ArcTan Relations by Rich Schwartz

0. The purpose of this note is to give a hyperbolic geometry interpretation of the relations $R(a,b,c)$ of the form

$$\{a\} + \{b\} + \{c\} = 0; \qquad \{x\} = \text{atan}(1/x). \tag{1}$$

Here atan is the arc-tangent, or inverse tangent function. Since the arctangent is an odd function, we could also write $\{-a\} = \{b\} + \{c\}$. I learned about these kinds of relations by reading Ron Knott's website. I worked out the connection to hyperbolic geometry myself, but I am very sure that all of this is known to number theorists.

1. Let me first explain the importance of these relations, or at least one of them, namely $R(-1,2,3)$. I think that Euler discovered this relation. Leibniz's famous formula says that

$$\frac{\pi}{4} = \frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \cdots$$

In spite of its great beauty, this formula is not very good for actually computing π, because it converges very slowly. Summing the first 1000 terms gives 2 digits of accuracy. Observe that $\pi/4 = \text{atan}(1)$, and that there is a more general relation

$$\text{atan}(x) = \frac{x}{1} - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \cdots.$$

This equation converges for $x \in [0,1]$. The smaller the value of x, the faster the convergence. When $x < 1$ the convergence is exponentially fast. Since $\{1\} = \{2\} + \{3\}$, we have

$$\frac{\pi}{4} = \frac{(1/2) + (1/3)}{1} - \frac{(1/2)^3 + (1/3)^3}{3} + \frac{(1/2)^5 + (1/3)^5}{5} - \frac{(1/2)^7 + (1/3)^7}{7} \cdots.$$

Summing the first 1000 terms gives the first 605 digits of π. This is a much better method for computing π.

2. The familiar formula

$$\tan(x+y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x)\tan(y)} \tag{2}$$

applied to the relation $\{-a\} = \{b\} + \{c\}$ yields

$$\frac{-1}{a} = \frac{b+c}{bc-1}$$

Rearranging, we get

$$\frac{ab+bc+ca-1}{a(bc-1)} = 0.$$

Cancelling out the (positive) denominator, we get

$$ab + bc + ca = 1. \tag{3}$$

This is a familiar Diophantine equation, and its solutions are well known. Now I'm going to explain what it means in terms of hyperbolic geometry.

3. Introduce the function $L : \boldsymbol{R}^3 \times \boldsymbol{R}^3 \to \boldsymbol{R}$ defined by

$$L(X,Y) = \frac{1}{2}(x_1y_2 + x_2y_3 + x_3y_1 + y_1x_2 + y_2x_3 + y_3x_1). \tag{4}$$

Here $X = (x_1, x_2, x_3)$ and $Y = (y_1, y_2, y_3)$. The function L has the following general properties.

- $L(X,Y) = L(Y,X)$
- $L(X_1 + X_2, Y) = L(X_1, Y) + L(X_2, Y)$
- $L(aX, Y) = aL(X,Y)$.

In short, L is a *bilinear form*.

Note that

$$L(X,X) = x_1x_2 + x_2x_3 + x_3x_1.$$

Changing variables, we have

$$L(X,X) = ab + bc + ca; \qquad X = (a,b,c). \tag{5}$$

Therefore, the solutions to Equation 3 are precisely those integer vectors $X = (a,b,c)$. Note that (a,b,c) is a solution iff $(-a,-b,-c)$ is a solution. So, we may work with those integer vectors such that $a+b+c > 0$ and a,b,c are all nonzero. We call these the *good vectors*.

4. The set of vectors $X \in \boldsymbol{R}^3$ satisfying $L(X,X) = 1$ is precisely a hyperboloid of 2 sheets. One of the sheets contains vectors whose coordinate sum is positive. Let $\boldsymbol{H}^2$ denote the sheet containing the vectors with positive coordinate sum. The good vectors all live in $\boldsymbol{H}^2$. The sheet $\boldsymbol{H}^2$ is another incarnation of the Lorentz model for the hyperbolic plane.

Consider the standard basis vectors

$$E_1 = (1,0,0); \qquad E_2 = (0,1,0); \qquad E_3 = (0,0,1). \tag{6}$$

Note that $L(E_k,E_k) = 0$ for $k = 1,2,3$. Consider also the vectors

$$F_1 = (1,-1,-1); \qquad F_2 = (-1,1,-1); \qquad F_3(-1,-1,1). \tag{7}$$

These vectors enjoy the property that

$$L(E_i,F_j) = 0; \qquad i \neq j; \qquad L(E_i,F_i) = -1. \tag{8}$$

Next, define the maps $R_k : \boldsymbol{R}^3 \to \boldsymbol{R}^3$ by the formula

$$R_k(X) = X - 2\frac{L(X,F_k)}{L(F_k,F_k)}F_k \tag{9}$$

Just using the axioms that L satisfies as a bilinear form, we see that R_k is an order 2 L-preserving linear transformation which fixes E_{k-1} and E_{k+1}. In terms of matrices,

$$R_1 = \begin{bmatrix} -1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}; \qquad R_2 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{bmatrix}; \qquad R_3 = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix} \tag{10}$$

For example, let $s_0 = (1,1,0)$. Then up to permutation we generate the Fibonacci relations (on Ron Knotts' website) by iteratively applying the sequence $R_1, R_2, R_3, R_1, R_2, R_3, R_1...$ to s_0. That is,

- $s_1 = R_1(s_0) = (-1,3,2)$.
- $s_2 = R_2(s_1) = (5,-3,8)$.
- $s_3 = R_3(s_2) = (21,13,-8)$.
- $s_4 = R_1(s_3) = (-21,55,34)$ ...

5. Here is a way to visualize what is going on. Let Π denote the plane $x+y+z=1$. There is a nice map $\boldsymbol{H}^2 \to \Pi$ given by

$$(x_1,x_2,x_3) \to \frac{(x_1,x_2,x_3)}{x_1+x_2+x_3}. \tag{11}$$

One can then draw pictures of vectors by projecting to Π and identifying Π with the piece of paper on which you are drawing. The image of $\boldsymbol{H}^2$ under this projection is an open disk, which we think of as the open unit disk. A vector X is *null* if $L(X,X)=0$. The projection map makes sense on the null vectors even though these points do not lie in $\boldsymbol{H}^2$. The null vectors project to the unit circle, which is the boundary of the unit disk. In particular, the vectors E_1, E_2, E_3 project to the points of an equilateral triangle inscribed in the unit circle. Figure 1 shows these objects and also the 3 points of $\boldsymbol{H}^2$ that project to the midpoints of the edges of the said equilateral triangle.

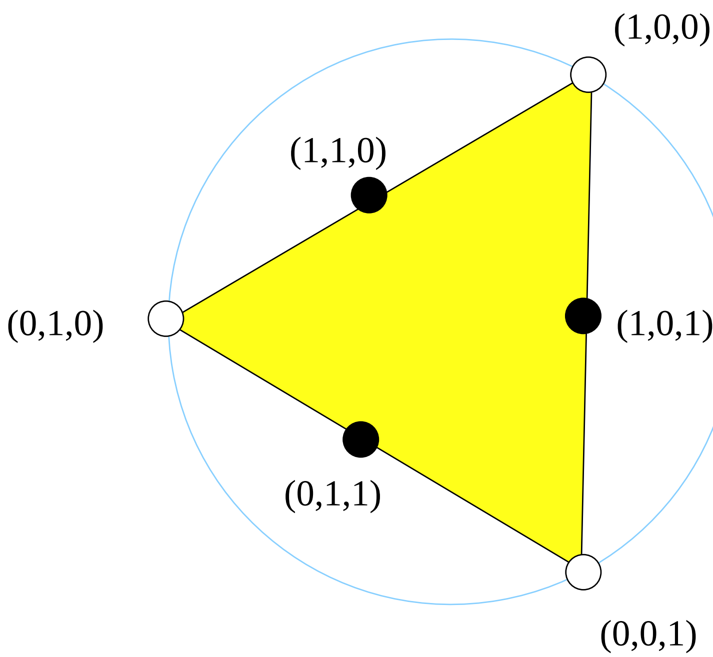

Figure 1: Projection to Π.

The reader familiar with this stuff will recognize that the disk I am describing is the *Klein model* of the hyperbolic plane. The yellow triangle is known as an *ideal triangle* in this model. The unit circle is known as the *ideal boundary*.

6. One can also visualize the action of the "reflections" R_1, R_2, R_3 on the unit disk. These maps act as real projective transformations of the unit disk – i.e. homeomorphisms that carry line segments to line segments. Call the yellow triangle Δ. Figure 2 shows the three triangles $R_k(\Delta)$ for $k = 1, 2, 3$. We have also drawn in some additional points. The new labels are in black. Our new points are the images of the old ones under the reflections. For instance $(-1, 3, 2) = R_1(1, 1, 0)$.

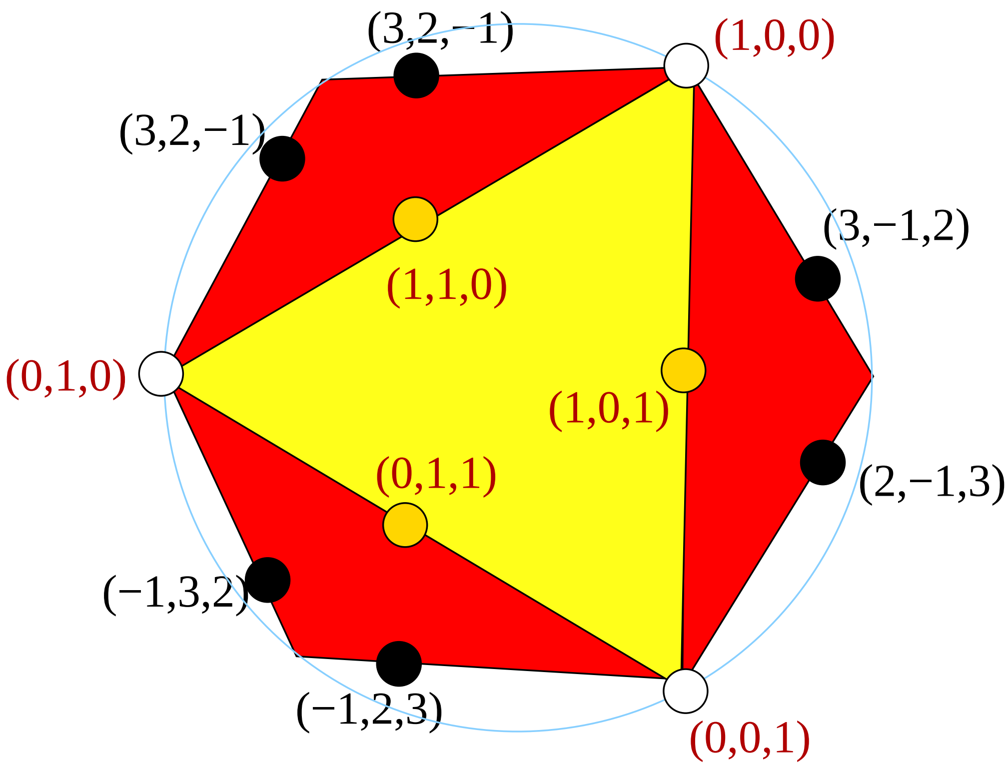

Figure 1: More triangles and points.

One can see that the new points are just the permutations of the good vectors corresponding to the Euler relations. Each ideal triangle has a natural "hyperbolic midpoint". For the yellow triangle the hyperbolic midpoints coincide with the Euclidean ones. For the red triangle, the hyperbolic midpoints appear to be "offcenter" from our Euclidean perspective. The labelled points, however, are the hyperbolic midpoints of the relevant triangles. This property comes from the fact that our reflections are *hyperbolic isometries*: They preserve the natural distance structure defined in the hyperbolic plane.

7. Say that a *word* is a finite composition of the reflections, e.g. R_2R_3 or $R_1R_2R_3R_2$. The group generated by R_1, R_2, R_3 consists of all words. This group is known as the ideal triangle group. The orbit of the central yellow triangle under this group is the famous Farey tiling. Every edge of the Farey triangulation has a hyperbolic center. This point has the property that hyperbolic rotation about this point is a symmetry of the Farey tiling. We say that an edge is *central* if it is an edge of the central yellow triangle; otherwise we say it is non-central. All but 3 Farey edges are non-central.

Say that a *Farey midpoint* is the hyperbolic midpoint of a Farey edge. Call the vectors $(0,1,1)$ and $(1,0,1)$ and $(1,1,0)$ the *central Farey midpoints*. These are the hyperbolic midpoints of the central Farey edges. Here is the main result.

Theorem 0.1 *The hyperbolic midpoints of the non-central Farey edges are in bijection with the good vectors.*

Proof: Let's first show that any good vector is a Farey midpoint. Suppose that V is a good vector. Then there is some word W such that $V' = W(V)$ is contained in the yellow triangle. By induction, W is L-preserving, and also preserves integer points. But then $V' = (a', b', c')$ is such that a', b', c' are all non-negative integers and $a'b' + b'c' + c'a' = 1$. This forces V' to be a central Farey midpoint. Hence $V = W^{-1}(V')$ is a Farey midpoint.

Applying elements of the group $\langle R_1, R_2, R_3\}$ repeatedly, and using induction, we see that any Farey midpoint (a,b,c) satisfies the relation $R(a,b,c)$ and has integer coordinates. We just have to check that a, b, c are all positive. if $a = 0$, then $bc = 1$, and this forces $|b| = |c| = 1$. The positive coordinate sum then forces $b = c = 1$. This gives us one of the central Farey midpoints. ♠