with a timelike separation are *causally* related in that the earlier event can influence the later one, since a decision at the later time might be made on the basis of information left at the given location at the earlier time.

In the Minkowski diagram, Figure 6, each of the events $X_1 = ct_1$ and $X_2' = ct_2'$ have a timelike separation from the event $X_3' = ct_3'$. The events X_1 and X_2' have a *lightlike* separation because the line segment connecting them, the worldline of a *light signal*, is parallel to the isotropic line $x = ct$. (The isotropic lines $x = \pm ct$ are the worldlines of light signals moving at the velocity of light and passing through the origin $X = 0 = X'$.) The events X_1 and X_1' have a *spacelike* separation because in the coordinate system $X' = ct'$ moving along the x-axis with velocity $v < c$ the two appear as simultaneous events at different locations, since the line segment joining them is parallel to the x' axis.

Whereas the spacetime distance reduces to the usual Euclidean distance (19) when an observer makes measurements of the distances between objects in her rest frame along her x-axis, the *Fitzgerald-Lorentz contraction* comes in to play when two coordinate systems are moving with respect to each other. An object at rest in either system appears *shorter* when measured from the other system than when measured in its rest system.

Suppose system S' is moving with velocity $v < c$ with respect to system S. The worldlines of an interval I at rest in S with ends at the origin and point $X = x_0u$ are the lines $X = ct$ and $X = ct + ux_0$. Now recall that the points on the hyperbola $(ct)^2 - x^2 = \rho^2$, $\rho > 0$, all have hyperbolic distance ρ from the origin. Thus the point on the x'-axis at hyperbolic distance x_0 is the point where this axis intersects the hyperbola $(ct)^2 - x^2 = x_0^2$, and it is clear from the Minkowski diagram (Figure 7a) that this is beyond the intersection of the x'-axis with the worldline $X = ct + ux$. Thus the hyperbolic length of the interval I in the rest frame S' is less than x_0, its hyperbolic length in S.

The worldline of a particle moving with constant velocity $v < c$ along the x-axis is $X(t) = ct + utv$. Differentiating $X(t)$ with respect to t gives $dX/dt = c + uv$, the *world-velocity* of this particle; a vector parallel to the ct'-axis. Thus the worldlines of the endpoints of an interval I' of the x'-axis (i.e., an interval at rest in system S') will be parallel to the ct'-axis, as in Figure 7b, and again the Minkowski diagram shows that the length of this interval in the coordinate system S is less than its length in S'. Thus the invariance of the hyperbolic distance under change of orthogonal coordinates in the Lorentz plane is seen to be a geometric expression of the relativity of physical length measurements.

Figure **7.** Fitzgerald-Lorentz contraction.

The length of an object is dependent upon the rest frame in which it is measured.

If we define the *world-momentum* $P = cm_rV$, for the *relative mass* $m_r = \frac{m}{\sqrt{1-v^2/c^2}}$ of a particle with *rest mass* m moving with world-velocity V, we arrive at Einstein's most famous equation by the simple calculation of the *total energy* E,

$$E \equiv |P|_h = \sqrt{|PP^-|} = \sqrt{c^2 m_r^2 (c^2 - v^2)} = mc^2.$$

We see that the total energy $E = mc^2$ is independent of the rest frame in which it is calculated, and depends only upon the rest mass m and the square c^2 of the speed of light, [**2, p. 289**].

Many other fundamental results in special relativity theory can be derived using geometric properties of the Lorentzian plane. As we have seen, the hyperbolic numbers play the same role for computations in the Lorentzian plane that the complex numbers have in the Euclidean plane. Thus it is high time for the mathematics community to recognize their importance and introduce them to students early in the curriculum.

ACKNOWLEGEMENTS

I gratefully acknowledge the support of CONACYT grant 3803-E9401 and of INIP of the Universidad de las Americas. I would also like to thank the Editor and referees for their work in making improvements and additions, particularly in the last section on special relativity.

REFERENCES

[1.] W. E. Baylis, J. Huschilt, and Jiansu Wei, *Why i?*, Am J. Phys. **60** (9), 788, (1992).

[2.] M. Born, *Einstein's Theory of Relativity*, Revised Edition, Dover, 1962.

[3.] T. Dantzig, *NUMBER: The Language of Science*, Fourth Edition, Free Press, 1967.

[4.] P. Fjelstad, *Extending relativity via the perplex numbers*, Am. J. Phys. **54** (5), 416, (1986).

[5.] A. P. French, *Special Relativity*, Norton, New York, 1968.

[6.] D. Hestenes, *Space Time Algebra*, Gordon and Breach, 1966.

[7.] D. Hestenes, P. Reany, G. Sobczyk, Unipodal Algebra and Roots of Polynomials, *Advances in Applied Clifford Algebras*, **Vol. 1**, No. 1, (1991) 31-51.

[8.] D. Hestenes, G. Sobczyk, *CLIFFORD ALGEBRA TO GEOMETRIC CALCULUS: A Unified Language for Mathematics and Physics*, Kluwer, 1992.

[9.] C. Lanczos, *Space through the Ages*, Academic Press, New York, 1970.

[10.] J. E. Marsden, *Basic Complex Analysis,* Freeman, San Francisco, 1973, p. 10.

[11.] D. M. Schneider, *Linear Algebra,* 2nd ed., Macmillan, New York 1987, p. 383.

[12.] G. Sobczyk, *Spacetime vector analysis*, Physics Letters, **84A**, 45 (1981).

[13.] G. Sobczyk, Unipotents, Idempotents, and a Spinor Basis for Matrices, *Advances in Applied Clifford Algebras*, **Vol. 2**, No. 1, (1992) 51-64.

[14.] G. Sobczyk, Noncommutative Extensions of Number: An Introduction to Clifford's Geometric Algebra, *Aportaciones Matemáticas* Communicaciones **11**, 207-218, (1992).

[15.] D. J. Struik, *A Concise History of Mathematics*, Dover, 1967.

[16.] G. B. Thomas, Jr., *CALCULUS*, Addison-Wesley, 1953.