Chapter Eleven: Fourier Series and Spectrograms

We have skirted around three obvious questions:

- Is every periodic function $y(x)$ satisfying $y(x+p) = y(x)$ for all x and some fixed period p given by a so-called *Fourier Series,* a sum of sines and cosines:
$y(x) = B_0 + A_1\sin(2\pi x/p) + B_1\cos(2\pi x/p) + A_2\sin(4\pi x/p) + B_2\cos(4\pi x/p) + \cdots$
or (using the rule that $A\sin(x) + B\cos(x)$ can be rewritten as $C\sin(x+D)$), by a sum of sines with phase shifts:
$y(x) = C_0 + C_1\sin(2\pi x/p + D_1) + C_2\sin(4\pi x/p + D_2) + \cdots$
- Is every function $y(x)$ with $y(0) = y(L) = 0$ a sum of sinusoids like this:
$y(x) = C_1\sin(\pi x/L) + C_2\sin(2\pi x/L) + C_3\sin(3\pi x/L) + \cdots$
- Is every solution of the vibrating string equation of the form written down by Euler:
$$y(x,t) = C_1\sin(\pi x/L)\cdot\cos(2\pi ft + D_1) + C_2\sin(2\pi x/L)\cdot\cos(4\pi ft + D_2) + \\ +C_3\sin(3\pi x/L)\cdot\cos(6\pi ft + D_3) + \cdots$$

The first arises in our description of a singing voice in Chapter 9, the second and third in our description of a vibrating string in Chapter 10. The answer to all these questions is YES to both IF you're a little careful about how jumpy and erratic the functions y are allowed to be and about how you go about adding up the infinity of terms of higher and higher frequency. (These 'IF's are the typical questions that can occupy *months* – years even – of study in higher math courses but that are usually irrelevant for applications and computations.) Their truth was one of the most important mathematical discoveries of the 18th and 19th century. The 'take home message' is that writing a function as a sum of sines and cosines is as important as writing a function as a polynomial: both are universal tools that display basic parameters in the functions makeup.

Relationship between the three bullets:
(a) The second bullet follows from the first because we can get sine series as a special case of Fourier series as follows. If $y(x)$ is defined between 0 and L and is zero at 0 and L, then first extend y to a function between $-L$ and 0 to have values $-y(-x)$ and then make y into a function defined for all values of x by making it periodic with period $2L$. Then the Fourier expansion of this periodic extension y turns out only to have sine's in it because y is 'odd', $y(-x)=-y(x)$ for all x, and so we get the sine series. [FIGURE FOR THIS]

(b) If we know that Fourier and sine series always exist, then we also know that these series give all the solutions of the vibrating string equation – the third bullet. This is because we saw from the difference equation approach to the PDE that if we know where y starts *and* its rate of change, i.e. $y(x,0)$ and $y_t(x,0)$, then the vibrating string equation has only one solution continuing these. So we just need to write out the two sine series:

$$y(x,0) = B_1\sin(\pi x/L) + B_2\sin(4\pi x/L) + \cdots$$
$$y_t(x,0) = 2\pi\left(A_1\sin(\pi x/L) + 2A_2\sin(4\pi x/L) + \cdots\right)$$

and contrive the C's and D's so that

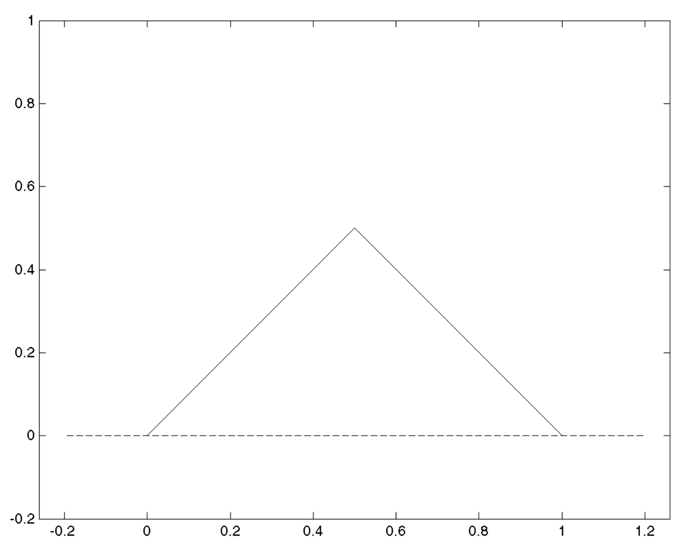
$$A_k \sin(2\pi kft) + B_k \cos(2\pi kft) = C_k \cos(2\pi kft + D_k)$$

With this choice of C's and D's, the function in bullet 1 is easily seen to have the right values of $y(x,0)$ and $y_t(x,0)$.

The theorem that all, not too erratic, periodic functions $y(x)$ have Fourier expansions has one of the most curious histories one could imagine. Euler, who found and published such expansions for all the basic functions and who loved manipulations of this kind, resisted strongly the idea that *every* function could be so expanded. The great mathematicians of the 18th century were polarized: on one side, the mathematicians who leaned to 'pure' mathematics, Euler, D'Alembert and Lagrange, insisted that functions given by Fourier series were special but, on the other side, the truly applied mathematicians and mathematical astronomers, Bernoulli and Clairaut, believed it was true. Fourier, after whom these series are named, was a 19th century polymath, who split his career between teaching and serving as prefect in various Departments of France, and applied these series to understand the spread of heat in the earth. Although not the inventor of 'Fourier series', he claimed strongly that they did represent all periodic functions and stimulated the rigorous theory of these series whose twists have continued to this day[1].

The best way to understand what was at issue is to look at an example. Euler had the idea that a sum of trig functions could be made to add up to any $y(x)$ that could be given by a single closed formula, such as a polynomial. But he also introduced what was then a radically new idea of what a function was: it could be given by one formula for some values of their argument and another function for others (or it might even be a *freehand curve*, drawn by hand). He called these *discontinuous* because the formula for them changed abruptly even though their value need not jump or anything. A typical example of a function that we call continuous, but Euler called discontinuous, is the tent function:

$$y(x) = |x|, \quad \text{if } 0 \le x \le L/2$$
$$y(x) = |L - x|, \ \text{if } L/2 \le x \le L$$



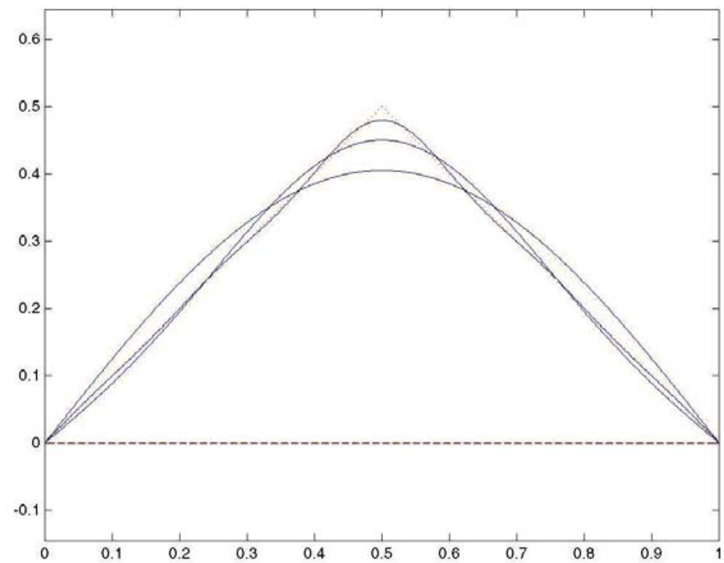
which is shown on the right for L = 1. Euler felt he could expand the first part and the second half into trig functions but not the combination of the two. *But he was wrong.* The answer is this (with L = 1 for simplicity):

[1] The key issues are how erratic a function can be to be expanded in such a series and in what sense this infinite series converges for increasingly wild functions. The upshot is that 'any' periodic function $y(x)$ has a unique Fourier expansion but at a very small number of points where y is erratic, e.g. it jumps, the series may not converge.

$$y(x) = C\left(\sin(\pi x) - \frac{1}{9}\sin(3\pi x) + \frac{1}{25}\sin(5\pi x) - \cdots\right), \text{ where } C = 4/\pi^2$$

This is not irrelevant to music: the above shape is a plausible way to pluck a string and the expansion shows that it produces all *odd* harmonics, that is the note itself, then its third harmonic, then its fifth, etc. Let's graph this expression truncating the infinite sum of sines – thus check it numerically, and then we'll see how to find such facts.

Here is a plot of three approximations to the tent curve, (a) with a single sine, (b) with two sines $\sin(\pi x) - \sin(3\pi x)/9$ and (c) with five sine terms of frequencies 1,3,5,7 and 9. It gets close to the tent everywhere except at the peak, but eventually, the trig sum will get close at the peak too (though each finite sum will be round at the peak of the tent if you look closely).



Where did we get these strange coefficients 1/9, 1/25 etc and outside everything $4/\pi^2$? There's a simple trick that can be used based on a trig identity. This is the formula:

$$\cos(2\pi at + D)\cos(2\pi bt + E) = \frac{1}{2}\left(\cos\left(2\pi(a+b)t + (D+E)\right) + \cos\left(2\pi(a-b)t + (D-E)\right)\right)$$

This means that if you *multiply* two sinusoidal functions of frequencies *a* and *b*, then the result is a sum of sinusoidal functions of frequencies *a*+*b* and *a*-*b*. We use cosines here instead of sines only to make the formula a bit simpler – just add $\pi/2$ to the arguments and the cosines become sines. This formula is just the result of rearranging the addition formula for cosines

Add: $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$

and: $\cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$

getting: $\cos(x+y) + \cos(x-y) = 2\cos(x)\cos(y)$

This formula is one of the key reasons why periodic sines and cosines are so versatile.

Now, going back to our Fourier expansion of the tent function, if we want the coefficient of the first term $\sin(\pi x)$, we multiply both sides of the equation by $\sin(\pi x)$, so that the equation reads:

$$\sin(\pi x)\cdot y(x) = C\left(\sin^2(\pi x) - \frac{1}{9}\sin(\pi x)\cdot\sin(3\pi x) + \frac{1}{25}\sin(\pi x)\cdot\sin(5\pi x) - \cdots\right)$$

The point is that $\sin^2(\pi x)$ is never negative while the other terms on the right are again sinusoids and have equal positive and negative swings. In fact, the average value of $\sin^2(\pi x)$ is ½ and the average value of the others is 0. So just integrate both sides of the equation between 0 and 1 and we'll get an equation that can be solved for *C*!

Let's check the details here:

a) By the identity above with $a=b=1/2$ and $D=E=\pi/2$, we get $\sin^2(\pi x) = (1-\cos(2\pi x))/2$. So $\sin^2(\pi x)$ has average ½. The same identity shows that the other terms are ½(cos((*n*-1)π*x*)–cos((*n*+1)π*x*)) and now both cosine terms have average 0.

b) We conclude $\int_0^1 y(x)\sin(\pi x)dx = C\int_0^1 \sin^2(\pi x)dx = C/2$.

c) Now the integral on the left is easy to compute *if* you remember a trick from calculus – integration by parts. It works out like this (but just accept this if you want):

$$\int_0^1 y(x)\sin(\pi x)dx = 2\int_0^{1/2} x\sin(\pi x)dx$$
$$= \frac{2}{\pi}\left(-x\cos(\pi x)\Big|_0^{1/2} + \int_0^{1/2}\cos(\pi x)dx\right) = \frac{2}{\pi}\left(0 + \frac{\sin(\pi x)}{\pi}\Bigg|_0^{1/2}\right) = \frac{2}{\pi^2}$$

d) Thus $C = 4/\pi^2$!

e) This trick works to get all the coefficients. Using (a) and (b), we first show that the coefficient of sin(*n*π*x*) is equal to $2\int y(x)\sin(n\pi x)dx$. This is easily worked out by integration by parts as in (d) and that it equals 0 if *n* is even, and equals $(2/n\pi)^2$ if *n* is odd.

In fact, we can expand any periodic function into a sum of sines and cosines, finding its coefficients by the above trick, known as the *orthogonality* of the trig functions, the fact that the average value of products of different trig functions is zero. We give another example in the problem below.

Let's formulate the general rule. Suppose *y*(*x*) is now any periodic function, with period *p*. That is, $y(x+p) = y(x)$ for all *x*. Then *y*(*x*) can be written uniquely as an infinite sum of trig terms:

$$\boxed{y(x) = B_0 + A_1\sin(2\pi\frac{x}{p}) + B_1\cos(2\pi\frac{x}{p}) + A_2\sin(4\pi\frac{x}{p}) + B_2\cos(4\pi\frac{x}{p}) + \cdots\cdots}$$

Moreover, the coefficients can be found by:

$$A_k = \frac{2}{p}\int_0^p \sin(2k\pi\frac{x}{p})y(x)dx,$$
$$B_0 = \frac{1}{p}\int_0^p y(x)dx$$
$$B_k = \frac{2}{p}\int_0^p \cos(2k\pi\frac{x}{p})y(x)dx, \text{ if } k > 0$$

This called the Fourier Series for $y(x)$. Of course, each sine/cosine pair can be rewritten:

$$A_k \sin(2k\pi\frac{x}{p}) + B_k \cos(2k\pi\frac{x}{p}) = C_k\left(\sin(2k\pi\frac{x}{p} + D_k)\right)$$

where $C_k = \sqrt{A_k^2 + B_k^2}$ is called the amplitude of the k^{th} harmonic and D_k its phase (C_k^2 is called its *power*).

There's a beautiful graphical way to display music using Fourier series that is the mathematical version of a musical score! It's called a *spectrogram.* All you need to do is break the sound up into 'windows' and expand each piece into a Fourier series as above and make a picture out of the amplitude of the various harmonics in each window. What you do exactly is take a 'window function', a smooth hump-like function (a) zero outside the window, (b) 1 on 90% of the window and (c) with shoulders near the beginning and end of the window. You multiply the sound by this window function and treat the product as though it were a periodic function, wrapping the beginning and end of the window together. Then it expands as a Fourier series by the formula above. [NEED FIGURE]

In the figure on the next page, the amplitude of the coefficients has been graphed by *colors* (another twist on Oresme's precepts!).

[EXPAND SECTION ON MODULATION] The formula above which showed that the product of two sinusoids of frequencies *a* and *b* was the sum of sinusoids with frequencies *a*+*b* and *a*-*b* has another immensely important application in communication. This is the idea of *modulation*, of combining many slow messages into one very fast message. The idea is to take, for instance, a human voice given by the air pressure function $y(t)$ which varies 'slowly', e.g. in the figure above, it is made up of vibrations in the range 60-6000 hertz, and multiply this function by a *carrier wave* $c(t)=\sin(2\pi ft)$ of much higher frequency, like 1,000,000 hertz (AM radios) or 100,000,000 hertz (FM radio, TV). Then – taking the old-fashioned AM case – the product $y(t).c(t)$ is a sum of sinusoids of frequencies between 994,000 and 1,006,000 hertz. [FIGURE HERE] If you want to broadcast many stations at once, each station uses carrier frequencies spaced at least 20,000 hertz apart. If station 1 sends voice *y* with carrier *c* at 1,000,000 hertz and station 2 sends voice *z* with carrier *d* at frequency 1,020,000 hertz then the total signal $s(t)$ is:

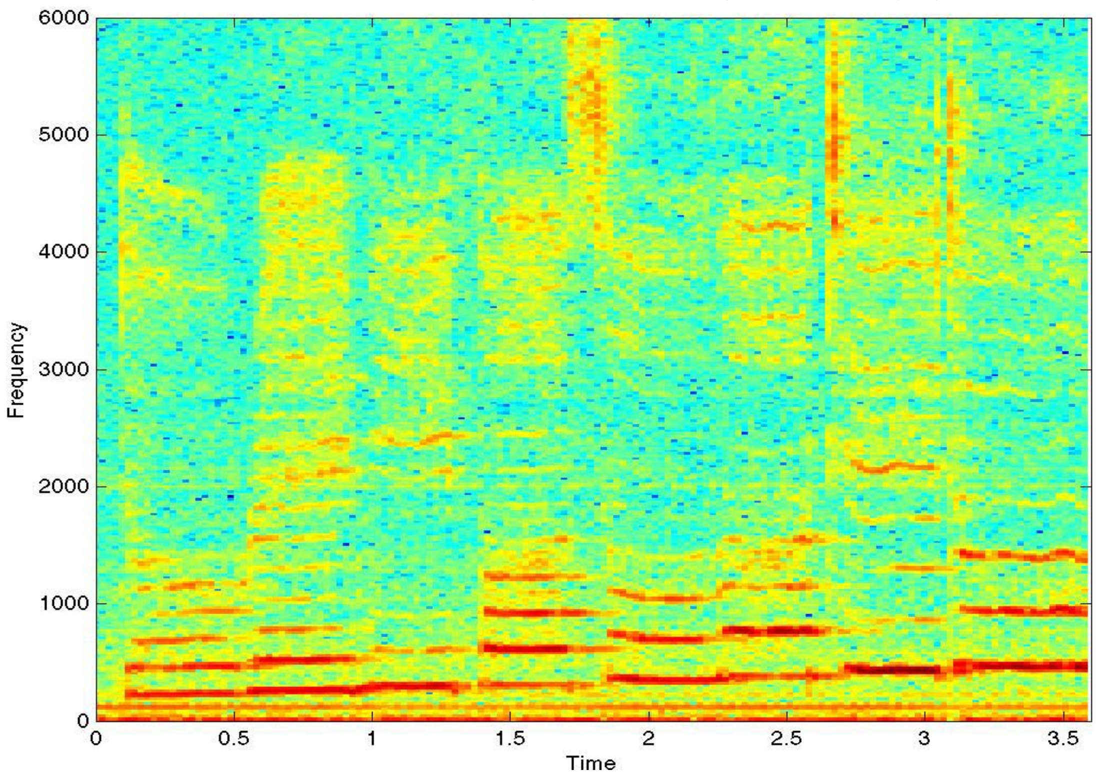
$$s(t) = y(t)c(t) + z(t)d(t)$$

The frequencies of the first term are in the range 994,000 to 1,006,000 and the second in the range 1,014,000 to 1,026,000. *These frequencies do not overlap*, so by the Fourier trick, we can pull them apart and recover the two voices y and z. By this trick, we multiply the signal $s(t)$ by sinusoids in the range 994,000-1,006,000 and average and we get back the Fourier coefficients of $y(t)$ hence we get back $y(t)$. It's actually easier than this: you just multiply s once by the carrier c, which, after all is just a single sine wave of frequency 1,000,000 and you get:

$$s(t)c(t) = y(t)c(t)^2 + z(t)c(t)d(t)$$

and by taking the right kind of average over windows of size 1/20,000-second (a so-called 'low-pass filter'), [MOLASSES] you can eliminate the second term and reduce $c(t)^2$ to a constant and out pops y again. All this is basically a fancy version of the method used to recover the coefficients of the tent function expansion by making one term positive while the others oscillate.

The idea of modulating multiple low frequency signals onto carriers at suitably spaced high frequencies and adding has been the source of all our ability to communicate efficiently until quite recently, when optical fibre, 'solitons' and digital signals have taken over (see Chapter ??).The spectrogram of the major scale sung by a female voice.



The 8 vertical strips show the 8 notes *do, re, mi, fa, sol, la, ti* and *do.* The dark horizontal lines in each strip represent strong harmonics present in each note: the lowest is the fundamental, then the second harmonic etc. *Ti* for example shows only the fundamental and the fifth harmonic and also has a marked trill. *La s*hows significant power in the 11th harmonic. Note the dark high frequency signal between *fa* and *sol*: this is called 'white noise' and what happens when you speak an 's'. It also occurs as a burst in the stop consonants 't' and 'd' of *ti* and the final *do*.

Chapter Twelve: The Square Root of -1 and Complex Numbers

We have been concerned with the line of development of mathematics that started with Galileo and Newton and centered on the issue of predicting the evolution of the physical world with the aid of differential equations. We could continue, but an new and completely unexpected element enters the story in the 18^{th} century, the gradual understanding that a completely new type of *number* was extremely useful for calculating the progression of events in the world. These new numbers are called "complex numbers" and are based on the introduction of the square root of minus one. Although we could skip this if we were trying to explain, in simplest terms, the mathematical models of waves with differential equations, the story of the square root of minus one is such an amazing one, such an unexpected twist, that it begs to be included. Of all the devices that mathematicians have found and used to model nature, this one seems – to me at least – as if God decided to throw us a curve ball, something unexpected, something that 'needn't have been so' but nonetheless really is true. My Aunt, studying maths at Girton College nearly a hundred years ago, called the square root of minus one a 'delightful fiction'. Modern scientists, however, take complex numbers for granted, FORTRAN makes them a predefined data type, and they are standard toolkit for any electronic engineer. Perhaps the most remarkable fact about complex numbers is that they are absolutely essential to quantum mechanics: in this theory, not only can the universe exist probabilistically in two states at once, but the uncertain composite state is constructed by adding the two simple states together with complex coefficients, introducing a complex 'phase', as we shall describe below.

Over millennia, many cultures have wanted to solve polynomial equations. Along with trying to make sense of the planets and the moon, solving quadratic equations appears as a strange mathematical obsession with so many cultures. Already in 1800 BC, we find Babylonian clay tablets posing problems like:

"I have multiplied length and breadth and the area is 10. The excess of length over breadth I have multiplied by itself and this result by 9. And this area is the area obtained by multiplying length by itself. What are the length and breadth?"

This translates in modern terms into:

$$
\begin{aligned}
&L = \text{length},\ B = \text{breadth} \\
&L \cdot B = 10 \\
&L > B \text{ and } 9(L-B)^2 = L^2
\end{aligned}
$$

which leads to a quadratic equation for L^2 and has solution $L = \sqrt{15}$. There is always an issue when you do this: what sort of numbers are allowed as solutions? Usually, this meant positive real numbers as both negative numbers and square roots of negative numbers were not legitimate. But both the Chinese and the Indians introduced negative numbers, for example to represent debts. The 12^{th} century Indian mathematician Bhaskara addressed the issue of square roots explicitly: he states that positive numbers have 2 square roots, the usual positive one and its negative – but that negative numbers have no square roots. He was well acquainted with the fact that the product of a negative and positive number is negative while the product of 2 negative numbers is positive. Incidentally, he also said when he found an equation with one positive and one negative solution: "*The (negative) value is in this case not to be taken, for it is inadequate; people do not approve of negative solutions*".

In the late Renaissance (c.1500-1550), it became a sporting competition for Italian mathematicians to challenge each other to solve various higher degree equations. Ferro, Tartaglia, Cardano and Ferrari were players. But Girolamo Cardano (1501-1576) spoiled the game by publishing their secrets in his famous book *Ars Magna* (The Great Art) in which the general procedure for solving third and fourth degree equations was explained. He was a boisterous figure, interested in everything, a gambler who was jailed for casting the horoscope of Jesus. But in his book, negative numbers were treated with suspicion and called *fictitious* solutions. Fourth powers were likewise a game – squares stood for areas of squares, cubes stood for volumes of cubes, but, in the absence of a fourth dimension, what should one make of fourth powers? Believe it or not, as recently as 1831, Augustus de Morgan, Professor of Mathematics at University College London could say:

"The imaginary expression $\sqrt{-a}$ and the negative expression –b have this resemblance, that either of them as occurring as the solution of a problem indicates some inconsistency or absurdity. As far as real meaning is concerned, both are equally imaginary, since 0-a is as inconceivable as $\sqrt{-a}$."

These sound like silly scruples to us now. But once you play games with formulas to solve equations, not only do you find it much easier to allow negative numbers and fourth powers, but you find it hard not to take square roots of all numbers. Cardano, in particular, found a very strange thing: his formula for the solutions of cubic equations, which usually worked fine, sometimes involved intermediate steps which were square roots of negative numbers, even when the final answer should be an honest positive real number! He ends his book saying "*So progresses arithmetic subtlety, the end of which, as is said, is as refined as it is useless*". His point of view was the same as my Aunt's.

So how should we look at 'numbers' like $\sqrt{-a}$ today. Firstly, we adopt modern notation and represent $\sqrt{-1}$ by the innocuous symbol *i*. (Some people, particularly electrical engineers, refer to the same quantity by *j*.) All you have to remember is that $i^2 = -1$. Then multiples of *i* like *i*/3 and $\sqrt{2} \cdot i$ are called *imaginary* numbers (in distinction to *real* numbers like -2.16 and +π) and they multiply in the usual way:

$$(ai) \cdot (bi) = (ab) i^2 = -ab$$

It is easy to see with this sort of calculation that all negative numbers now have square roots of the form *ai*. If we now *add* a real and an imaginary number we get a number like -3.1 + 2.6*i* which is called *complex*. We simply *define* a complex number to be a sum *a*+*bi* for any real numbers *a* and *b.* To show that a complex number should be thought of as a single entity, we often represent *x*+*iy* by a single symbol z. The (real) number *x* is called the real part of *z*, written Re(*z*), and *y* is called the imaginary part, written Im(*z*).

The miracle is that once we have allowed ourselves to take this bold step, we are done! Using the usual rules of arithmetic, we can mechanically add, subtract and multiply to our heart's content: no contradictions arise and no new kinds of `numbers' will ever appear.

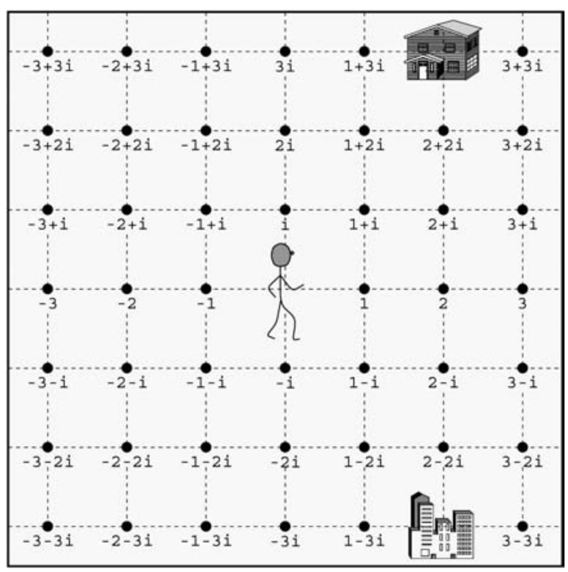
$$(a+ib)+(c+id)=(a+c)+i(b+d)$$
$$(a+ib)-(c+id)=(a-c)+i(b-d)$$
$$(a+ib)\cdot(c+id)=ac+ibc+iad+i^2bd=(ac-bd)+i(bc+ad)$$

Division is a bit harder, but there is a simple trick that makes it clear there is only one possible way it could work. Every complex number $a+ib$ has a sort of twin, called its 'conjugate' which is given by $a-ib$ – just change the sign of the imaginary part. This is because there is really no difference between i and $-i$. Then you divide like this:

$$\frac{a+ib}{c+id}=\frac{(a+ib)\cdot(c-id)}{(c+id)\cdot(c-id)}=\frac{ac-iad+ibc-i^2bd}{c^2-i^2d^2}=\frac{ac+bd+i(bc-ad)}{c^2+d^2}=\left(\frac{ac+bd}{c^2+d^2}\right)+i\left(\frac{bc-ad}{c^2+d^2}\right)$$

Even more strikingly, in the mid 18^{th} century, D'Alembert showed that really remarkable fact that no further fictitious or imaginary numbers need to be invented in order to solve all possible polynomial equations: every n^{th} degree equation has n roots if you allow them to be complex numbers. This was called the *fundamental theorem of algebra*, a reasonable name since it capped more than 3 millennia of solving polynomial equations. In spite of all this, there was an air of unreality about them. Felix Klein writes about the attitude of enlightenment mathematicians toward complex numbers as follows:

> "*Imaginary numbers made their own way into arithmetic calculations without the approval, and even against the desires of individual mathematicians, and obtained wider circulation only gradually and to the extent that they showed themselves useful".*



A key step in taming complex numbers was them see them geometrically. Negative numbers seem much less daunting when you just say they represent points to the left side of the origin on the x-axis (or numbers below the zero point on your thermometer). Representing all the points on a line by positive and negative numbers makes it seem inevitable to include them on an equal footing as legitimate numbers. This can be done with complex numbers by representing them as points in a plane. The idea goes back to John Wallis in his 1685 book *Treatise on Algebra*. He was concerned with making a picture to describe the roots of a quadratic equation even when the quadratic formula requires you to take the square root of a negative number. He says, take a line with an origin, the positive and negative reals being to its right and left. Then represent *a*+*ib* by going a distance *a* along this line and then a distance *b* along a *perpendicular line*. This is shown in the figure above where the man is standing at the origin, his house is at the point 2+3*i*, and his office is at the point 2-3*i*. This geometrical representation appears again, much more fully developed, in the work of a Norwegian surveyor, Caspar Wessel (1745-1818), who found that using complex numbers to represent his observations made many calculations much simpler. It finally became the accepted way of understanding complex numbers in the work of Carl Frederich Gauss (1777-1830), the greatest mathematician since Newton.

What Wessel realized is that, in this planar representation of complex numbers, multiplication becomes extremely simple: multiplying all the points in the plane by a fixed complex numbers *x*+*iy* has two effects: it stretches (or shrinks) everything by some factor *r* and rotates the plane around the origin through some angle θ. To see this, we merely need to change to polar coordinates:

$$\text{Let } x = r \cdot \cos(\theta), y = r \cdot \sin(\theta),$$
$$\text{then } z = x + iy = r \cdot (\cos(\theta) + i\sin(\theta))$$

Here $r = \sqrt{x^2 + y^2}$ is called the *absolute value* |*z*| of *z,* the distance of *z* from the origin and $\theta = \arctan(y/x)$ is called the *argument* arg(*z*), the angle which a line from *z* to the origin makes with the positive *x*-axis The special complex numbers cos(θ)+isin(θ) are the points on the unit circle (because $\cos^2 + \sin^2 = 1$). Now multiply 2 complex numbers expressed in polar coordinates:

$$r\left(\cos(\theta) + i\sin(\theta)\right) \times s\left(\cos(\phi) + i\sin(\phi)\right) =$$
$$rs\left((\cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi)) + i(\cos(\theta)\sin(\phi) + \sin(\theta)\cos(\phi))\right)$$

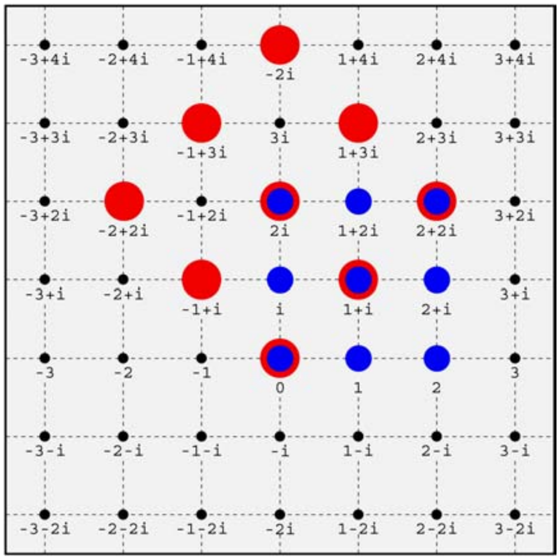
At first sight, this is a mess. But we use again the somewhat cumbersome addition formulas for sine and cosine:

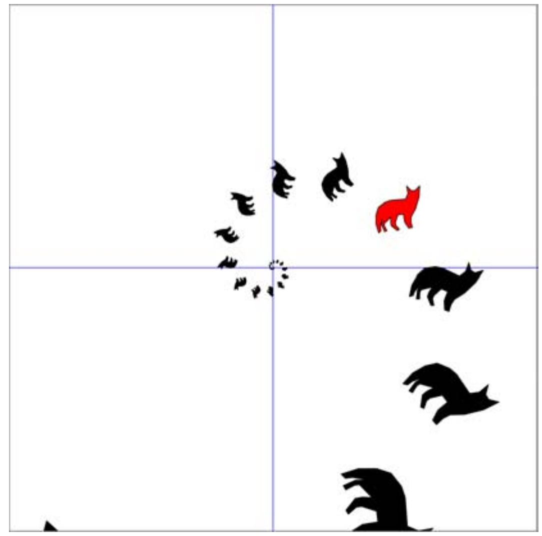
$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y),$$
$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

Then the above simplifies to:

$$r\left(\cos(\theta) + i\sin(\theta)\right) \times s\left(\cos(\phi) + i\sin(\phi)\right) =$$
$$rs\left((\cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi)) + i(\cos(\theta)\sin(\phi) + \sin(\theta)\cos(\phi))\right) =$$
$$rs\left(\cos(\theta + \phi) + i\sin(\theta + \phi)\right)$$

In other words, multiplying geometrically multiplies the distances *r* and *s* from the origin and adds the angles θ and ϕ the points make with the *x*-axis. The pair of figures below show the result of multiplication by 1+*i* (on the left, blue circles going to red circles) and by powers of $0.8\left(\frac{\sqrt{3}}{2} + \frac{i}{2}\right)$ and its inverse (on the right: the red fox being carried in the spiral by successive powers). Note that the absolute value of these are $\sqrt{2}$ and 0.8, so one expands, the other contracts. One rotates by 45 degrees, the other by 30 degrees. It is this ability to handle rotations as well as expansions and contractions that makes complex arithmetic so handy – to mathematicians as well as surveyors.





Euler, who loved formulas with all his heart, found a link between trigonometry and the square root of minus one and comes up with the most beautiful formula of all, the formula which established the importance of complex numbers once and for all.

To explain his idea, we have to go back to some a really simple differential equation which we skipped over when we were developing the consequences of Newton's laws. In Newton's language, suppose you have a fluent whose rate of increase is proportional to its value. This sort of growth is familiar to us in interest bearing bank accounts: the rate of growth is constant so the absolute growth is proportional to money in the account. And if the rate is negative, so we get shrinkage instead of growth, it well known from radioactive decay: in any period of time, a fixed percentage of the Uranium (for example) will decay. This leads to a simple formula connecting the fluent to its fluxion, i.e. the function to its derivative:

$$\boxed{\dot{x}(t) = ax(t)}$$

This equation is not in Newton's form: it doesn't express the *acceleration* of *x* in terms of positions and velocities but is simpler and expresses the velocity or rate of change of *x* directly in terms of the value of *x*. In finite difference, computer-ready form we can express it:

$$\frac{x\big((k+1)\Delta t\big) - x(k\Delta t)}{\Delta t} = ax(k\Delta t), \text{ which works out to be:}$$

$$x\big((k+1)\Delta t\big) = x(k\Delta t) + \Delta t \cdot \big(ax(k\Delta t)\big) = \big(1 + a\Delta t\big) \cdot x(k\Delta t)$$

Thus the solution is simply $x(k\Delta t) = (1 + a\Delta t)^k x(0)$. If a is positive, x is growing exponentially and if a is negative, it is shrinking exponentially. Suppose t is measured in years and a=0.1. Then this shows the interest on a bank account yielding 10% per year. Δt is simply the interval at which interest is posted. Thus if interest is paid monthly, then Δt = 1/12 and after k months you have $(1 + 0.1/12)^k x(0)$ dollars. The differential equation corresponds to interest being *compounded continuously*. Suppose the interest rate is

100%, i.e. a = 1, and suppose you start with \$1.00 in the bank. If the interest is credited once a year, you have \$2.00 after 1 year. If it is credited quarterly, you have $(1+1/4)^4 \approx \$2.36$ at the end of the year. If interest is compounded daily, you have $(1+1/365)^{365} \approx \$2.71$ at the end of the year. If, however, it is compounded after every infinitesimal passage of time, the money you have turns out to be \$2.71828.... which is one of those numbers, like π, which don't have a simple expression, so we give them a name – in this case *e*, after Leonard Euler. The solution to the boxed equation, which stands for continuous exponential growth, growth proportional to how much you have ('*To him who has, more shall be given*', Matt 20:21) is just:

$$x(t) = e^{at}x(0)$$

where e = 2.71828..., which you should remember from 'intro calc'. If $a<0$, then a common example is that of radioactive decay, the amount for example of Uranium present after some lapse of years. By the way, the computer gives us the approximation $x(t) \approx (1+at/n)^n x(0)$ if we use $\Delta t = t/n$, so we see that

$$e^{at} \approx (1+at/n)^n, \text{ with equality in the limit for } n \to \infty$$

Note that changing the constant a in the equation has the effect of making time go faster or slower. If a is increased, then the rate of change of x is faster and the whole future unravels at a faster pace; while if a is decreased, the future comes to pass at a slower rate. We will see this again and again, when the constants in these differential equations are considered.

Now, following Euler, we are going to put together exponential growth with the complex numbers. We saw above that:

$$(*) \qquad (\cos(\theta)+i\sin(\theta))\cdot(\cos(\phi)+i\sin(\phi)) = (\cos(\theta+\phi)+i\sin(\theta+\phi))$$

This has as a Corollary a formula that DeMoivre had found a few years before Euler got involved. If $\theta=\phi$, then we get

$$(\cos(\theta)+i\sin(\theta))^2 = \cos(2\theta)+i\sin(2\theta)$$

and if we take $\theta=2\phi$ and combine it with the previous formula, we get:

$$(\cos(\theta)+i\sin(\theta))^3 = \cos(3\theta)+i\sin(3\theta)$$

Proceeding by induction, we get deMoivre's formula:

$$(\cos(\theta)+i\sin(\theta))^n = \cos(n\theta)+i\sin(n\theta)$$

But it was Euler who made the next wonderful leap: fix the product $\phi=n\theta$ but let n get larger and larger, θ get smaller and smaller. Then $\sin(\theta)$ is very close to θ and $\cos(\theta)$ is very close to 1, so:

$$\begin{aligned}&\cos(\phi)+i\sin(\phi)\\ &\quad = (\cos(\phi/n)+i\sin(\phi/n))^n\\ &\quad \approx (1+i\phi/n)^n\end{aligned}$$

Now remember the rule we just found for compound interest, for continuous exponential growth. This was that:

$$e^{ax} = \lim_{n\to\infty}\left(1+a/n\right)^n$$

So now Euler makes the conclusion:

$$\boxed{e^{i\phi} = \cos(\phi) + i\sin(\phi)}$$

What in heaven's name does this mean? Well, it is really a definition of exponentials of imaginary numbers, the definition that one is compelled to make in order to keep arithmetic working smoothly. For example, look at what the formula which we labeled (*) above becomes if we use exponentials:

$$e^{i\theta}\cdot e^{i\phi} = \left(\cos\theta + i\sin\theta\right).\left(\cos\phi + i\sin\phi\right) = \left(\cos(\theta+\phi) + i\sin(\theta+\phi)\right) = e^{i(\theta+\phi)}$$

which is what must happen if the power law $a^b.a^c = a^{b+c}$ is true. And look what it says for derivatives:

$$\frac{d}{dt}\left(e^{iat}\right) = \frac{d}{dt}\left(\cos(at) + i\sin(at)\right) = -a\sin(at) + ia\cos(at) = ia\left(\cos(at) + i\sin(at)\right) = iae^{iat}$$

which is exactly the same equation we had above without the *i*.

Its most astonishing corollary is the special case when $y = \pi$, when it says:

$$e^{i\pi} = -1$$

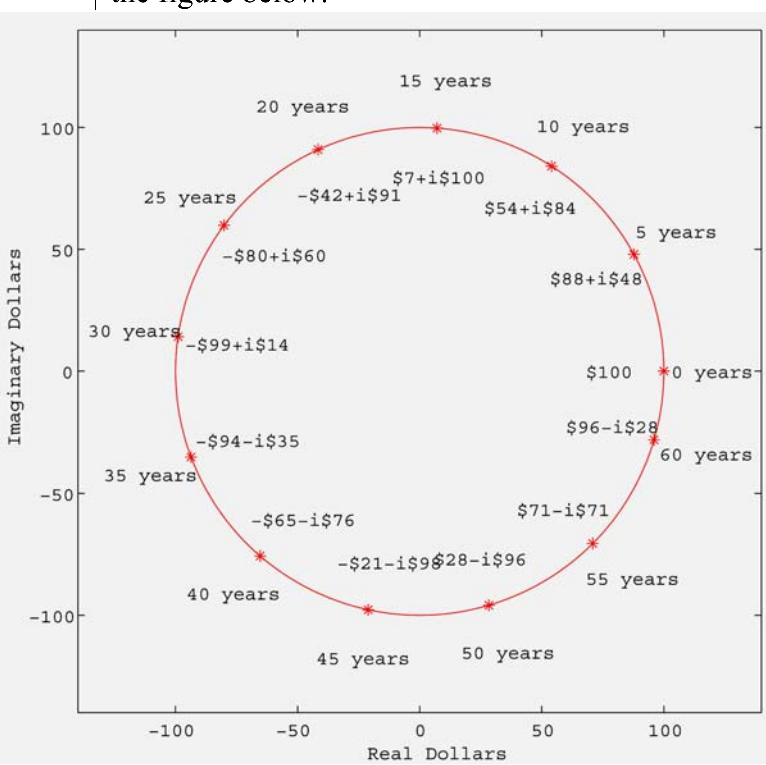
I once had to give an after dinner talk to a distinguished group of non-mathematicians about mathematics, and thought, can I explain this weird formula to them? Here's the explanation:

> Suppose an imaginative and enterprising banker decides to offer an exciting new type of savings account – one that pays imaginary interest, at the rate of $(10\sqrt{-1})$% each year. The public, fascinated by imaginary money, wants to participate in this new financial offering. Joe Bloggs deposits \$100 in such an account. After one year, he has earned 10 imaginary dollars in interest and his balance stands at \$$(100+10\sqrt{-1})$. The next year he gets 10 more imaginary dollars and is thrilled: but to his chagrin, the imaginary balance of 10 imaginary dollars also earns interest of $(0.1\sqrt{-1})$ x \$$10\sqrt{-1}$ dollars, or –1 real dollars. So his balance after two years stands at \$$(99+20\sqrt{-1})$. As the years go by, he keeps building up his pile of imaginary dollars but, as this gets bigger, he also sees the interest on this imaginary balance whittle away his real dollars at an ever-increasing rate. In fact, if the bank used continuous compounding of interest rather than adding the interest once a year, then after 5 years Joe would have \$$(88+48\sqrt{-1})$, having lost 12 real dollars in return for his 48 imaginary ones. Joe doesn't quite know what these imaginary dollars are good for, but maybe they aren't a bad deal in return for the 12 real ones he lost! Time passes and, after 10 years, he has \$$(54+84\sqrt{-1})$ and now his real money is bleeding away fast because of the interest on his imaginary balance. In fact, at 15 years, his balance is \$$(7+99.5\sqrt{-1})$ and finally at

15 years, 8 months and 15 days he checks his balance, only to find he no real money at all, but 100 imaginary dollars. This length of time is in fact 10π/2 years and what we have done is track his balance by Euler's formula. Explicitly, since continuous compounding is the same as using exponentiation, we have:

$$\begin{aligned}\text{Balance after } t \text{ years} &= (\text{Initial deposit}) \times \mathrm{e}^{(\text{Interest rate})t} \\ &= \$100 \times e^{0.1\sqrt{-1}t} \\ &= \$100\left(\cos(0.1t) + \sqrt{-1}\sin(0.1t)\right)\end{aligned}$$

Let's go on. More years elapse and now the interest on Joe's imaginary dollars puts him in real debt. *And* the interest on the real debt begins to take away his imaginary dollars. At 20 years, his balance stands at $\$(-41+91\sqrt{-1})$, at 25 years $\$(-80+60\sqrt{-1})$ and at 30 years $\$(-99+14\sqrt{-1})$. Finally at 10π years, which works out to be 31 years, 5 months, he finds himself 100 dollars in debt with no imaginary money. Not willing to give up, and finding the banker willing to extend him credit with only imaginary interest to pay, he perseveres and after about 47 years, finds that he has only imaginary debt now, and no real money either positive or negative. And now the interest on negative amounts of imaginary money is positive real money (because $(0.1\sqrt{-1}) \times (-100\sqrt{-1}) = +10$). So he finally begins to win back his real money. On his deathbed, after 20π years, that is 62 years and 10 months, he has back his original deposit and has paid off his imaginary debt. He promptly withdraws this sum, sues his banker and vows never to have any truck with complex numbers again. His odyssey is traced in the figure below.



Complex arithmetic takes a bit of getting used to, but, when you do, it's like driving a sports car instead of a sedan. Here's some points on how this works.

First note that when writing complex numbers in polar coordinates, instead of $x+iy = r(\cos(\theta)+i\sin(\theta))$ we can write $x+iy = re^{i\theta}$.

Second, the basic periodic function, with period p is not $\sin(2\pi t/p)$ or $\cos(2\pi t/p)$ but $e^{2\pi it/p}$. We have repeatedly used the fact that sums of sines and cosines are equal to a sine (or a cosine) with a phase shift. In complex terms, this comes out like this: if $A-iB = Ce^{iD}$, then

$$C\left(\cos(\theta+D)+i\sin(\theta+D)\right)=Ce^{i(\theta+D)}=\left(Ce^{iD}\right)e^{i\theta}=$$
$$=(A-iB)(\cos(\theta)+i\sin(\theta))=A\cos(\theta)+B\sin(\theta)+i(A\sin(\theta)-B\cos(\theta))$$

so, looking just at the real parts of the extreme right and left hand sides, we get

$$C\cos(\theta+D)=A\cos(\theta)+B\sin(\theta)$$

But isn't it simpler to just use $Ce^{i(\theta+D)}=(A-iB)e^{i\theta}$?

Thirdly, simple harmonic motion is much simpler with complex exponentials. Instead of the second order differential equation:

$$\frac{d^2x}{dt^2}=-x$$

we can start with the first order equation $\frac{dx}{dt}=ix$, because this implies:

$$\frac{d^2x}{dt^2}=\frac{d}{dt}\left(\frac{dx}{dt}\right)=\frac{d}{dt}\left(ix\right)=i\frac{dx}{dt}=i^2x=-x$$

In fact $\frac{dx}{dt}=-ix$ also implies $\frac{d^2x}{dt^2}=-x$. The solutions of these first order equations are e^{+it} and e^{-it} and adding and subtracting these, we get back our old solutions cos(t) and sin(t):

$$e^{it}+e^{-it}=\cos(t)+i\sin(t)+\cos(-t)+i\sin(-t)=2.\cos(t)$$
and
$$e^{it}-e^{-it}=\cos(t)+i\sin(t)-\cos(-t)-i\sin(-t)=2i.\sin(t)$$

Fourthly, Euler's formula allows us to express all those complicated expressions for Fourier series in a much simpler way. The Fourier series which we wrote as:

$$y(x)=B_0+A_1\sin(2\pi x/p)+B_1\cos(2\pi x/p)+A_2\sin(4\pi x/p)+B_2\cos(4\pi x/p)+\cdots$$

is now:

$$y(x)=\text{real part of } B_0+(B_1-iA_1)e^{2\pi ix/p}+(B_2-iA_2)e^{4\pi ix/p}+\cdots$$

In the most compact form, one just says that a periodic function $y(x)$ can be expanded as:

$$\boxed{y(x)=\sum_n a_n e^{2\pi inx/p}}$$

where a_n are complex coefficients which have both amplitude and phase built in.

The take home message is that, out of the blue, came this miracle $\sqrt{-1}$; this led to complex numbers; and these now prove to be by far the best way to handle much of two dimensional geometry, simple harmonic motion and Fourier series.

An epilog to this story is how quantum mechanics incorporates the complex numbers. Quantum mechanics is based on the startling idea that the most we can know about the world is represented by an object $\left|s\right\rangle$ and that any 2 such objects can be both added

$|s\rangle + |t\rangle$ and multiplied by complex numbers $re^{i\theta}|s\rangle$. The strange idea is that, as Schrodinger put it, in the world represented by $|s\rangle$, a cat might be alive, while in the world $|t\rangle$, it might be dead: it's a bit like probability of the cat being dead or alive. So in the world $|s\rangle + |t\rangle$ there is no way of knowing whether the cat is dead or alive: it's 50-50. Now the world $re^{i\theta}|s\rangle$ is indistinguishable from the world $|s\rangle$, but in there are many half-dead, half-alive cat worlds $|s\rangle + e^{i\theta}|t\rangle$ which really are different for different complex "phases" $e^{i\theta}$. This is hardly something that can be swallowed with this glib description but I hope it convinces you that the square root of minus one is deeply embedded in God's plan for the universe.