Pythagoras, Euclid, Archimedes and a new Trigonometry

N J Wildberger
School of Mathematics UNSW
Sydney 2052 Australia
webpages: http://web.maths.unsw.edu/~norman/

October 13, 2006

Abstract

Pythagoras' theorem, Euclid's formula for the area of a triangle as one half the base times the height, and Heron's or Archimedes' formula are amongst the most important and useful results of ancient Greek geometry. Here we look at all three in a new and improved light, giving a dramatically simpler and more elegant trigonometry.

Three great theorems

There are three classical theorems about triangles that every student meets. We work with a triangle $\overline{A_1 A_2 A_3}$ with side lengths $d_1 = |A_2, A_3|$, $d_2 = |A_1, A_3|$ and $d_3 = |A_1, A_2|$.

Pythagoras' theorem *The triangle* $\overline{A_1 A_2 A_3}$ *has a right angle at* A_3 *precisely when*

$$d_1^2 + d_2^2 = d_3^3.$$

Euclid's theorem *The area of a triangle is one half the base times the height.*

Heron's theorem *If* $s = (d_1 + d_2 + d_3)/2$ *is the semi-perimeter of the triangle, then*

$$\text{area} = \sqrt{s(s-d_1)(s-d_2)(s-d_3)}.$$

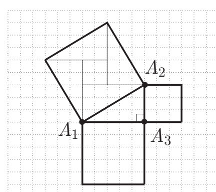
Unfortunately these usual formulations are flawed from a pure mathematical perspective. In this paper we will recast all three in a simpler and more general form. As a reward, we find that *rational trigonometry* falls into our laps, essentially for free. Note however that we do not prove these results, as to do this carefully requires some work—the interested reader may consult [2].

Our reformulations of these results turn out to work over a general field (not of characteristic two), in arbitrary dimensions, and even with an arbitrary quadratic form. The justification for these assertions is found in [2] and [3].

Pythagoras' theorem

The ancient Greeks regarded *area*, not *distance*, as the fundamental quantity in planar geometry. Indeed they worked with a straightedge and compass in their constructions, not a ruler. A line segment was measured by constructing a square on it, and determining the area of that square. Two line segments were considered equal if they were congruent, but this was independent of a direct notion of distance measurement.

For the ancient Greeks, Pythagoras' theorem was a relation not about distances but rather about areas: the squares built on each of the sides of a right triangle. From a Cartesian point of view, this is still a particularly attractive way to introduce students to the subject, since with a sheet of graph paper, area of simple figures is directly accessible by subdividing, rearranging, and counting cells.



A triangle with side areas $9, 25$ and 34.

The triangle $\overline{A_1A_2A_3}$ shown has smaller sides of lengths 5 and 3, but the length of the longer side is not so easy to determine by measurement. The diagram shows the two smaller squares with areas 25 and 9 respectively, as we just count cells. To find the area of the larger square, subdivide it into a smaller 2×2 square and four right triangles which combine to form two 5×3 rectangles, for a total of 34. So in this case Pythagoras' theorem is a result which can be established by *counting*.

Every important notion should have a name. Following the Greek terminology of 'quadrature', we define the **quadrance** Q of a line segment to be the area of the square constructed on it. Pythagoras' theorem allows us to assert that if $A_1 = [x_1, y_1]$ and $A_2 = [x_2, y_2]$, then the quadrance between A_1 and A_2 is

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

So for example the quadrance between the points $[0,0]$ and $[1,2]$ is $Q = 5$. The usual distance is the 'square root' of the quadrance. No doubt you regard the 'square root' as such a basic function as to not justify any quotes, and you believe that you *really understand* a 'number' like $\sqrt{5}$.

Let me be honest with you—as a pure mathematician, I *do not understand* this 'number'. For me $\sqrt{5} = 2.236\,067\,977\,499\,789\,696\,409\,173\,668\,731\,276\,235\,440\,618\,359\,611\,525\,724\,270\,897\,245\,410\,520\,925\,637\,804\,899\,414\,414\,41\ldots$ is specified by its decimal expansion, which is largely a mystery to me. Almost any interesting question you ask me about this sequence I would find difficult, perhaps impossible, to answer, so I much prefer to work with the number 5.

In statistics, variance is more fundamental than standard deviation, and least squares analysis rules. In quantum mechanics, wavefunctions are more fundamental than probability amplitudes. In harmonic analysis, L^2 is more pleasant than L^1. In geometry, *quadrance* is *more fundamental* than *distance*. The ubiquitious nature of squared quantities in Euclidean geometry show that this is not altogether a new point of view, and it deserves your contemplation!

For a triangle $\overline{A_1A_2A_3}$ we define the quadrances $Q_1 = Q(A_2, A_3)$, $Q_2 = Q(A_1, A_3)$ and $Q_3 = Q(A_1, A_2)$. Here then is the *true Pythagoras' theorem*—it holds over arbitrary fields, in arbitrary dimensions, and in fact turns out to be valid with arbitrary quadratic forms, although that is beyond the scope of this article.

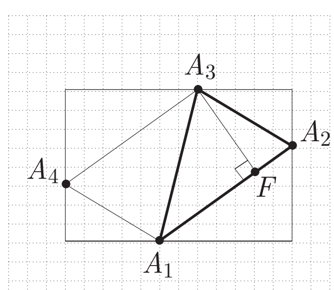
Theorem 1 (Pythagoras) *The lines A_1A_3 and A_2A_3 of the triangle $\overline{A_1A_2A_3}$ are perpendicular precisely when*

$$Q_1 + Q_2 = Q_3.$$

It is well worth noting that Pythagoras' theorem in the Cartesian setting is more than just a definition of quadrance/distance, as some texts erroneously suggest. Although we implicitly use Pythagoras' theorem to *define* quadrance, so that the theorem is automatically true for triangles with two legs along the coordinate axes, one is still required to *prove* it for the general case when triangle sides are not along the coordinate axes.

Euclid's theorem

The area of a triangle is one-half the base times the height. This is a splendid and most useful engineering, or applied mathematics, rule. However as a theorem of pure mathematics, it is seriously flawed! Let's see why, by looking at the area of the triangle $\overline{A_1A_2A_3}$ in the following figure.



A triangle and an associated parallelogram

This area is half of the area of the associated parallelogram $\overline{A_1A_2A_3A_4}$, which is the result of removing from a 12×8 rectangle four triangles, which can be combined to form two rectangles, one 5×3 and the other 7×5. The area of $\overline{A_1A_2A_3}$ is thus 23.

To apply the one-half base times height rule, the base $\overline{A_1A_2}$ by Pythagoras has length

$$d_3 = |A_1, A_2| = \sqrt{7^2 + 5^2} = \sqrt{74} \approx 8.602\,325\,267\,04\ldots.$$

To find the length h of the altitude $\overline{A_3F}$ is somewhat more work. If we set the origin to be at A_1, then the line A_1A_2 has Cartesian equation $5x - 7y = 0$ while $A_3 = [2, 8]$. A well-known result from coordinate geometry states that then the distance $h = |A_3, F|$ from A_3 to the line A_1A_2 is

$$h = \frac{|5 \times 2 - 7 \times 8|}{\sqrt{5^2 + 7^2}} = \frac{46}{\sqrt{74}} \approx 5.347\,391\,382\,22\ldots.$$

If the engineer doing this calculation works with the surd forms of both expressions, she will notice that the two $\sqrt{74}$'s conveniently cancel, giving an area of 23. However if she works immediately with the decimal forms (which is probably on the whole more natural) she obtains with her calculator

$$\text{area} \approx \frac{8.602\,325\,267\,04\ldots \times 5.347\,391\,382\,22\ldots}{2} \approx 23.0.$$

So you can see what the problem is—the one-half base times height rule means we descend to the level of square roots, even when the eventual answer is a rational number. This introduces unnecessary approximations and inaccuracies into the subject. Here is a better version.

Theorem 2 (Euclid) *The square of the area of a triangle is one-quarter the quadrance of the base times the quadrance of the altitude.*

As a formula, this would be

$$\text{area}^2 = \frac{Q \times H}{4}$$

where Q is the quadrance of the base and H is the quadrance of the altitude to that base.

Heron's or Archimedes' Theorem

The same triangle $\overline{A_1A_2A_3}$ of the previous section has side lengths

$$d_1 = \sqrt{34} \qquad d_2 = \sqrt{68} \qquad d_3 = \sqrt{74}.$$

The semi-perimeter s is then

$$s = \frac{\sqrt{34} + \sqrt{68} + \sqrt{74}}{2} \approx 11.339\,744\,206\,6\ldots.$$

Using the usual Heron's theorem, a computation with the calculator shows that

$$\text{area} = \sqrt{s\left(s-\sqrt{34}\right)\left(s-\sqrt{68}\right)\left(s-\sqrt{74}\right)} = 23.0.$$

Again we have a formula involving square roots in which there appears to be a surprising rational outcome. The historical record makes it pretty clear that Archimedes' knew Heron's theorem before Heron did, and the most important mathematician of all time deserves credit for more than he currently gets. So here is a better form of Heron's theorem, with a more appropriate name.

Theorem 3 (Archimedes) *The area of a triangle $\overline{A_1A_2A_3}$ with quadrances Q_1, Q_2 and Q_3 is given by*

$$16 \times \text{area}^2 = (Q_1+Q_2+Q_3)^2 - 2\left(Q_1^2+Q_2^2+Q_3^2\right).$$

In our example the triangle has quadrances $34, 68$ and 74, each obtained by Pythagoras' theorem. So the formula becomes

$$16 \times \text{area}^2 = (34+68+74)^2 - 2\left(34^2+68^2+74^2\right) = 8464$$

and this gives $\text{area}^2 = 23^2$.

It is instructive to see why the two formulations of Heron's/Archimedes' theorems are equivalent. From Heron's theorem,

$$\begin{aligned}
16 \times \text{area}^2 &= (d_1+d_2+d_3)(-d_1+d_2+d_3)(d_1-d_2+d_3)(d_1+d_2-d_3) \\
&= \left((d_1+d_2)^2-d_3^2\right)\left(d_3^2-(d_1-d_2)^2\right) \\
&= \left((d_1+d_2)^2+(d_1-d_2)^2\right)Q_3-(d_1+d_2)^2(d_1-d_2)^2-Q_3^2 \\
&= 2(Q_1+Q_2)Q_3-\left(d_1^2-d_2^2\right)^2-Q_3^2 \\
&= 2(Q_1+Q_2)Q_3-(Q_1-Q_2)^2-Q_3^2 \\
&= 2Q_1Q_2+2Q_1Q_3+2Q_2Q_3-Q_1^2-Q_2^2-Q_3^2 \\
&= (Q_1+Q_2+Q_3)^2-2\left(Q_1^2+Q_2^2+Q_3^2\right).
\end{aligned}$$

Archimedes' theorem implies a formula of remarkable simplicity and importance. The following is the fundamental result of *one-dimensional geometry*—a surprisingly rich topic as it turns out. (I will explain this remark more fully elsewhere).

Theorem 4 (Triple quad formula) *The three points A_1, A_2 and A_3 are collinear precisely when*

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

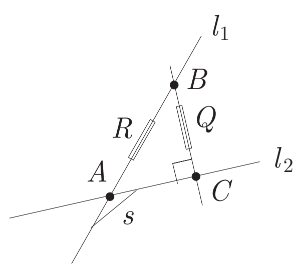
The proof is of course immediate, as collinearity is equivalent to the area of the triangle being zero.

Rational Trigonometry

An angle is a *circular distance*, that is distance measured along a circular arc, and this is *too complicated* a concept to qualify as fundamental for measuring the separation of two lines. To define angles properly you require calculus, a logical point that is rarely acknowledged by educators. There is a reason that classical trigonometry is so complicated and painful to students—it is based on the wrong notions!

Teachers of trigonometry constantly rely on $90-45-45$ and $90-60-30$ triangles for examples and test questions. Once you get the hang of rational trigonometry and the much wider scope for explicit triangles that can be completely analysed, you will appreciate just how limiting classical trigonometry is. You will marvel at how generations of mathematicians accepted this theory with scarcely a peep of protest! See [2] for a complete development of this exciting new theory. In what follows, we show how the basic ideas follow naturally from our presentation of the theorems of Pythagoras, Euclid and Archimedes.

The true separation between lines l_1 and l_2 is captured by the concept of *spread*, which may be defined as the ratio of two quadrances. 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.



Spread s between two lines l_1 and l_2

Then the **spread** s between l_1 and l_2 is

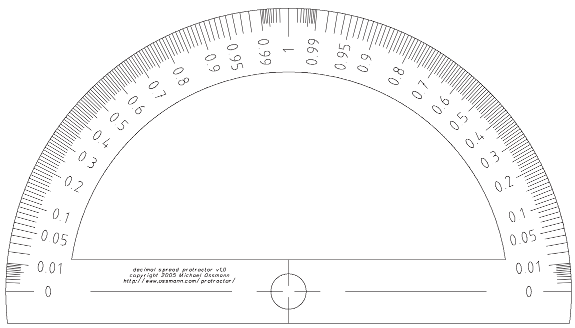
$$s = s(l_1, l_2) = \frac{Q(B,C)}{Q(A,C)} = \frac{Q}{R}.$$

This ratio is clearly independent of the choice of B, by Thales' theorem. The spread is defined between lines, not rays. Parallel lines are defined to have spread $s = 0$, while perpendicular lines have spread $s = 1$. you may check that the spread corresponding to $30°$ or $150°$ is $s = 1/4$, while the spread corresponding to $60°$ or $120°$ is $3/4$.

When lines are expressed in Cartesian form, the spread becomes a rational expression in the coefficients of the lines. It therefore makes sense over arbitrary fields, although there is the possibility of null lines for which the denominator involved in the spread is zero. Note that in the triangle $\overline{ABC}$ above, the spread at the vertex A and the spread at the vertex B sum to 1, on account of Pythagoras' theorem.

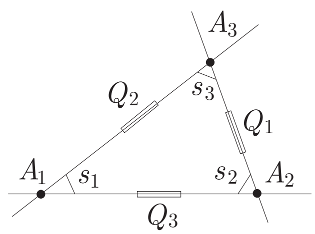
In diagrams a spread $s = s(l_1, l_2)$ is displayed beside a small line segment joining the two lines to distinguish it from angle.

The following spread protractor was created by Michael Ossmann and is available online at [1].



A spread protractor

So now a triangle $\overline{A_1A_2A_3}$ has quadrances Q_1, Q_2 and Q_3 as well as spreads s_1, s_2 and s_3, as in the following diagram.



Quadrances and spreads of a triangle

If H_3 is the quadrance of the altitude from A_3 to the line A_1A_2, then Euclid's theorem and the definition of spread give

$$\text{area}^2 = \frac{Q_3 \times H_3}{4} = \frac{Q_3Q_2s_1}{4} = \frac{Q_3Q_1s_2}{4}.$$

It follows by symmetry that

$$\frac{s_1}{Q_1} = \frac{s_2}{Q_2} = \frac{s_3}{Q_3} = \frac{4 \times \text{area}^2}{Q_1Q_2Q_3}.$$

This is the **Spread law**, the analog of the Sine law. By equating the formulas for $16 \times \text{area}^2$ given by Euclid's and Archimedes' theorems, we get

$$\begin{aligned} 4Q_3Q_2s_1 &= (Q_1 + Q_2 + Q_3)^2 - 2\left(Q_1^2 + Q_2^2 + Q_3^2\right) \\ &= 2Q_1Q_2 + 2Q_1Q_3 + 2Q_2Q_3 - Q_1^2 - Q_2^2 - Q_3^2. \end{aligned}$$

This can be rearranged in the form

$$(Q_1 - Q_2 - Q_3)^2 = 4Q_2Q_3(1 - s_3)$$

which is the **Cross law**, the analog of the Cosine law.

Now substitute $Q_1 = s_1D$, $Q_2 = s_2D$ and $Q_3 = s_3D$ from the Spread law into the Cross law and cancel the common factor of D^2. The result is the **Triple spread formula**

$$(s_1 - s_2 - s_3)^2 = 4s_2s_3(1 - s_3)$$

which can be rewritten more symmetrically as

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

This formula is a deformation of the Triple quad formula by a single cubic term, and is the analog in rational trigonometry to the classical fact that the three angles of a triangle sum to $3.141\,592\,653\,59\ldots$.

The *Triple quad formula*, *Pythagoras' theorem*, the *Spread law*, the *Cross law* and the *Triple spread formula* are the five main laws of rational trigonometry. These are implicitly contained in the geometrical work of the ancient Greeks.

As demonstrated at some length in [2], these formulas and a few additional secondary ones suffice to solve the majority of trigonometric problems, usually more simply, more accurately and more elegantly than the classical theory involving $\sin\theta$, $\cos\theta$, $\tan\theta$ and their inverse functions. As shown in [3], the same formulas extend to geometry over general fields and with arbitrary quadratic forms.

In retrospect, the blind spot first occurred with the Pythagoreans, who initially believed that all of nature should be expressible in terms of natural numbers and their proportions. When they discovered that the ratio of the length of a diagonal to the length of a side of a square was the incommensurable proportion $\sqrt{2} : 1$, they panicked, and threw the exposer of the secret overboard while at sea.

They should have maintained their beliefs in the workings of the Divine Mind, and concluded that the *squares of the lengths are the crucial quantities in geometry.* Had they grasped this essential point, mathematics would have had a significantly different history, Einstein's special theory of relativity would have been discovered earlier, algebraic geometry would have quite another aspect, and students would be studying a much simpler and more elegant trigonometry.

References

[1] M. Ossmann, 'Print a Protractor', download online at http://www.ossmann.com/protractor/

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

[3] N. J. Wildberger, *Affine and Projective Rational Trigonometry*, preprint, 2006.