The Beauty of Geometrical Curves

Professor Sarah Hart

14th March 2022

Introduction

Today, I want to tell you about roulettes. In geometry, a roulette is a curve obtained by rolling one curve along another. This simple idea leads to some lovely curves. Among other things, we'll see the curve so appealing it was nicknamed the "Helen of Geometry", not just for its beauty but for the squabbles it caused among mathematicians, as well as the curve so versatile it has applications in everything from clockwork to nuclear reactors.

The Cycloid

The first kind of roulette I want to show you is the cycloid. It is the path traced out by a given point on the rim of a circle as you roll it along a straight line (a rather special case of a "curve rolling along a curve" roulette, because one of the "curves" is a straight line). Given that wheels have been rolling along roads for thousands of years, it's perhaps surprising that there's no convincing evidence of this curve being studied until the 16th century at the earliest. The first written mention of the question comes with Charles de Bouvelle, in 1501, when he was working on the problem of squaring the circle. But he didn't get very far with it – he thought that the shape was made from arcs of circles. Galileo studied the question in more depth. He knew that these weren't circle arcs, and gave the curve the name "cycloid".

So, what is this curve? Drawing it accurately, we can see the arches aren't circle arcs. The French mathematician Marin Mersenne thought that each arch might be half an ellipse. Lots of mathematicians in the century after Galileo named the cycloid tried to find out its properties. There are two obvious questions: to find the area under its arches (the "quadrature of the cycloid" problem), and to find the length of each arch of the cycloid. The answers to both these questions were found by more than one person, and they often fought bitterly over who had done it first. Both of these questions can be solved nowadays through techniques of calculus, but this is well before those tools were available. Galileo tried to work it out by doing a physical experiment. He got a circle cut out of metal, then rolled it along another piece of metal and marked out a cycloid. He then cut this out too, and simply weighed both the circle and the cycloid. The cycloid seemed to be about three times the weight of the circle, so he wondered if perhaps the area under it might be some irrational multiple, like π, times the area of the circle. But he couldn't resolve the matter. Remember, this was before even concepts like "the equation of a curve" had been invented, so the tools really weren't available yet to solve this problem.

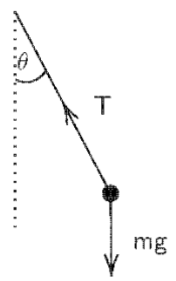
The next development was that Mersenne got interested. Mersenne was an important figure in European mathematics in the 17th century. Firstly, he was a good mathematician himself, making important contributions to many areas of mathematics (Mersenne's Laws governing how the frequency of strings on musical instruments changes with their length, tension, and linear density, are named after him, as are the famous Mersenne primes that he studied). But secondly, he corresponded with dozens of mathematicians across Europe, keeping them in touch with the latest developments. This was one way that news about cycloids got around. One person that Mersenne talked about cycloids with was Gilles de Roberval. Within a few years Roberval had managed to work out the area under the arches of the cycloid – it turns out to be exactly three times the area of the rolling circle that made it, which is a lovely and surprising outcome. But Roberval had a habit of being very secretive about his work. Sometimes he didn't reveal that he had solved a problem, and then when someone else announced their proof, he would angrily intervene saying that he had done it first. One reason for this odd behaviour is that Roberval's job depended on it. He held a post that was reappointed every three years, the best candidate being decided by their answers to a set of questions posed by…the current incumbent. This meant there was a strong incentive, if you had developed a particular technique, not to reveal your methods. You could then set questions that could only be solved with your method, banking on the fact that nobody else would be able to do them. And you'd be back in post again for another three years. I'm glad the Gresham Professorships aren't arranged like that!

We'll come back to Roberval later, because I want to show you one of the arguments he used. But first I'll tell you a little more of its surprising history.

Many famous mathematicians were drawn to the beguiling cycloid. Fermat, Descartes, Pascal, Newton, all worked on it, as well as Galileo, Roberval, and Mersenne. It even lured Blaise Pascal back to mathematics. He had decided to give it up, but one night was troubled by a terrible toothache. He distracted himself by thinking about cycloids, and the toothache miraculously went away. Taking this as a sign from God that studying mathematics is OK, he continued thinking about the cycloid for eight more days, proving lots of lovely results in the process, and he carried on studying mathematics after that, which is good news for mathematics (though I have to say cycloids have not replaced the local anaesthetic in dentistry). Some of these mathematicians were pretty rude about each other. When Fermat and Descartes both came up with ways to find tangents to the cycloid, Descartes said that Fermat's method was "ridiculous gibberish". Meanwhile, when an Italian mathematician called Torricelli published a proof that the area under the arches of a cycloid is three times the area of its corresponding circle, an irate Roberval went round telling everyone that Torricelli had stolen the idea from him – it's not clear how – and it was even claimed that Torricelli had literally died of shame when his plagiarism was discovered. It's a good story, but he did also happen to have contracted typhoid at the time, which may not have been unrelated. Still, these beautiful curves can drive a person a bit crazy. These, among many other stories, are how the cycloid came to be called the "Helen of Geometry".

But the story of the cycloid doesn't end there. Let me tell you something about clocks. Until 1656, sundials were the most accurate way to tell the time – if the sun happened to be shining, that is. The problem of telling time at night or in inclement weather was dealt with in a huge range of ways. There were water clocks, sand clocks or hourglasses, candle clocks, and so on. In the 15th and 16th centuries, mechanical clocks began to come into use. They relied on springs to regulate movement, but they were very inaccurate, losing up to 15 minutes a day. So what happened in 1656? Well, the scientist Christiaan Huygens built the world's first pendulum clock. Galileo had earlier studied pendulums, and suggested that the regularity of their motion could be useful in timekeeping, but had not put these ideas into practice.

Let's see why pendulums are useful. Imagine a very simple mathematical version of a pendulum. We'll assume the string has negligible mass compared to the bob on the end, that it doesn't stretch at all, and that there's no friction or air resistance at play. The only forces acting, then, are the tension T in the string, and gravity. Newton's law that force is mass times acceleration tells us that the force downwards, if the mass of the bob is m, is mg. The bob has to stay the same distance from the fixed top of the pendulum, so the component of the gravitational force pulling it in the direction of the line of the string is exactly cancelled out by the tension. This leaves the force along the circular arc, which a bit of trigonometry tells us is $mg\sin\theta$, where θ is the angle between the string and the vertical at the point of release.



Now, for small angles, $\sin\theta$ is approximately equal to θ. If we approximate the force by $mg\theta$, then we can use this to calculate the period (the total time for the pendulum to swing all the way to one side and return to the other), and it comes out to be

$$2\pi\sqrt{\frac{L}{g}}.$$

This is independent of both the initial angle of the pendulum and on the mass of the bob. And that's why pendulums are useful in timekeeping: once you set a pendulum swinging, even though the system gradually loses energy and so the amplitude (the height of the swing) gets lower over time, the period stays constant. So you can use this constant period to measure time. This one development meant clocks were now accurate to 15 seconds a day, not 15 minutes, a vast improvement. However, this neat formula is only an approximation that's accurate when the angles involved are very small. Huygens wanted to improve on clock design, and reasoned that what would be really helpful is if he could engineer things so that the pendulum could be made to move along a path with the following property: whatever point along the path the bob is released, it reaches the bottom in the same time. This is known as the "tautochrone" problem (from the Greek for "same" (tauto) and "time" (chronos). Huygens managed to solve it, and it turns out that the required curve is a cycloid (inverted of course). How do you make a pendulum move along a cycloidal path, though? Huygens managed to resolve that as well, and here is our next example of a roulette: an *involute*. This is the special case of a roulette where the curve that's doing the rolling is a straight line – a tangent to a given curve. We can think of it as the path of the end of a piece of taut string as it is unrolled from a curve. Involutes can look very different from the original curves. The involute of a circle looks like a spiral, for example, but the involute of a cycloid turns out to be another cycloid. This means that to create a cycloid path for your pendulum, you just need to suspend it at the cusp of the inverted cycloid (and make the pendulum half the length of a cycloid arch). Unfortunately, it turned out that cycloid-based pendulum clocks had more friction than normal ones, so weren't any more accurate in practice. But it's still a great piece of mathematics.

While we are talking about curves relating to speed, I can't resist mentioning a problem posed in 1696 by Johann Bernouilli, who asked not about the curve where all particles will reach the bottom in the same time, but instead about a curve of quickest descent. If I have two fixed points, what shape of curved wire should I link them with, such that a particle falling under gravity along the wire will reach the bottom in the shortest time? This was called the brachistochrone problem (from the Greek brachistos, for shortest). Johann and his brother Jacob both solved it, as did Leibniz, and so too did Newton – he published his solution anonymously but upon seeing it, Bernouilli is supposed to have said "tanquam ex ungue leonem" – we know the lion by his claw. Guess what curve solves the brachistochrone problem? That's right, it's the cycloid.

Before we think about other kinds of roulette curve, I want to show you Roberval's ingenious argument for finding the area under the cycloid (true to form, his methods weren't published until 1693, years after his death). The argument is perhaps not quite as mathematically rigorous as we might desire nowadays, but it can be made rigorous with calculus. If you have studied that subject, you might like to try it.

What Roberval did was to consider something he called the companion curve, and use it to work out the area under half the cycloid arch, and then double it. First, picture a semicircle standing upright at the starting point of the cycloid. (The semicircle should have the same radius as the circle that generates the cycloid.) This semicircle is made up of a lot (infinitely many, in fact) of horizontal line segments. What we do is, at each point of the half-cycloid, draw a horizontal line precisely the same length as the width of the semicircle at that point. This marks out a new curve – the companion curve. Now we use an argument that is extremely plausible, and is in fact correct, but proving it rigorously is more of a challenge, and wouldn't be possible for another few decades. The idea is that if you have two shapes that you can build up by parallel lines, and at the same level the length of the lines are equal to each other, then those shapes have the same area. This is sometimes called Cavalieri's principle – more often applied to three dimensional shapes. In that case, it would say that if corresponding slices through two shapes have equal areas, then the shapes have equal volumes. One example of this would be the formula for the area of a triangle. We know that if we chop a rectangle in half, we get two equal right-angled triangles. So the area of a right-angled triangle is half the base times the height. We can make any triangle with a given base and height by starting with a right-angled triangle of that base and height, and then just sliding the horizontal slices along until we have the triangle we want. Because we have exactly the same slices involved, the area must stay the same, goes the argument. And indeed this is the case. The area of any triangle is half its base times its height, whether it's right-angled or not.

Anyway, the construction of the companion curve guarantees that the area between it and the cycloid is exactly equal to the area of the semicircle, which, if the generating circle has radius r, is $\frac{1}{2}\pi r^2$. The next step is to think about the rectangle that encloses the half-arch of the cycloid. At the top of the cycloid, the rolling circle has gone half a revolution. So the distance rolled along the ground is half a circumference, or πr. Meanwhile the height of the rectangle is a diameter, which is $2r$. The great thing is that this picture is symmetrical – you could rotate the whole thing around its centre and get the same image. As a consequence, the companion curve actually divides the rectangle into two parts of equal area. (If you work out the equation, you discover it is in fact a sine curve, suitably scaled and translated – but Roberval probably didn't know that.)

The total area of the rectangle is $\pi r \times 2r = 2\pi r^2$. So half of that is πr^2. Now we can add that half-rectangle to the area between the cycloid and the companion curve to show that the area under half the cycloid arch is $\frac{1}{2}\pi r^2 + \pi r^2 = \frac{3}{2}\pi r^2$, meaning that the area under the whole arch is precisely $3\pi r^2$, which is exactly three times the area of the generating circle. Just as surprising is that the length of the cycloid arch turns out to be a whole number multiple of the radius – exactly $8r$, with not a π in sight. This fact was first proved by a former Gresham Professor of Astronomy, Sir Christopher Wren. (Wren died in 1723, so watch out in 2023 for several Gresham events marking this 300-year anniversary.)

Cycloids have caught the imagination of not only mathematicians, but writers. I mentioned in one of my lectures last year about mathematics and literature, that cycloids are discussed in Moby-Dick. But they are also mentioned in two great 18th century novels: Gulliver's Travels and Tristram Shandy. For those wishing someone would write a book about the many links between mathematics and literature, fear not: my book "Once Upon a Prime" will be published in 2023.

Epicycloids

So far, we've looked at a circle rolling along a straight line. What happens if a circle rolls along a more exciting curve? The curve made by a circle rolling along the outside of a circle is called an epicycloid; if it rolls round the inside of the circle it's called a hypocycloid. There are variants of all these too – the popular Spirograph Toy almost draws hypocycloids. However, the point we follow is not quite on the rim of the rolling circle, it's a little way inside. So, technically, these curves are called hypotrochoids and epitrochoids (a trochoid being what you get if you roll a circle along a straight line but the point you follow is not on the circumference, but somewhere inside or beyond the edge of the circle). We won't worry about those today. I'll focus on epicycloids, and in particular, ones where the radius of the fixed circle is a whole number multiple of the radius of the rolling circle. The first two examples are the cardioid, when the rolling circle has the same radius as the stationary one, and the nephroid, where the stationary circle has twice the radius of the rolling one.

The name cardioid (meaning heart-shaped) was first used by de Castillon in 1741, in a paper in the Philosophical Transactions of the Royal Society, though the curve had been studied for a long time before that. If the radius of the circles is r, then the length of the cardioid is 16r, and the area between it and the stationary circle is five times the area of the rolling circle that generates it. This is nice, and rather reminiscent of the cycloid. Suppose our circle of radius r is rolling round the outside of a circle of radius kr. Then we'll get k "arches" around the bigger circle. It can be shown that the total length is $8(k+1)r$, and the total area enclosed in the curve is $\pi r^2(k^2+3k+2)$. You can check this fits with the cardioid calculation. For values of k higher than 1, we'll get k arches, each of which has length $8\left(1+\frac{1}{k}\right)r$. Meanwhile, the area under enclosed between the arches and the circle (the "petals", if you like, of this flower) will be the total area, minus the area of the big circle, which is $\pi k^2 r^2$. This comes out at $\pi r^2(3k+2)$. So the area under each arch is $\frac{1}{k}$ of this. That is, $\pi r^2\left(3+\frac{2}{k}\right)$. Think what happens as k gets larger and larger. In the limiting case, we get a circle rolling along a straight line. As k tends to infinity, $\frac{1}{k}$ tends to zero. So we would get an arch length of $8r$ and area $3\pi r^2$. And sure enough, these are indeed the correct values for the cycloid.

There's a really nice way to draw a cardioid (or at least a very good approximation of one). Get a circle and mark off points equally spaced around the circumference. Let's try with 12 to begin with, like a clock face. Just like with a clock, we count round the hours and when we get to 12 the next hour is 1 again, but in the 24-hour clock it's 13, then 14, 15 etc. So each point does double duty, 4 = 16, 5 = 17 and so on. Now we follow this process – join each point to its double. So join 1 to 2, 2 to 4, 3 to 6 and so on, 6 to 12, 7 to 14 (which is the same as 2 because 14 hours is 2 o'clock), 8 to 16 and so on. The last one would be 12 to 24, so there's actually no line there, because it's the same point. This gives a series of lines, and we are interested in the shape they form. With 12 points it's not particularly good, but if we increase the number of points we get an increasingly curve-like shape. It's not obvious, but it turns out (and again I won't prove this as it's a bit technical) that in the limit, we precisely get our friend the cardioid.

You might also sometimes see things that look a bit like cardioids in your kitchen, in the light patterns created in a cup of tea in some specific kinds of lighting. In general, a curve whose boundary is made by light rays reflected off a curved surface is called a caustic. If you have a light source at a point inside a circular mirror, rays from that source will bounce off the inner surface. The curve to which each of those reflected rays are tangent will be visible as a concentrated region of light. If we look at the geometry of the situation, imagine we have our ray coming out from the point 0, and we have divided up the circumference again into equally spaced points. What happens to the ray that hits point p? It bounces off the circle and heads towards some point q. It's a law of optics that "angle of incidence equals angle of reflection". So we get symmetry here and the distance round the circle travelled between 0 and p equals the distance round the circle between p and q. So $q = 2p$. This means that the curve we end up with is the cardioid again. Now, in our cup of tea example, the light source isn't on the rim of the cup, but some way away. If it were very far away, we could assume that the light rays hitting the rim of the cup are parallel. In that situation, it can be shown that the caustic is actually not a cardioid but a nephroid. Since a strong overhead light is somewhere between these two extremes, the curve we get is usually going to be somewhere between a cardioid and a nephroid.

Are there any applications of cardioids? Yes, if you are a sound engineer you'll be familiar with cardioid microphones. These microphones are sensitive to sound only in a cardioid shape around the microphone – they pick up sound at the sides and in front, but not behind the microphone. It's precisely this arrangement that you want when recording something like live music, where you want to capture, say, the singer's voice, but not so much the sound of the audience.

One final place you might have seen a cardioid – the Mandelbrot set. If you know about complex numbers you'll know that squaring a complex number squares its distance from the origin and doubles the angle. The Mandelbrot set is generated by repeatedly squaring things and adding a constant, and seeing if the sequence zooms off to infinity or not, so it's at least plausible that there's a link between the doubling process I described, and areas of stability in the Mandelbrot set, and indeed that turns out to be the case.

Involutes

I already mentioned the idea of involutes. Any curve has an involute, but I want to focus in the rest of our time just on involutes of circles, and that's what I'll mean from now on when I say involute. The point at the end of a straight line segment rolling along a circle will sweep out a curve, called the involute. The straight line will always be tangent to the circle, and another way of looking at this is that the involute is the curve produced by the end of a taut thread being unravelled from a circle (or maybe cotton reel). Involutes look a bit like spirals, but they are something different.

I want to tell you about two uses of involutes – the first is something you'll almost certainly have somewhere in your house, and that's gears, or cogwheels, the sort you find in clockwork toys and other machinery. Gears transmit energy when their teeth mesh together. For maximum efficiency you want the teeth to be in constant contact and, if the input gear turns at a constant speed, you want the output gear to do the same (it might be a different constant speed if the gear wheels are different sizes). If the teeth of the gears have an involute profile, then where they meet, the line of force is going to be a line tangent to both circles, and the energy will be transmitted very smoothly. While there are other curves that do this (including cycloids), the involute design has several advantages, the main one being that as long as you keep the same number of teeth per inch (the so-called diametral pitch), you can use identical teeth on any size gear, which is a huge advantage in manufacturing. On a bigger circle with correspondingly more teeth, the part of the involute that's used is smaller but is the same shape as the larger part of the involute required on the smaller circle. The involute profile also does better than the cycloid if there are slight inaccuracies in the manufacture, particularly if the axle of the gear is slightly off-centre. Cycloid gears are still used in some mechanical clocks, though – they are stronger, so can have fewer teeth, and this allows a gear chain that has large gear reductions to fit into a small space. (In other words, the ratio of speeds between adjacent gears can be higher if you use cycloidal gears.)

To finish, I want to tell you about one other very important application of involutes: in nuclear reactors. At the High Flux Isotope Reactor at Oak Ridge National Laboratory, in Tennessee, their job is not to create energy but to create high mass elements, for use mainly in scientific research. They do this by using a nuclear reaction to create neutrons with which they bombard lighter elements to create heavier ones. The nuclear reactor cores are cylinders into which are put the fuel – Uranium Oxide (U_3O_8). The fuel is sandwiched between aluminium plates and then rolled out into long strips, which are then shaped into a particular curved shape to fit into the cylindrical core. Lots of heat is produced by 10^{15} (a million billion) neutrons bouncing around every square centimetre, and it's vital to be able to draw this heat away. So what's then needed is for these fuel strips to be curved in such a way that they are equally separated, in order that the water coolant can pass between them with no hotspots or places where the strips are too close together. And here's where involutes come in. They have an extremely useful property. If you draw a series of involutes starting at equally spaced points on the circumference of a circle, then the distances between the involutes remains constant. The involute is the only curve for which this is true – it wouldn't be the case for arcs of circles, for instance. That's why it's used in the nuclear reactor core.

We've really only scratched the surface of roulette curves today – I'll have to tell you another time what happens if you roll an ellipse along a sine curve, but I hope you've enjoyed the talk and next time you have a coffee break, look out for those cardioids.

References and Further Reading

- The Mactutor History of Mathematics site has a list of famous curves, including the ones we have discussed today, with interactive controls that allow you to experiment with the effect of changing the parameters. https://mathshistory.st-andrews.ac.uk/Curves/
- For more on cycloids in Moby-Dick you can read my article *Ahab's Arithmetic: The Mathematics of Moby-Dick*, at https://scholarship.claremont.edu/jhm/vol11/iss1/3/; I also discussed it briefly in my Gresham lecture Mathematical Journeys in Fictional Worlds https://www.gresham.ac.uk/lectures-and-events/maths-worlds
- You can watch the full video about the High Flux Isotope Nuclear Reactor on YouTube at https://www.youtube.com/watch?v=P99C051arMo, one of the excellent films made by Periodic Videos. Their website is http://www.periodicvideos.com/
- If you want to see the derivation of the nephroid/cycloid in the teacup, there's a very good Chalkdust magazine article which goes through some of the mathematics at https://chalkdustmagazine.com/features/cardioids-coffee-cups/

Image Credits

Unless otherwise specified, the technical diagrams are drawn by me and are ©Sarah Hart, and the photos are either taken by me or are public domain images.

- The sculpture of Pascal studying the cycloid is by Augustin Pajou (1730 – 1809). The photo is Public domain, via Wikimedia Commons.
- The picture of Huygens' pendulum clock design is public domain, https://commons.wikimedia.org/wiki/File:Huygens_first_pendulum_clock_-_front_view.png
- Cardioid gif: https://commons.wikimedia.org/wiki/File:Cardiod_animation.gif by AtomicShoelace, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons.
- Nephroid gif: https://commons.wikimedia.org/wiki/File:Nephroid.gif by Zorgit, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons.
- Involute gif: https://commons.wikimedia.org/wiki/File:Involute_of_circle.gif by 09glasgow09, CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons.
- Involute gears gif: https://commons.wikimedia.org/wiki/File:Involute_wheel.gif by Claudio Rocchini, CC BY-SA 3.0 <http://creativecommons.org/licenses/by-sa/3.0/>, via Wikimedia Commons
- Mandelbrot set: https://commons.wikimedia.org/wiki/File:Mandel_zoom_00_mandelbrot_set.PNG Created by Wolfgang Beyer with the program Ultra Fractal 3., CC BY-SA 3.0 <http://creativecommons.org/licenses/by-sa/3.0/>, via Wikimedia Commons
- Pictures of the involutes in the nuclear reactor are stills from the High Flux Isotope Reactor video linked in the references.