A Rational Approach to Trigonometry

N J Wildberger
School of Maths UNSW Sydney
Australia 2052

January 24, 2007

The trigonometry which you learnt at school is *the wrong trigonometry*. Young people over the ages have been taught a theory which unnecessarily complicates the subject and leads to loss of accuracy in practical applications. Unfortunately continual repetition has cemented this approach in the minds of educators as the only one possible—as you shall see, this is a mistake.

The thinking for the last two thousand years rests on the false assumptions that *distance* is the best way of measuring the separation of points, and that *angle* is the best way of measuring the separation of lines. So in order to study triangles students must first understand circles; they learn about π, lengths of circular arcs and the transcendental circular functions such as $\cos\theta$ and $\sin\theta$ that relate arc length on a circle to x and y projections. They study the relations between the circular functions and their inverse functions, ponder complicated graphs, and try to remember lots of special values. Advanced students see the infinite power series that calculators use to approximate true values. This is complicated stuff, especially if you try to do it correctly.

Yet a triangle itself is seemingly quite an elementary object. Why should the theory of a triangle be so complicated? Until now there has been no reasonable alternative. So educators have resigned themselves to the difficulties, and each year millions of students memorize the formulas, pass the tests (or not), and then promptly forget the unpleasant experience. And mathematicians wonder why the general public regards their beautiful subject with distaste bordering on hostility.

In this article, I am going to explain to you the right approach—called **rational trigonometry**. It is based on the idea that *algebra* is more basic than *analysis*, and that the true measurements in elementary geometry should be *quadratic* rather than *linear* in nature. Because this is a short paper, and I want to impart to you a working knowledge of the subject, the proofs are short. Giving more details is not hard, and the results are simple enough for a high school course.

Much more information can be found in '*Divine Proportions: Rational trigonometry to Universal Geometry*' ([Wildberger]). There the theory is developed over a general field (not of characteristic two) and many useful new formulas and applications appear. This is then used to develop Euclidean geometry in a general and powerful way which fully exploits the power of the Cartesian coordinates.

This paper is an introduction to these ideas, accessible to high school students.

1 Distance and angle

Aside from being clumsy, the current treatment is also logically dubious. Currently the basic concepts of classical planar trigonometry are distance and angle. Informally, *distance* is what you measure with a ruler. More precisely it is given by the formula

$$|A_1, A_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

for points $A_1 \equiv [x_1, y_1]$ and $A_2 \equiv [x_2, y_2]$. Informally, *angle* is what you measure with a protractor. But what is an angle precisely? As the ancient Greeks realized, this is a tricky question.

One way is to define it in terms of arclengths of circular arcs, but this requires a prior understanding of calculus. Another common approach is to try to define an angle θ as $\arctan(y/x)$, possibly plus or minus π, but this requires a prior understanding of the arctan function. Well, $\arctan x$ is the inverse function of $\tan\theta$, which is itself the ratio $\sin\theta/\cos\theta$, and $\sin\theta$ and $\cos\theta$ are defined in terms of ratios or projections involving the angle θ. And the angle θ is defined in terms of ... heh, what's going on here?

Waffling about angles is supplemented with imprecision about the definitions or the circular (trig) functions and their properties—a subject which really belongs to the theory of special functions. The rational approach, described below, banishes hand waving—along with distance, angle and the circular functions—from the study of triangles.

2 Points and quadrance

A **point** A is an ordered pair $[x, y]$ of numbers. The **quadrance** between two points $A_1 \equiv [x_1, y_1]$ and $A_2 \equiv [x_2, y_2]$ is the number

$$Q(A_1, A_2) \equiv (x_2 - x_1)^2 + (y_2 - y_1)^2 .$$

In diagrams a quadrance is displayed beside a thin rectangle along the line between two points, or alternatively in a rectangular box, to distinguish it from distance. If three points A_1, A_2 and A_3 lie on a line, meaning that they satisfy a linear equation of the form $ax + by + c = 0$, then the three quadrances

$$Q_1 \equiv Q(A_2, A_3) \qquad Q_2 \equiv Q(A_1, A_3) \qquad Q_3 \equiv Q(A_1, A_2)$$

satisfy the **Triple quad formula**

$$(Q_1 + Q_2 + Q_3)^2 = 2\left(Q_1^2 + Q_2^2 + Q_3^2\right).$$

Conversely if the Triple quad formula is satisfied then the three points are collinear. This is the first and most fundamental fact in the subject. It is a good exercise in algebraic manipulation once you write down what it means for three points to be collinear. The Triple quad formula can also be rewritten in the asymmetrical but still useful form $(Q_1 + Q_2 - Q_3)^2 = 4Q_1Q_2$.

If A_1, A_2, A_3 and A_4 are four collinear points, then the relation between the quadrances $Q_{12} = Q(A_1, A_2)$, $Q_{23} = Q(A_2, A_3)$, $Q_{34} = Q(A_3, A_4)$ and $Q_{14} = Q(A_1, A_4)$ is more subtle. It turns out to be

$$\left((Q_{12} + Q_{23} + Q_{34} + Q_{14})^2 - 2\left(Q_{12}^2 + Q_{23}^2 + Q_{34}^2 + Q_{14}^2\right)\right)^2 = 64 Q_{12} Q_{23} Q_{34} Q_{14}.$$

This **Quadruple quad formula** plays an important role in the study of quadrilaterals.

3 Lines and spreads

For distinct points A_1 and A_2 the unique line passing through them both is denoted $A_1 A_2$. Two lines l_1 and l_2 with respective equations $a_1 x + b_1 y + c_1 = 0$ and $a_2 x + b_2 y + c_2 = 0$ are **parallel** precisely when

$$a_1 b_2 - a_2 b_1 = 0$$

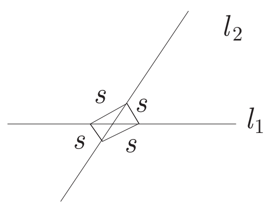
and **perpendicular** precisely when

$$a_1 a_2 + b_1 b_2 = 0.$$

The **spread** between these two lines is the number

$$s(l_1, l_2) = \frac{(a_1 b_2 - a_2 b_1)^2}{(a_1^2 + b_1^2)(a_2^2 + b_2^2)}.$$

In diagrams a spread $s = s(l_1, l_2)$ is displayed beside a small line segment joining the two lines to distinguish it from an angle; there are four equivalent positions as shown.



The spread $s(l_1, l_2)$ is 0 precisely when the lines l_1 and l_2 are parallel, and because of Fibonacci's identity

$$(a_1 b_2 - a_2 b_1)^2 + (a_1 a_2 + b_1 b_2)^2 = \left(a_1^2 + b_1^2\right)\left(a_2^2 + b_2^2\right)$$

it is 1 precisely when the lines are perpendicular. Otherwise the spread is a number between 0 and 1, and can be measured in practise with a **spread protractor as** shown in Figure 1. (Thanks to M. Ossmann at http://ossmann.com/protractor for this diagram.)

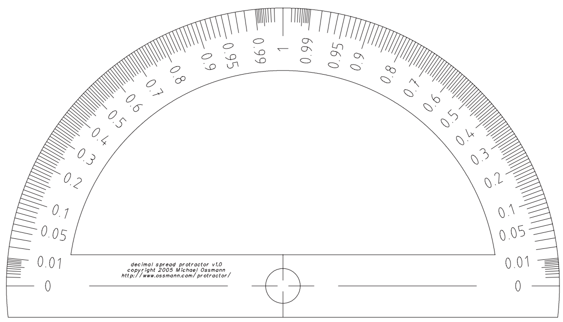


Figure 1: A spread protractor

The spread has a simple geometrical interpretation. Suppose l_1 and l_2 intersect at the point A. Choose a point $B \neq A$ on one of the lines, say l_1, and let C be the foot of the perpendicular from B to l_2 as in Figure 2.

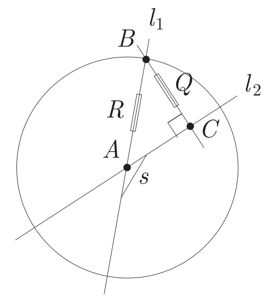


Figure 2: Spread as ratio

Then the spread between l_1 and l_2 is also

$$s\left(l_1, l_2\right) \equiv s \equiv \frac{Q\left(B, C\right)}{Q\left(A, B\right)} = \frac{Q}{R}.$$

Note that the circle plays no role in the definition. You may easily check that the spread corresponding to 30° or 150° is $s = 1/4$, the spread corresponding to 45° or 135° is $1/2$, and the spread corresponding to 60° or 120° is $3/4$. Spread does not distinguish between an angle and its supplement, since the spread is naturally measured between *lines*, not *rays*.

However when discussing rays it is possible for the spread of a sector to be either *acute* or *obtuse*. In Figure 3 the left diagram illustrates an acute spread of $s = 0.625$ and the right diagram an obtuse spread of 0.845. As Euclid realized, the acuteness of the sector on the left is equivalent to the condition

$$Q\left(A_1, A_2\right) + Q\left(A_1, A_3\right) \geq Q\left(A_2, A_3\right)$$

while the obtuseness of the sector on the right is equivalent to

$$Q\left(A_1, A_2\right) + Q\left(A_1, A_3\right) \leq Q\left(A_2, A_3\right).$$

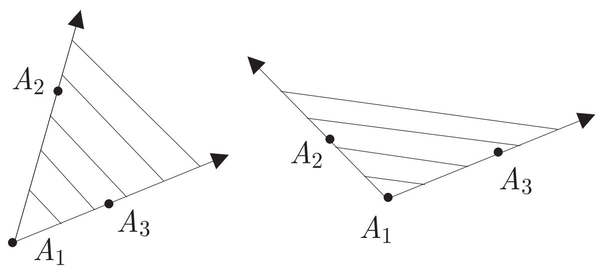


Figure 3: Acute and obtuse sectors

4 Laws of rational trigonometry

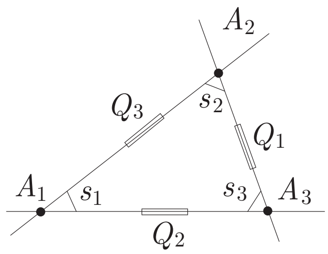


Figure 4: Quadrances and spreads of a triangle

Given three distinct points A_1, A_2 and A_3, define the quadrances

$$Q_1 \equiv Q(A_2, A_3) \qquad Q_2 \equiv Q(A_1, A_3) \qquad Q_3 \equiv Q(A_1, A_2)$$

and the spreads

$$s_1 \equiv s(A_1A_2, A_1A_3) \qquad s_2 \equiv s(A_2A_1, A_2A_3) \qquad s_3 \equiv s(A_3A_1, A_3A_2)$$

as in Figure 4. Here are the main laws of planar rational trigonometry, aside from the Triple quad formula stated earlier.

Pythagoras' theorem The lines A_1A_3 and A_2A_3 are perpendicular precisely when

$$Q_1 + Q_2 = Q_3.$$

Spread law For any triangle $\overline{A_1A_2A_3}$

$$\frac{s_1}{Q_1} = \frac{s_2}{Q_2} = \frac{s_3}{Q_3}.$$

Cross law For any triangle $\overline{A_1A_2A_3}$ define the **cross** $c_3 \equiv 1 - s_3$. Then

$$(Q_1 + Q_2 - Q_3)^2 = 4Q_1Q_2c_3.$$

Triple spread formula For any triangle $\overline{A_1A_2A_3}$

$$(s_1 + s_2 + s_3)^2 = 2\left(s_1^2 + s_2^2 + s_3^2\right) + 4s_1s_2s_3.$$

The Spread law is the rational analog of the Sine law, the Cross law is the analog of the Cosine law, and the Triple spread formula is the analog of the fact that the sum of the angles is π. However the proofs we now give are *independent of the usual treatment.*

5 Proofs of the laws

We'll assume that Pythagoras' theorem is known. Consider either of the following diagrams (the proofs work for both simultaneously).

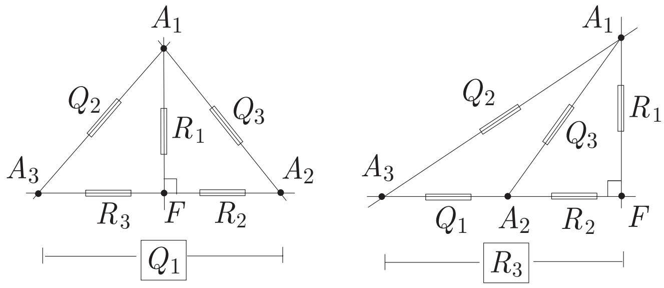


Figure 5: Spread and Cross laws

The spreads at A_2 and A_3 are respectively

$$\begin{aligned} s_2 &= \frac{R_1}{Q_3} \\ s_3 &= \frac{R_1}{Q_2}. \end{aligned}$$

Solve for R_1 to get

$$R_1 = Q_3 s_2 = Q_2 s_3$$

so that

$$\frac{s_2}{Q_2} = \frac{s_3}{Q_3}.$$

In a similar manner

$$\frac{s_1}{Q_1} = \frac{s_2}{Q_2}.$$

This proves the Spread law. To prove the Cross law, use the same diagram. Pythagoras' theorem shows that

$$\begin{aligned} Q_3 &= R_1 + R_2 \\ Q_2 &= R_1 + R_3. \end{aligned}$$

By the definition of the cross

$$c_3 = 1 - s_3 = 1 - R_1/Q_2 = R_3/Q_2.$$

Solve sequentially for R_3, R_1 and then R_2 to get

$$\begin{aligned} R_3 &= Q_2 c_3 \\ R_1 &= Q_2 (1 - c_3) \\ R_2 &= Q_3 - Q_2 (1 - c_3). \end{aligned}$$

Since A_2, A_3 and F are collinear, apply the Triple quad formula to the three quadrances Q_1, R_2 and R_3, yielding

$$(Q_1 + R_3 - R_2)^2 = 4Q_1R_3.$$

Substitute the values of R_3 and R_2, to get

$$(Q_1 + Q_2 - Q_3)^2 = 4Q_1Q_2c_3.$$

This proves the Cross law.

From the Spread law, there is a non-zero number D such that

$$\frac{s_1}{Q_1} = \frac{s_2}{Q_2} = \frac{s_3}{Q_3} = \frac{1}{D}. \tag{1}$$

The Cross law can be rewritten as

$$(Q_1 + Q_2 + Q_3)^2 = 2\left(Q_1^2 + Q_2^2 + Q_3^2\right) + 4Q_1Q_2s_3. \tag{2}$$

Use (1) to replace Q_1 by s_1D, Q_2 by s_2D and Q_3 by s_3D in (2), and then divide by D^2. The result is the Triple spread formula

$$(s_1 + s_2 + s_3)^2 = 2\left(s_1^2 + s_2^2 + s_3^2\right) + 4s_1s_2s_3.$$

6 Rational trigonometry in action

That's it. You have now learnt the basics of rational trigonometry. Tables of trig functions and calculators can be dispensed with (although calculators are still useful for those large fractions that often come up). Knowing that

$$\sec\theta = 1 + \frac{1}{2}\theta^2 + \frac{5}{24}\theta^4 + \frac{61}{720}\theta^6 + \frac{277}{8064}\theta^8 + \frac{50\,521}{3628\,800}\theta^{10} + \cdots$$

or that $\sin(\arctan x) = x/\sqrt{1+x^2}$ or that $\tan(\pi/5) = \sqrt{5 - 2\sqrt{5}}$ is not necessary. You can delete all such formulas from your over-crowded memory banks, and still solve geometrical problems faster and with greater accuracy than your classical classmates.

With a bit of practice studying the examples in '*Divine Proportions*', I assure you that without a calculator—i.e. just pencil and paper—you will demolish all similarly equipped competition when it comes to solving trigonometric problems quickly to high accuracy. I guess that on average you will be at least twice as fast, and often much faster. If you want, answers with rational trigonometry can be 100% accurate. The implications for computing should be obvious.

Of course some additional formulas are useful, but with the basic laws you can go far. Here are just two examples of the power of this method, which illustrate some additional points. The first treats a classical problem involving slopes, and the second one shows the rich interplay between number theory and geometry that rational trigonometry brings to the fore. For more examples, see [Wildberger].

Problem 1 *Suppose an inclined plane has a spread of s with the horizontal plane. An insect climbing up the plane walks on a straight line which makes a spread of r with the path of greatest slope. At what spread to the horizontal does the insect climb on this path?*

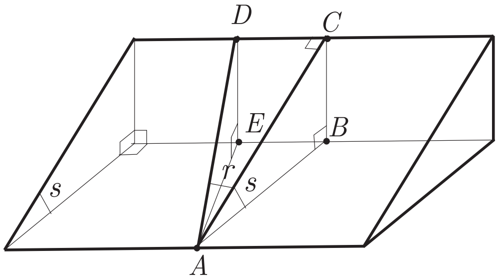


Figure 6: Path on a wedge

Solution. As in Figure 6, let AC be a line making the maximum possible spread s with the horizontal, and AD the path of the insect. The spread between AC and AD is r, and you need to find the spread $s(AE, AD)$. Suppose that $Q(A, C) = Q$. Then from the definition of the spread s,

$$Q(B, C) = sQ.$$

In the right triangle $\overline{ACD}$, the spread $s(AC, AD)$ is r, so by Pythagoras' theorem the spread $s(DA, DC)$ is $1 - r$. The Spread law then gives

$$\frac{1}{Q(A, D)} = \frac{r}{Q(C, D)} = \frac{1 - r}{Q}.$$

So

$$Q(A, D) = \frac{Q}{1 - r}.$$

Thus in the right triangle $\overline{ADE}$

$$\begin{aligned} s(AD, AE) &= \frac{Q(D, E)}{Q(A, D)} \\ &= \frac{Q(B, C)}{Q(A, D)} = \frac{sQ}{Q/(1 - r)} \\ &= s(1 - r). \end{aligned}$$

Note that this is a linear expression in both s and r. ∎

Problem 2 *The triangle $\overline{A_1 A_2 A_3}$ shown to scale in the left diagram of Figure 7 has an angle at A_1 of $60°$, with $|A_1, A_3| = \sqrt{57}$, $A_3 D$ the bisector of the angle at A_3 and $|D, A_3| = \sqrt{43}$. The triangle $\overline{A_1 A_3 D}$ is acute. What is the distance $d_1 = |A_2, A_3|$ exactly? The corresponding rational formulation of the same problem is in the right diagram of the same figure. The triangle $\overline{A_1 A_2 A_3}$*

has a spread of $s_1 = 3/4$, *quadrance* $Q_2 \equiv 57$ *and quadrance* $Q(A_3, D) \equiv 43$. *In this case we are to determine the quadrance* Q_1 *exactly.*

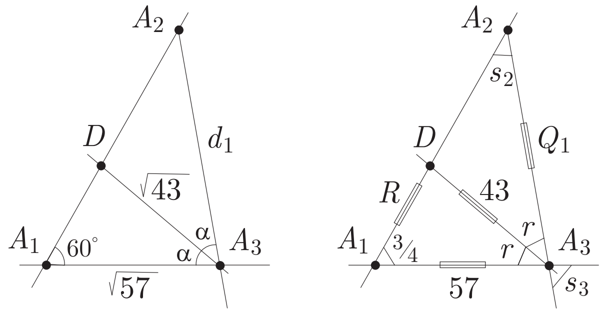


Figure 7: Classical and rational versions

Solution. Suppose that $R \equiv Q(A_1, D)$ and $s(A_3A_1, A_3D) = s(A_3A_2, A_3D) \equiv r$ as in the Figure. The Cross law in $\overline{A_1A_3D}$ gives

$$(R + 57 - 43)^2 = 4 \times R \times 57 \times (1 - 3/4)$$

or

$$R^2 - 29R + 196 = 0.$$

This has solutions $R = \frac{29}{2} \pm \frac{1}{2}\sqrt{57}$ which are approximately 18.27 and 10.72. But we are told that $\overline{A_1A_3D}$ is acute; this implies that $R + 43 \geq 57$. So the correct solution is $R = \frac{29}{2} + \frac{1}{2}\sqrt{57}$. The Spread law in $\overline{A_1A_3D}$ shows that

$$\frac{r}{R} = \frac{3/4}{43}$$

so that

$$r = \frac{87}{344} + \frac{3}{344}\sqrt{57} \approx 0.318\,7.$$

The three lines meeting at A_3 make spreads of r, r and s_3. The Triple spread formula applies, and in this case reduces to

$$\begin{aligned} s_3 &= 4r(1-r) \\ &= \frac{10\,923}{14\,792} + \frac{255}{14\,792}\sqrt{57} \approx 0.868. \end{aligned}$$

Note the appearance of the *logistic map* $f(x) = 4x(1-x)$ in this context; this is an example of a *Spread polynomial.* Having $s_1 = 3/4$ and s_3, the Triple spread formula

$$(s_1 + s_2 + s_3)^2 = 2\left(s_1^2 + s_2^2 + s_3^2\right) + 4s_1s_2s_3$$

in $\overline{A_1A_2A_3}$ gives for s_2 the quadratic equation

$$s_2^2 - s_2\left(\frac{11\,265}{14\,792} - \frac{255}{14\,792}\sqrt{57}\right) + \frac{1867\,833}{109\,401\,632} - \frac{43\,605}{109\,401\,632}\sqrt{57} = 0.$$

Somewhat surprisingly, the solutions are

$$s_2 = \frac{171}{7396} \approx 0.023$$

or

$$s_2 = \frac{10\,923}{14\,792} - \frac{255}{14\,792}\sqrt{57} \approx 0.608.$$

The second one is the relevant one, and then by the Spread law

$$\begin{aligned} Q_1 &= \frac{3}{4} \times \frac{57}{s_2} \\ &= \frac{383\,735\,913}{6422\,528} + \frac{8958\,405}{6422\,528}\sqrt{57} \approx 70.279. \end{aligned}$$

If a distance is required, take the square root of the quadrance Q_1. This turns out to be, also somewhat surprisingly,

$$d_1 = \frac{2451}{3584} + \frac{3655}{3584}\sqrt{57} \approx 8.383. \quad \blacksquare$$

It should be remarked that the above numbers were essentially chosen at random—this is not a 'rigged' example. It is interesting to try this problem with classical trigonometry to compare the two approaches.

7 Conclusion

Rational trigonometry is generally superior to classical trigonometry for solving the majority of geometric problems. It cleanly separates the physical subject of uniform circular motion and the mathematical subject of triangles. For the former, the circular functions are of course useful, for the latter they are—or should be—largely irrelevant.

In scientific circles, it is reasonably common for new developments to replace rather than to augment old theories. This gives scientific research a youthful vigor that is sometimes lacking in mathematics, where tradition generally rules. Here will be a good test case. Will the merits of rational trigonometry be properly analyzed and discussed, and the theory implemented if found superior?

If you try your hand at a wide range of problems using rational trigonometry, you will see for yourself the power of this technology. The new algebraic approach is cleaner, faster and more accurate than the old analytic one. Rational trigonometry is far easier to learn and to teach, and so opens up an exciting new vista for mathematics education.

References

[Wildberger] N J Wildberger, *Divine Proportions: Rational Trigonometry to Universal Geometry*, Wild Egg (http://wildegg.com), Sydney 2005.