SO YOU WANT TO KNOW THE CIRCUMFERENCE OF AN ELLIPSE?

CHRIS RACKAUCKAS

The circumference of an ellipse is a surprising problem due to the complexity that it has. Although the question "what is the circumference of an ellipse?" sounds fairly simple, it introduces mathematicians to what have became known as elliptic functions and the diverse properties of such functions. I will begin by giving an outline of how the mathematics behind the circumference an ellipse evolved to the form the definition of elliptic functions. After understanding what an elliptic function is, I will prove one of the key properties of an elliptic function that shows that elliptic functions cannot be expressed as a linear combinations of "elementary functions" (a set of functions which will be defined later). After realizing the complex nature of using elliptic functions to describe the circumference of an ellipse, I will turn towards approximation methods to see how the circumference of an ellipse can be approximated using elementary functions.

1. The Discovery of Elliptic Functions

A brief glimpse of how the arc length of a circle is found gives the method which is used to find the arc length of an ellipse. A circle centered at the origin with radius one is described in Cartisian coordinates as $x^2 + y^2 = 1$. Thus the equation for the upper half of the circle is given by $y = \sqrt{1 - x^2}$. To find the arc length of the curve, we use the formula that the arc length, L, is

$$L = \int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Thus by the normal rules of the derivative we get that

$$y' = -\frac{x}{\sqrt{1 - x^2}}.$$

Thus for $y = f(x)$, we get that

$$(f'(x))^2 = \frac{x^2}{1 - x^2},$$

$$1 + (f'(x))^2 = \frac{1 - x^2}{1 - x^2} + \frac{x^2}{1 - x^2} = \frac{1}{1 - x^2}.$$

and then by substitution

$$L = \int_{-1}^{1} \frac{1}{\sqrt{1 - x^2}} dx$$

where $a = -1$ and $b = 1$ since x goes from -1 to 1. However, by symmetry we know that

$$L = 2\int_0^1 \frac{1}{\sqrt{1-x^2}}dx.$$

Thus since by using the formula for arcsin, we get that

$$L = 2\arcsin(x)|_0^1 = 2(\arcsin(1) - \arcsin(0)) = 2(\frac{\pi}{2}) = \pi.$$

Since the formula for the circumference of the entire circle is two times the circumference of the upper half, we get that the circumference of a circle of radius 1 is 2π which is what we expected.

Now lets naively do the same for an ellipse. The Cartesian coordinates of an ellipse are given by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Thus we find that the circumference of the upper half is given by

$$y = \sqrt{b^2(1 - \frac{x^2}{a^2})},$$

$$y' = \frac{-bx}{a^2\sqrt{1-\frac{x^2}{a^2}}},$$

$$y'^2 = \frac{b^2x^2}{a^2(a^2-x^2)},$$

$$\sqrt{1+y'^2} = \sqrt{\frac{b^2x^2}{a^2(a^2-x^2)} + 1}.$$

Let $c = \frac{b}{a}$ and parametrize x by t where $t(-a) = 1$ and $t(a) = 1$ and we get that the arc length of the upper half is

$$L = \int_A^B \sqrt{1+(y')^2}dx = \int_{-a}^{a} \sqrt{\frac{b^2x^2}{a^2(a^2-x^2)} + 1}dx = \int_{-1}^{1} \sqrt{\frac{1-(c^2-1)t^2}{1-t^2}}dx.$$

This can be simplified by

$$\int_{-1}^{1} \sqrt{\frac{1-(c^2-1)t^2}{1-t^2}}\sqrt{\frac{1-(c^2-1)t^2}{1-(c^2-1)t^2}}dx = \int_{-1}^{1} \frac{1-(c^2-1)t^2}{\sqrt{(1-t^2)(1-(c^2-1)t^2)}}dx,$$

or

$$L = \int_{-1}^{1} \frac{1}{\sqrt{(1-t^2)(1-(c^2-1)t^2)}}dx - \int_{-1}^{1} \frac{(c^2-1)t^2}{\sqrt{(1-t^2)(1-(c^2-1)t^2)}}dx$$

With this we have stumbled onto an integral that is of the form of an elliptical integral of the second kind and thus we must stop since, as we will see later, this integral cannot be solved in terms of elementary functions.

2. Elliptic Integrals and Elliptic Functions

So now we see that in order to understand what the circumference of an ellipse we have to understand this type of integral, mainly an elliptic integral. By definition, an elliptic integral, such as the one derived above, is defined as the integral of $R[t, \sqrt{p(t)}]$ where R is a rational function and p is a polynomial of degree 3 or 4 without repeated roots. In other words, an elliptic integral can be expressed as

$$\int_0^x \frac{dt}{\sqrt{p(t)}}.$$

This fits with what we derived above where $p(t) = (1-t^2)(1-(c^2-1)t^2)$, and thus the first integral is an elliptic integral.

Properties of elliptic integrals can be understood by comparison to other functions. For example, the arcsin function is as such

$$\arcsin(t) = \int_0^x \frac{dt}{\sqrt{1-t^2}}$$

closely resembles elliptic integrals like the lemniscatitc integral, the integral to describe the arc length of the lemniscate of Bernoulli, which is

$$\int_0^x \frac{dt}{\sqrt{1-t^4}}.$$

Since we know that

$$2\arcsin(t) = 2\int_0^x \frac{dt}{\sqrt{1-t^2}} = \int_0^{2x\sqrt{1-x^2}} \frac{dt}{\sqrt{1-t^2}},$$

it helps lead us to the fact that

$$2\int_0^x \frac{dt}{\sqrt{1-t^4}} = \int_0^{2x\sqrt{1-x^4}/(1+x^4)} \frac{dt}{\sqrt{1-t^4}}.$$

which is the formula for doubling the length of the lemniscate of Bernoulli. Other properties such as the addition formulas can be understood by looking at functions in the same way, such as how there is an analogy between the angle addition formula for arcsin and the addition of arc lengths for the lemniscate.

However, one of the most important properties of elliptic integrals, one that applies to all elliptic integrals, has to be approached from a different direction. The property that I am talking about is that elliptic integrals cannot be solved in terms of elementary functions. Elementary functions are functions that are written as any linear combination of rational, circular (trigonometric), exponential, and/or logarithmic functions and their inverses. So this means that no elliptic integral can be written as elementary functions, or the functions generally seen in mathematics. This is the reason that all of the elliptic integrals before were left in terms that included the integral: there is no way to write it in terms of functions without the integral (unless you define some new function that is defined by that integral).

The implication of this is far reaching. However, for our purposes, it gives the most fascinating result of the whole paper: the circumference of an ellipse cannot be expressed in closed form by elementary functions. Unlike the circumference of a circle which we know of as $2\pi r$ where r is the radius, there is no way to express the circumference of an ellipse as a formula with known functions and no integrals. This is a direct result of the fact that elliptic integrals cannot be expressed by elementary functions since the circumference of an ellipse itself is found by an elliptic integral.

3. The Non-Integrability of Elliptic Integrals

To show that elliptic integrals cannot be written as a linear combination of elementary integrals, we will take for granted the theorem known as Loiuville's Theorem.

Theorem 1. *Let F be a differential field of characteristic zero. Take $\alpha \in F$ and $y \in F[t]$ where $F[t]$is some elementary differential extension field of F having the same subfield of constants. If the equation $y' = \alpha$ has a solution, then there exists constants $c_1, c_2, ..., c_n \in F$ and elements $u_1, u_2, ..., u_n \in F$ $v \in F[t]$ such that*

$$\alpha = \sum_{i=1}^{n} c_i \frac{u_i'}{u_i} + v'.$$

For the use of this paper, take F is be the field of rational functions, take the differential extension field of F to be F with an extra element t, or $F[t]$. Thus for our purposes, since the derivative of one of these functions is a rational function, and a rational function times another rational function and divided by a rational function is still an rational function, we can note that $\sum_{i=1}^{n} c_i \frac{u_i'}{u_i}$ is just some element of F. Thus we can simplify our expression by saying

$$\alpha = (element\ of\ F) + v'.$$

Since $v \in F[t]$, we can write v as

$$v = \sum_{j=0}^{m} b_j t^j,$$

where each term $b_j \in F$. By differentiating, we get that

$$v' = \sum_{j=0}^{m} b_j' t^j + j b_j t' t^{j-1}.$$

If j,$b_j \neq 0$, then $b_j' + j b_j g' \neq 0$ since otherwise g' would every single one of its poles order one, which is impossible. Thus we have that $m = 1$. We can then in most cases ignore the extra element of F added onto v' since most cases will get a contradiction regardless if its presence, and if it can be solved without the additional terms of F then it can be solved by adding just the 0 element of F. Thus we get our formula that we have to solve, that is

$$\alpha = a't + at'$$

where $a = b_1$.

Thus for our case we now have a much simpler formula to deal with. In order to understand the next step, we must now refine our definition of an elementary function to be the elements of differential extension fields of the differential field of rational functions that satisfy that equation. Thus, for any given α you wish to check as an element of the set of elementary functions, break into into two functions, $f(x)$ and t where $f(x)$ is a rational function and t is an extension of the field of elementary functions, and show whether this formula can be solved.

To show this, take the more concrete example of $\alpha = f(x)e^{g(x)}$ where $f(x)$ and $g(x)$ are rational functions. Here, $f(x) = f(x)$ and the extension t is defined by $t = e^{g(x)}$. Thus, there must be an a in the field of rational functions such that

$$f(x)e^{g(x)} = a'e^{g(x)} + ag'(x)e^{g(x)},$$

$$f(x) = a' + ag'(x).$$

Since a is in the field of rational functions, a' is in the field of rational functions, and two rational functions added together make an rational function, we have a statement that can be true. Thus by Loiuville's Theorem, functions of the form $f(x)e^{g(x)}$ can be in a differential extension field of rational functions. This leads us to the conclusion that functions of the form $f(x)e^{g(x)}$ can be an elementary function.

Notice however that this definition of an elementary function then leads directly to its integrability. By showing that this equation has a solution we show that there is a $y \in F[t]$ such that $y' = \alpha$, or that α be written as the derivative of the function in the differential extension field. Thus there is a function in the extension field, namely $e^{g(x)}$ in this case, whose derivative is $f(x)e^{g(x)}$ and thus $f(x)e^{g(x)}$ has an anti-derivative or is integrable.

However, another case could be $\alpha = e^{x^2}$. This is the infamous Gaussian curve or the bell curve. Here, $f(x) = 1$ and $t = e^{x^2}$, making the equation as follows:

$$e^{x^2} = a'e^{x^2} - 2xae^{x^2},$$

$$1 = a' - 2xa.$$

This equation, however, leads to problems. Since a is a rational function, write $a = \frac{p}{q}$ where p and q are polynomials sharing no roots. From our rules of the derivative, we have that

$$a' = \frac{p'q - q'p}{q^2}.$$

Thus by algebra we get

$$q^2 = p'q - q'p - 2xpq,$$

$$q = p' - \frac{q'p}{q} - 2xp.$$

However, this gives us a contradiction. The function q is a polynomial, and thus $\frac{q'p}{q}$ must be a term in the polynomial. By the definition of p and q, these polynomials share no roots. However, q' has at least one less root than q by the definition of the derivative on a polynomial (it is one degree less). Thus there at least one root in q that is not in $q'p$, but this violates the definition of a polynomial. Thus q is not a polynomial and hence there is a contradiction. Thus by Loiuville's Theorem we can conclude that $\alpha = e^{-x^2}$ is not in a differential extension field of the rational functions that satisfies our required equation, and thus α is not an elementary function. Likewise, we have shown that there is no $y \in F[t]$ where $t = e^{-x^2}$ such that $y' = \alpha$, and thus the integral of α cannot be written in terms of elementary functions.

Now that we have the background knowledge I can show that an elliptic function is not integrable. Take $\alpha = \frac{1}{\sqrt{m}}$ where m is a polynomial of degree 3 or 4. Without loss of generality, take m to be a polynomial of degree 4. Thus, using our formula, we get

$$\frac{1}{\sqrt{m}} = \frac{a'}{\sqrt{m}} - \frac{3am'}{2\sqrt{m}},$$

$$1 = a' - \frac{3am'}{2}.$$

Since a is a rational function, take $a = \frac{p}{q}$ where p and q share no roots. Thus this can be rewritten as

$$1 = \frac{p'q - q'p}{q^2} - \frac{3pm'}{2q},$$

$$q^2 = p'q - q'p - \frac{3}{2}pqm'.$$

Here we get the same contradiction as in the Gaussian that q must be a polynomial but the term $\frac{q'p}{q}$ cannot be a term in a polynomial (notice that if $\alpha = \frac{f(x)}{\sqrt{m}}$ there would be no contradiction since then $f(x)q$ would have to have the term $\frac{q'p}{q}$ which is not contradictory). Thus we have shown that $\frac{1}{\sqrt{m}}$ is not in a differential extension field of the rational functions that satisfies our required equation, and thus $\alpha$is not an elementary function. Furthermore, there is no $y \in F[t]$ where $t = \frac{1}{\sqrt{m}}$ such that $y' = \frac{1}{\sqrt{m}}$, and thus the integral of $\frac{1}{\sqrt{m}}$ cannot be written in terms of elementary functions, which is the property we set out to prove.

4. Approximation Methods

At this point we seem to have hit a road block. We have calculated exactly what the integral must be to give us the circumference of an ellipse, but now we have just shown that the integral cannot be solved in terms of functions that are commonly used. However, we are not left hopeless in our adventure to find the circumference of an ellipse. There exist approximation methods to find the circumference of an ellipse. These can be separated into two categories, those being more on the theoretical side and relying on numerical approximations of the integral in question, and those more computationally practical.

The "more theoretical" approximation methods rely on numerical approximations of the integral that we found, namely

$$L = \int_{-1}^{1} \frac{1 - (c^2 - 1)t^2}{\sqrt{(1 - t^2)(1 - (c^2 - 1)t^2)}} dx.$$

These approximation techniques will give exact answers to as many decimal places as needed at the cost of high computational requirements. One of these methods is to use algebra to show that

$$\frac{1 - (c^2 - 1)t^2}{\sqrt{(1 - t^2)(1 - (c^2 - 1)t^2)}} = \frac{b^2\sqrt{(b^2 + t^2) + nt^2}}{(b^2 + t^2)^{\frac{3}{2}}}$$

where n is defined by $a^2 = (n+1)b^2$. Then by using the binomial theorem for non-integer powers it is shown that the function can be split in the following way:

$$\frac{1-(c^2-1)t^2}{\sqrt{(1-t^2)(1-(c^2-1)t^2)}} = \frac{b^2}{b^2+t^2} + \frac{1}{2}\frac{b^2nt^2}{(b^2+t^2)^2} + \frac{1\times 1}{2\times 4}\frac{b^4n^2t^4}{(b^2+t^2)^3} + \frac{1\times 1\times 3}{2\times 4\times 6}\frac{n^3t^6}{(b^2+t^2)^{\frac{5}{2}}} + \ldots.$$

This new function can then be split into separate integrals, each of which can be integrated for an answer in elementary functions. However, this is an infinite series so the computation can never be completed. However, an approximation can be given by using just the first n terms of the sequence. Likewise, this computation still requires integrals and it has many complicated terms to deal with and as such it is not very practical to use.

On the other hand, practical and easy approximation methods exist. The earliest of which can be attributed to Kepler. Kepler simply used the geometric mean, making the equation for the circumference C of an ellipse equal to

$$C = 2\pi\sqrt{ab}.$$

This simple approximation was within the error bounds of Kepler when he famously determined the orbits of celestial bodies to be ellipses. However, this approximation works well for cases where the ellipse is close to a circle (like in the case of celestial bodies) but goes towards 100% error quickly as the ellipses become flatter. Even the simpler approximation method using the arithmetic mean

$$C = 2\pi\frac{a+b}{2} = \pi(a+b)$$

which gives less error than the geometric mean. Interestingly, by taking a combination of the geometric and arithmetic means, we can arrive at an even better approximation, namely

$$C = 2\pi(3\frac{a+b}{2} - \sqrt{ab}).$$

Even better is to use

$$C = 2\pi(1.32\frac{a+b}{2} - .32\sqrt{ab})$$

The last useful approximation that can be done by hand is called the Linder approximation, given by

$$\pi(a+b)[1+\frac{h}{8}]^2,$$

where h is given by the following equation (this same h is used in further equations as well):

$$h = (\frac{a-b}{a+b})^2.$$

This method however is really just an extension of what was shown earlier, as it is the arithmetic mean combined with a correction to more accurately match the first three terms of the Taylor expansion. This approximation does extremely well for only using basic algebraic methods that can be done by hand and as such is one of the most practical approximation equations.

However, there exists an approximation that is hard to beat that only requires a calculator. This method is known as the Ramanujan Second Approximation, given by

$$C = \pi(a+b)[1+\frac{3h}{10+\sqrt{4-3h}}].$$

This approximation matches the Taylor expansions for the first 9 terms for ellipses of moderate eccentricity (meaning, not extremely flat).

Although Ramanujan's Second approximation normally given as the example approximation for a circumference, there is a better and easier approximation for the circumference of an ellipse. A group of approximations known as the Pade approximations give a better approximation and do not require taking the square root. An easy approximation from this group is given by

$$C = \pi(a+b)\frac{64+3h^2}{64-16h},$$

though there are even better of which can be used in computer programs, namely

$$C = \pi(a+b)\frac{135168-85760h-5568h^2+3867h^3}{135168-119552h+22208h^2-345h^3}.$$

Thus, the use of these approximation techniques can be described as follows. For simple "in your head" or on paper calculation, taking the arithmetic mean or the combination of the arithmetic mean and the geometric mean will do well for most cases. With slightly more thought put into it the Linder approximation will get the user a few more decimal places of accuracy. However, if one has a calculator or is writing a computer program, it is recommended that they use one of the Pade approximation methods. Although the Ramanujan approximation methods are famous, they are not as computationally efficient as the Pade approximations and do not receive as much accuracy, and as such should not be used in practical applications.

5. Conclusion

As can be seen by the depth of this paper alone, the circumference of an ellipse is a complex problem. By naively starting the calculation for the arc length of an ellipse we stumbled onto an elliptic integral. It was then proved that these integrals could not be solved in terms of elementary functions. However, since we could not go out with a defeat, we lastly turned to approximation methods. These ranged from approximations that can be done in one's head to the integral of an infinite series. These approximations given should be good enough for most people who need to find the circumference of an ellipse.