# Chapter Seven: The Pendulum and phase-plane plots

There is a story that one of the first things that launched Galileo on his scientific career was sitting in church and watching an oil lamp swinging at the end of the cord by which it was suspended from the high ceiling. He wondered how fast it swung from side to side, how this period depends on the weight of the lamp, the length of the cord and the size of the swing. As we saw above, his first sweeping principle was that all the phenomena related to the pull of gravity – the free fall of a body, the trajectory of a cannon ball, a body rolling down an inclined plane and the swing of a pendulum – all took place at a speed independent of the weight of the body. As we mentioned in the last Chapter, this is because gravitational force increases with the mass of the object, hence the acceleration, which is force divided by mass, is independent of mass.
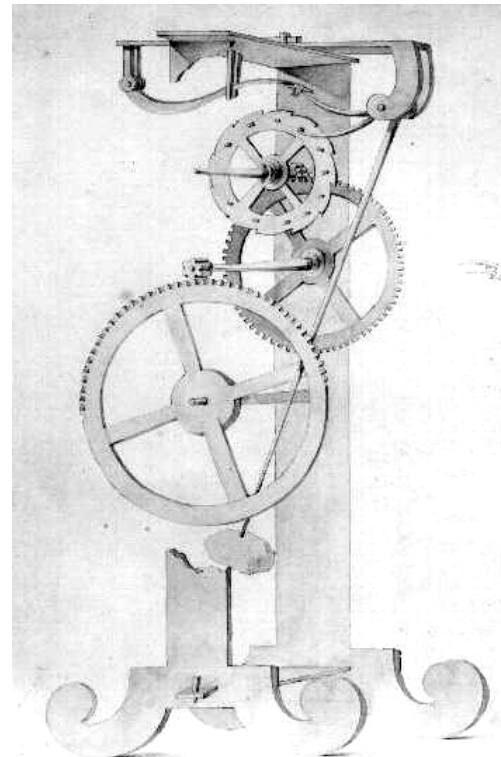
Now, experimentally, if the length is fixed, the period of the oscillation of a pendulum is found to be very nearly independent of how far it is swinging, that is of the amplitude of the oscillation. This is known as *isochrony*. Apparently, early in his work, Galileo believed this to be exactly true (allowing for air resistance) but then could not prove it and perhaps suspected later that there were systematic deviations from this isochrony – as indeed there are! In any case, he knew isochrony held sufficiently well that one could use a pendulum for the

timing mechanism of better clocks. The idea is that because of near isochrony, the clock will keep good time regardless of whether the pendulum inside is swinging strongly or weakly. The figure above is the pendulum based clock which Galileo designed. The technology for such clocks was perfected in the century after his life and this became one of the most successful designs for clocks (until they were superseded by electronic technology).

What is the differential equation for a pendulum? The figure on the right shows the forces: there is a downwards gravitational force of size mg on the bob. This, however, breaks up into a force which merely puts tension on the cord holding the bob plus a force which actually swings the bob. The breakup depends on the angle at which the
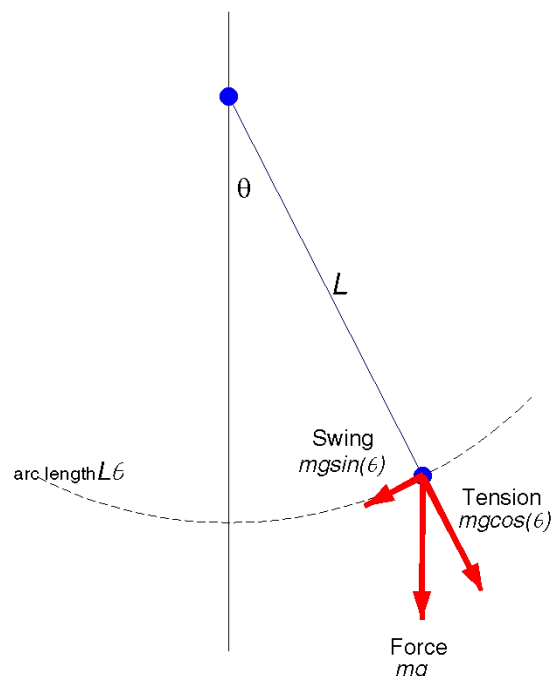
pendulum is found. If θ is this angle, then the stretch force is $mg \cos(\theta)$ and the swing force is $-mg \sin(\theta)$ (note the minus sign: when θ is positive, the force seeks to decrease θ and thus is negative). Now as θ is measured in radians, θ gives arc length on a *unit* circle; as the pendulum swings in a circle of radius L, arc length along the trajectory of the pendulum is θL. Thus Newton's equation $F=ma$ becomes:

$$F = -mg \cdot \sin(\theta), \ a = \left(L\ddot{\theta}\right) \text{ hence}$$

$$\ddot{\theta} = -\left(g/L\right)\sin(\theta)$$

Now here's the interesting mathematical side of this: if θ is small, $\sin(\theta)$ is very close to θ (remember that the limit of their ratios is 1!). If we replace $\sin(\theta)$ by θ, we have the simple harmonic oscillator $\ddot{\theta} = -(g/L)\theta$. This equation has the very simple but crucial property that if $\theta(t)$ solves this equation, then $2\theta(t)$, $3\theta(t)$, ..., in fact any multiple $c\theta(t)$ of $\theta(t)$ also solves the same equation. This is what isochrony means in formulas: if we scale up or down the oscillation by a factor c, we get a solution with larger or smaller amplitude and the same period as before. The irony is that this is *not* true for the exact pendulum equation especially when the amplitude gets large so that $\sin(\theta)$ and θ are not so close. Below is a graph of (i) the sinusoidal function $\sin(\theta)$, (ii) the linear function θ and (iii) a better approximation $\theta - \theta^3/6$ to sine.
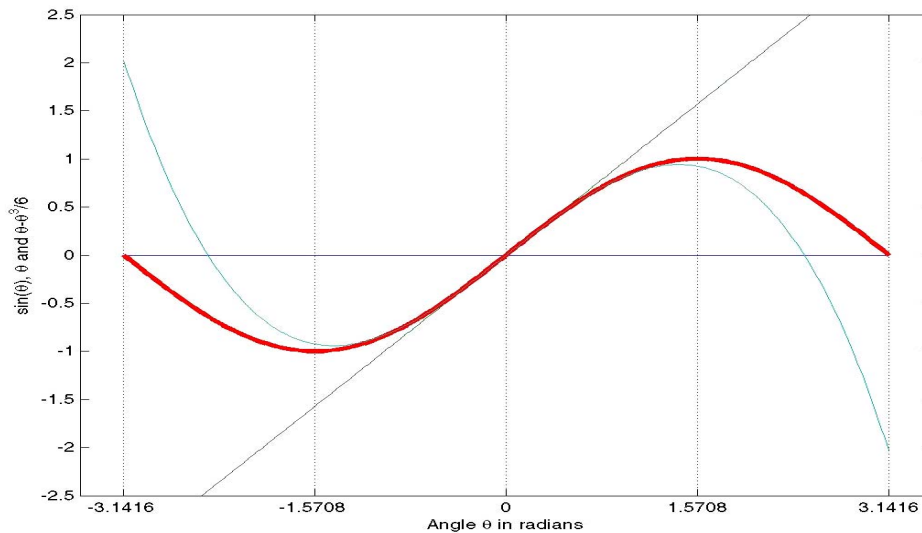
Note that the linear approximation of $\sin(\theta)$ is pretty good out to about $\pi/4$ or 45 degrees but then gets pretty bad.

From today's perspective, we can look back at Galileo and see how he was always looking for the *main effect* and that to find these, he put aside physical corrections like air resistance. But he also did not really have the mathematics to do the pendulum 'right'. It was Huygens, in the next generation, who worked it out, although he too did not have the full generality of Newton's laws. In particular, he worked out how to 'fix' the pendulum to make it exactly isochronous, even for large amplitude oscillations. The solution is to put 'cheeks' on each side of the cord near the point of suspension. As the pendulum

swings, a small part of the cord pulls taut against the cheeks and so the effective length *L* is shortened.

To get a good understanding of the pendulum and its dynamics, writing down some messy formula for $\theta(t)$ is not the best way. What works much better is to follow Oresme's dictum and seek a bird's eye view of its 'configuration' by making a suitable geometric representation. But this is not done best by plotting $\theta(t)$ against time *t*, that is graphing $\theta(t)$. It's better to do what we have slipped in a few times already: plot the pairs $(\theta(t), \dot{\theta}(t))$ as *t* varies. We did this for simple harmonic motion. In this case, plotting $\theta(t)$ gave us a sinusoidal curve while plotting these pairs gave us circles. This second plot is sometimes called a 'phase-plane plot'; with a raw data set of pairs (x_i, y_i), it is called a 'scatter plot'. To see what this does for us, write $v(t)$ for the time derivative $\dot{\theta}(t)$ and write the basic differential equation $\ddot{\theta} = -\sin(\theta)$ (we are assuming $g/L=1$ which can always be achieved by measuring time in suitable units) as a pair of differential equations:

$$\dot{\theta}(t) = v(t)$$
$$\dot{v}(t) = -\sin(\theta)$$

The first equation is just the definition of $v(t)$ and the second equation comes from rewriting $\ddot{\theta}$ as $\dot{v}$. Now in these equations, the rate of change of both variables is expressed as a function of their values at that point of time. In other words, it gives the direction in which the pair (θ, v) is moving whenever you know where it is. This can be expressed by drawing a whole lot of little arrows in the plane: at each point (θ, v), the arrow points to $(\theta + \Delta t \cdot v, v - \Delta t \cdot \sin(\theta))$, which is where it will go next. The result looks like this:
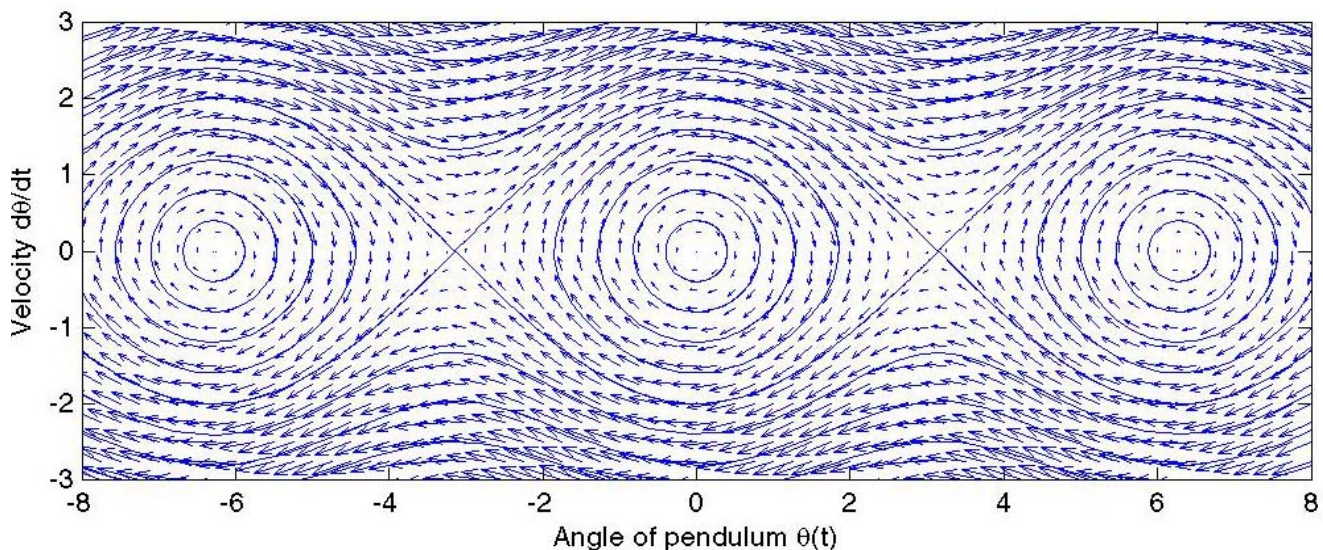
What are we looking at here? Each point in this plot corresponds to some pair of values $(\theta, v = d\theta/dt)$, a specification of both the position and the velocity of the pendulum. You can imagine the pendulum being released in any such state and then track what happens.

The arrows everywhere show the direction of motion of the two variables as the pendulum swings – Newton's term fluents seems especially apt here. The curves in the plot above show particular examples of ways the pendulum can swing. Thus the nearly circular curves around (0,0) represent the pendulum making regular small swings near its rest position. The circular curves around the points (2π,0) and (–2π,0) really represent the same motion because when an angle is increased or decreased by 2π, this is the same as 360 degrees and it's just the same angle. Whenever you make a graph and one of axes represents an angle, you must either just repeat the graph when the angle repeats or cut the graph off at some angle. We chose to allow the angle to repeat because we want to show the pendulum motions with higher velocity: when you start the pendulum off at $\theta =$ 0 but with high enough velocity, the pendulum goes all the way around. Of course, its velocity will slow down on the way up but then it will speed up on the way down again. In the absence of friction, it just keeps spinning around indefinitely. The counter-clockwise motions of the pendulum of this kind are shown in the graph by the wavy lines at the top that keep going from left to right indefinitely, while the curves on the bottom which go from right to left represent clockwise rotations.

How far is the pendulum from being isochronous? We can compute its period numerically as the amplitude of the swing increases. (There are fancier methods too, but this works just fine.) We'll describe the computer algorithm below. The result is this. So long as the pendulum doesn't swing more than 20 degrees or so, the error is less than 1%. No wonder Galileo had trouble measuring this.
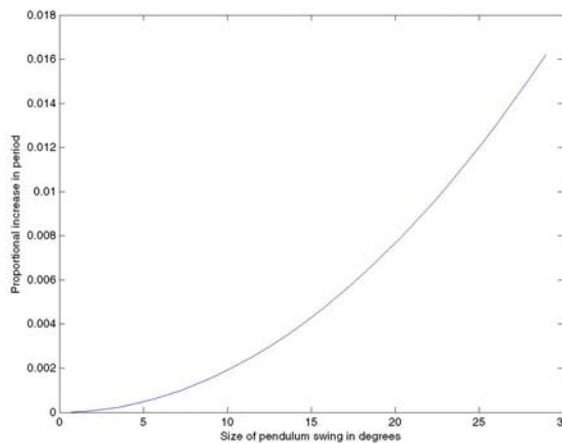
There is an interesting historical irony in the eventual mathematical analysis of the pendulum. The analysis that proved that clocks work depends on the invention of calculus, which in turn was driven by the experimental and theoretical results of Galileo, which in turn he could only have conceived of because time had come to be looked at as a precisely measurable quality – and this could only happen after clocks were invented around 1300. In other words, the whole development was circular: clocks had to be accepted as accurate before anyone could check that they in fact worked! The whole process took about 400 years. Presumably many scientific discoveries are like this.

What happens if we add in friction? Then, instead of swinging with fixed amplitude, back and forth, the pendulum should gradually slow down, taking smaller and smaller swings. Considered in the phase-plot, this comes out as a spiral. In each swing, the pendulum angle θ goes to a max, then the pendulum stops momentarily, then swings back gaining speed. But the speed when it comes back to the middle is slightly less. The result is that on the phase plot, it follows a spiral, getting closer and closer to stopping at (0,0). This is shown in the new phase-plot below:
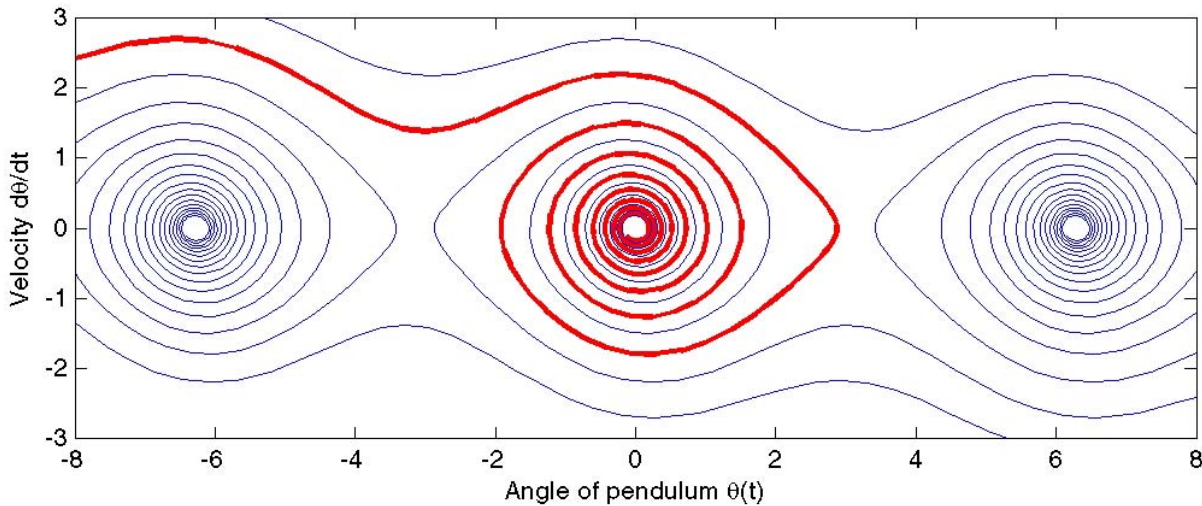
The swings which previously went over the top now can go all the way around only a few times before friction slows them down and they no longer have the speed to make it to the top. One such event is shown by the red curve; three similar events are shown in blue.

What is the equation for the pendulum with friction? You need to simply add in the force of friction. This is a force which is proportional to the velocity and in the opposite direction to the velocity:

$$\ddot{\theta} = -a\dot{\theta} - (g/L)\sin(\theta)$$

Note that there is a new constant a, the coefficient of friction that enters the equation. Now how did we make these figures? The sine term prevents us from having a simple formula for the solution, as in the case of simple harmonic motion. In fact, it is much more common that there is no explicit formula for the solution. As applied mathematicians, we resort to the computer and make numerical approximations which we plot. As pure mathematicians, we can at least prove the correctness of the behavior of these solutions, e.g. that without friction, we get periodic motion of the pendulum which repeats indefinitely, that the pendulum will not reach the top until it has a critical velocity and then it will continue whipping up and over the top indefinitely too. For the computation, you can just use:

$$\theta_{k+1} = 2\theta_k - \theta_{k-1} - a \cdot \frac{\theta_k - \theta_{k-1}}{\Delta t} - (g/L)\cdot\sin(\theta_k)$$

I want to finish off this discussion of clocks with another differential equation which is a fairly good model of the original clock, the 'foliot-and-verge' clock shown in Chapter 3, p.18. We will describe the position of the clock by the angle θ of the *foliot*, which, if you refer to that figure, is the horizontal bar on the top which swings back and forth, alternately releasing one or the other pallets from locking with the crown wheel. The crown wheel is constantly being pulled counter-clockwise by a heavy weight (not shown). There are 2 crucial angles of the foliot: if the top pallet is holding back the crown wheeel, then there is an angle at which this pallet can no longer hold the teeth of the crown wheel. And similarly for the bottom pallet. We will assume these angles are

$\theta = \pm 1$. The upshot is that when the top pallet is engaged, then $\theta<1$ and the crown wheel is putting a constant force on the pallet to increase θ. When θ hits 1, this pallet releases and the bottom pallet engages and the crown wheel is now putting a constant force to decrease θ until θ hits -1, and then everything repeats. The other force in the system is a lot of friction on the foliot which keeps it from going too fast. This is not a standard sort of differential equation but so what. Having chosen constants appropriately (after all, the medieval clock builders tinkered too), the force equation *F=ma* looks like:

$$\ddot{\theta} = 4(1 - \dot{\theta}) \text{ when pallet 1 engaged,}$$

$$\ddot{\theta} = 4(-1 - \dot{\theta}) \text{ when pallet 2 engaged}$$

The ± 4 forces represent the crown wheels push and pull on the foliot; the $-4\dot{\theta}$ is just friction. If we make a phase plot, the whole thing gets a lot clearer:

Angle θ of the foliot,
normalized so that at θ=± 1, the pallets release the crown wheel
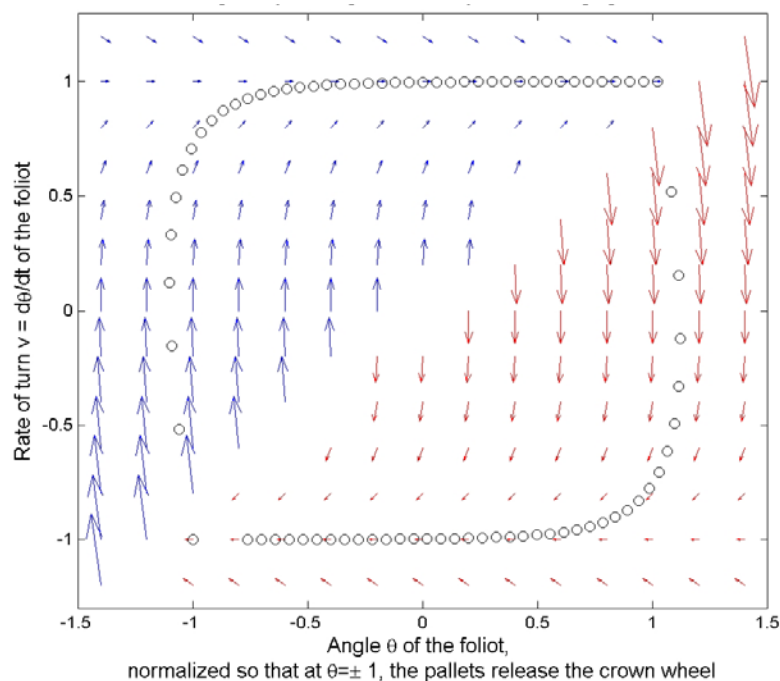
The circles represent the clock in action, each showing the next state after the same amount of time. The top right corner is where pallet 1 is released and the bottom left corner is where pallet 2 is released. The slow horizontal motions in the phase plot, where the dots are close together, show the foliot swinging at nearly constant velocity +1 or –1, but limited by friction, so there is almost no acceleration. These are interspersed by the rapid vertical part of the motion on the two sides, where the dots are widely spaced, in which the angle doesn't change much but the velocity switches from +1 to nearly –1 or vice versa. You have to trace this through to believe it all works and is an intuitively reasonable model of the medieval clock.

Problem: Differential equations have been proposed to model many things in nature besides mechanical contrivances. An influential model is the so-called *predator-prey* model, which models 2 species, say foxes and rabbits, and their population cycles over a period of years. When there a lot of rabbits, there is plenty for the foxes to eat and they

multiply rapidly. But then the rabbit population is depleted and, as a result, the foxes starve. When the foxes are nearly gone, the rabbits flourish again. And so 'the cycle of nature repeats' as nature specials on public TV have told us so often. OK, let's do math. Instead of forces, we now have birth and death. Let $r(t)$ be the population of rabits and $f(t)$ that of foxes. If a is the reproduction rate of each rabbit, then r should increase at a rate ar due to births. If b is the probability of one fox eating one rabbit in some unit of time, then bfr is the death rate due to rabbits being eaten. Thus:

$$\dot{r} = ar - bfr$$

Now let c be the excess of the death rate of foxes over their birth rate when there are no rabbits to eat and let d be the increase in births minus deaths of foxes for each rabbit in the population. Then:

$$\dot{f} = -cf + dfr$$

(We're clearly simplifying a lot, e.g. ignoring the natural deaths of rabbits from old age.) These equations were first written down by Lotka in 1910 and further studied by Volterra, hence are also called the *Lotka-Volterra* equations. At first sight, there are 4 constants a,b,c,d here but there's really only one essential constant. If we change the units in which f and r are measured, replacing f by af/b and r by dr/c, then b and d drop out and the equations become $\dot{r} = a(1-f)r, \dot{f} = c(r-1)f$ and if, further, we rescale time to proceed c times faster, then c drops out and only the ratio a/c remains: $\dot{r} = (a/c)(1-f)r, \dot{f} = (r-1)f$. Take $a/c{=}2$ as a simple case and first, by hand, make a rough phase-plot with arrows as above, showing at each point (f,r) where the point is moving. Then solve this equation starting at $(1/2,1/2)$ approximately in Excel with some small time step Δt. Experiment with various Δt until the result stabilizes: what is the system doing? Try other values for a/c. What would you guess about the long term behavior of solutions $f(t),r(t)$?

Chapter Eight: Gravity and the Planets

Predicting the position of the sun, moon and planets in the skies has been a goal of every civilization of mankind. Without electric lights or pollution, the heavenly bodies in the sky are very prominent. The connection of the sun's position, high or low in the sky, with the seasons, the apparent connection of the moon with women's menstrual cycles made them central events in life. And, unlike the 'fixed stars', the planets zip around in complex ways, seeming to have a life of their own. But the more mankind recorded celestial events, the positions of these bodies at precise times, the more complications seemed to appear. The sun and the moon, for example, slow down and speed up at various times each year or month, the path traced by the sun and the moon against the stars changes slowly over the years, the motion of Mars and its brightness is extremely irregular, eclipses seem totally unpredictable. It is a just source of pride for European civilization that Newton found the golden key that finally predicted every little wrinkle in these motions (well nearly every one – Einstein improved a remaining glitch in the motion of Mercury).

Newton apparently told several people that the key insight came to him when sitting in the garden of his family home in Lincolnshire during the plague years. An apple fell from a nearby tree and suddenly he saw an analogy between the falling of the apple towards the center of the earth and the force that kept the moon in its orbit instead of sailing off on a straight line away from the earth. Maybe the same sort of force was pulling the apple to earth and pulling the moon towards the earth with exactly the right force to keep it in a roughly circular orbit! How does this work out?

We know how to calculate the acceleration of a body towards a point P which is needed to keep that body moving in a circle with center P. As we saw when measured the size of the earth, if you move *in a straight line* tangent to the earth a distance x, then you wind up above the earth's surface by $x^2/2r$ – so long as x is assumed small compared to the radius r of the earth. This applies just as well to the moon: let r be the distance from the moon to the earth and let x be the distance the moon would move in 1 second if it went straight. Then it would have moved beyond its circular orbit by the distance $x^2/2r$. Let's work this out approximately: in 28 days, the moon moves a distance $2\pi r$. r is roughly 240,000 miles, so in one second it moves how many feet?
$$2\pi(240000)(5280)/(28.24.60.60) \approx 3300 \text{ feet}$$
This means that to stay in its orbit, it must fall towards the earth in 1 second:
$$(3300)^2\big/(2.240000.5280) \approx .0043 \text{ feet}$$
Now the apple on the earth's surface falls 16 feet in one second (remember Galileo's $(g/2)t^2$), so this is less by a factor of about 3700. *But the moon is about 60 times further away from the center of the earth than the apple and* 60^2 *is* 3600! Newton's observation was that the force needed to keep the moon in its orbit was the inverse square of the how much further away it was (to within the accuracy of his observations). This it: the inverse square law which unlocked everything.

When Newton began to think about gravity about 1665, Kepler's laws had been known for nearly 50 years. Kepler had struggled to find a better way of expressing the then abundant data on the exact positions and motions of the sun, moon and planets, with many false starts, (such as his idea that the planets moved on spheres spaced out by the regular solids, the tetrahedron, cube, octahedron, dodecahedron and icosahedron, which now seems almost ludicrous). In spite of his desires for something more elegant he was finally forced to consider ellipses. In 1609, he published his three laws about planetary motion that were much much more accurate than anything before:

1. Each planet moves in an ellipse with the sun at one of its foci,
2. A line drawn from each planet to the sun sweeps out equal areas in equal times.
3. The period of each planet is proportional to the 3/2 power of the average of its closest and furthest distance from the sun.

The question in front of Newton was – did his idea of universal gravity explain these three laws? In *Principia*, Newton begins by studying all possible centripetal force laws, i.e. there is a fixed attracting body at the origin in the plane and a second body moving around it and being attracted towards the central body by a force which depends in some way on the distance between the two bodies. If the second body, let's call it the planet, is at (x,y), then we can write this as:

$$\ddot{x} = -f\left(\sqrt{x^2 + y^2}\right) \cdot x,$$
$$\ddot{y} = -f\left(\sqrt{x^2 + y^2}\right) \cdot y$$

for some function f. In other words, if $r = \sqrt{x^2 + y^2}$ is the distance of the planet from the origin, then it is attracted to the origin with a force $(-f(r)x, -f(r)y)$. This is a force with magnitude $r.f(r)$. Note that the *inverse square law* is the case $r.f(r) = G/r^2$, for some constant G, i.e. $f(r) = G/r^3$.

His first theorem on this is very pretty: for *any such force law*, the motion of the planet will satisfy Kepler's second law: if you join the planet to the origin, the areas swept out by that line in equal times will be equal. His proof of this is shown on the next page. It is a good example of what mathematicians like to call an elegant proof! Note that the planet's motion is shown approximated by the polygon *ABCDEF* and Newton then constructs its motion by combining (a) a straight linear continuation, e.g. *Bc* after *AB* plus (b) a displacement caused by the centripetal force *BV* towards *S*. Read and see how simple facts about areas of triangles show that *ABS*, *BcS* and *BCS* all have the same area.

To find centripetal forces

Proposition 1[a] *The areas which bodies* [b]*made to move in orbits*[b] *describe by radii drawn to an*
Theorem 1 *unmoving center of forces lie in unmoving planes and are proportional to the*
times.
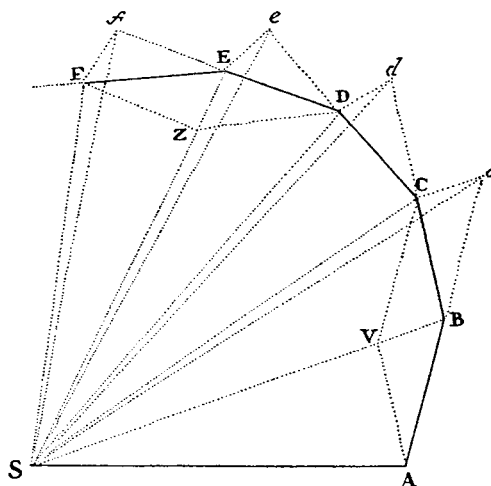
Let the time be divided into equal parts, and in the first part of the
time let a body by its inherent force describe the straight line AB. In the
second part of the time, if nothing hindered it, this body would (by law 1)
go straight on to *c*, describing line B*c* equal to AB, so that—when radii AS,
BS, and *c*S were drawn to the center—the equal areas ASB and BS*c* would
be described. But when the body comes to B, let a centripetal force act with
a single but great impulse and make the body deviate from the straight line
B*c* and proceed in the straight line BC. Let *c*C be drawn parallel to BS and
meet BC at C; then, when the second part of the time has been completed,
the body (by corol. 1 of the laws) will be found at C in the same plane as

triangle ASB. Join SC; and because SB and C*c* are parallel, triangle SBC
will be equal to triangle SB*c* and thus also to triangle SAB. By a similar
argument, if the centripetal force acts successively at C, D, E, . . . , making
the body in each of the individual particles of time describe the individual
straight lines CD, DE, EF, . . . , all these lines will lie in the same plane;
and triangle SCD will be equal to triangle SBC, SDE to SCD, and SEF to
SDE. Therefore, in equal times equal areas are described in an unmoving
plane; and by composition [or componendo], any sums SADS and SAFS of
the areas are to each other as the times of description. Now let the number
of triangles be increased and their width decreased indefinitely, and their
ultimate perimeter ADF will (by lem. 3, corol. 4) be a curved line; and
thus the centripetal force by which the body is continually drawn back from
the tangent of this curve will act uninterruptedly, while any areas described,
SADS and SAFS, which are always proportional to the times of description,
will be proportional to those times in this case. Q.E.D.

This is Theorem is interesting because it also has a very simple algebraic proof, using a simple formula for the area of a triangle in Cartesian coordinates – the technique which Descartes had developed some years before Newton wrote. So perhaps this is a good example of how Newton sought to cover his tracks, to rewrite his algebraic and analytic arguments using the 'good Geometry' of the ancients. How would the proof have gone if Newton had used algebra? He would start by citing the formula[1] for the area of a triangle with vertices (0,0), (x,y) and (u,v), namely

$$\text{area} = \tfrac{1}{2}\,|xv\text{-}yu|.$$

Then letting S be the origin, A be (x,y) and B be $(x+\dot{x}\cdot\Delta t, y+\dot{y}\cdot\Delta t)$, we find the area of ABS to be $|x\dot{y}-y\dot{x}|\cdot\Delta t$. c is then $(x+2\dot{x}\cdot\Delta t, y+2\dot{y}\cdot\Delta t)$ and the force and hence the acceleration of the planet is a vector $-e\cdot(x+\dot{x}\cdot\Delta t, y+\dot{y}\cdot\Delta t)$ for some small constant e. Thus C is $(x+2\dot{x}\cdot\Delta t, y+2\dot{y}\cdot\Delta t)-e\cdot(x+\dot{x}\cdot\Delta t, y+\dot{y}\cdot\Delta t)$. Plugging these into the area formula, it's just a mechanical verification that area(ABS) = area(BCS). Algebra has a way of making a lot of things easy!

Even easier is to use the machinery of calculus! Let $A(t)$ be the area swept out by the line from the planet to the sun. Then, because of the area formula above, we get:
$$\dot{A} = (x\dot{y}-y\dot{x})/2$$
Differentiating again, we get
$$\ddot{A} = ((\dot{x}\dot{y}+x\ddot{y})-(\dot{y}\dot{x}+y\ddot{x}))/2 = \left(x(-f(x,y)y)-y(-f(x,y)x)\right)/2 = 0$$
Thus $\dot{A}$ is constant, and area is swept out at a constant rate.

Thus Kepler's second law is consistent with any centripetal force law f. But his third law shows that immediately that the force law must be an inverse square. The reason is that this law says that if we have one solution $x(t),y(t)$ of the equation, then $A.x(bt), A.y(bt)$ must also solve the equation provided $A^3 b^2 = 1$, i.e. you speed up (or slow down) time by a factor b, hence you multiply its period by $1/b$, and you make the orbit bigger or smaller by a factor A, then Kepler's law says that the ratio $1/b$ of the periods should equal the 3/2 power of the ratio A of the orbit sizes. If we assume that both $x(t),y(t)$ and $A.x(bt), A.y(bt)$ satisfy the same equation of the form:

$$\ddot{x} = -f\left(\sqrt{x^2+y^2}\right)\cdot x,$$

$$\ddot{y} = -f\left(\sqrt{x^2+y^2}\right)\cdot y$$

---

[1] This is usually shown when cross product is introduced. You can use the area formula: ½ times base times height. Taking the line from (0,0) to (x,y) as the base, the height is found by writing the vector (u,v) as a sum of a multiple of (x,y) and of the perpendicular line (-y,x). Subtracting any multiple of (x,y) from (u,v) doesn't change either the height, the area of the value of |xv-yu|, so we can assume (u,v) is perpendicular to (x,y), i.e. u=-ay,v=ax. Then it follows that |xv-yu|=a||x,y||²=base x *height*.

then, by comparing the resulting two equations, we find that we must have $b^2 = f(Ar)/f(r)$. This holds for any r,A,b such that $A^3 b^2 = 1$. Letting $r = 1$, we find $f(A) = f(1)/A^3$ or, in other words, the force is given by:
$$r.f(r) = G/r^2.$$

It is quite a bit harder to derive Kepler's first law, namely that if the force law is inverse square $r.f(r) = G/r^2$, then the planet must move in an ellipse with the origin (e.g. the sun) at one focus. 'Focus' you say, what's that? When I was in high school, you learned conic sections but no longer. In a few words, an ellipse is a circle squashed in some direction. In coordinates, if a circle is squashed in the x or y direction, you get:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$
which goes around the origin and through $(\pm a, 0)$ and $(0, \pm b)$. Assume $a > b > 0$, so it is squashed in the y direction relative to the x. The foci of this ellipse are $(\pm c, 0)$ where $c = \sqrt{a^2 - b^2}$. Why are the foci important? Tie a string of length $2a$ between the two foci and stretch it out with a pencil and move the pencil around, keeping the string taut, i.e. getting all points such that the sum of their distances to the 2 foci is $2a$. This draws the ellipse! Anyway, let's just check that the planetary orbits for Newton's equations:
$$\ddot{x} = -Gx\big/r^3,$$
$$\ddot{y} = -Gy\big/r^3$$
are ellipses with (0,0) as a focus. It was Halley's asking Newton whether he knew this was true (of course he did, although he had forgotten his proof!) that triggered Newton's decision to finally publish *Principia*.

This is a somewhat nasty piece of algebra (the French call a calculation like this *assez penible*). We start with the points $(x,y) = (a.\cos\theta, b.\sin\theta)$ on the ellipse above. But we must shift one of the foci to the origin, so better, let $x(\theta) = c + a.\cos\theta$, $y(\theta) = b.\sin\theta$. We need to work out how far this point is from the origin:

$$\|(x,y)\|^2 = (c + a\cos\theta)^2 + (b\sin\theta)^2 = (c^2 - b^2) + 2ac\cos\theta + (a^2 - b^2)\cos^2\theta = (a + c\cos\theta)^2$$
(using $a^2 = b^2 + c^2$) hence $\|(x,y)\| = a + c\cos\theta$.

Next, work out the rate of change of this point with respect to a change in θ:
$$\dot{x} = -a.\sin(\theta), \dot{y} = b.\cos(\theta), (\theta \text{ rate of change})$$
However, what we want is to make θ into a function of time in such a way that Kepler's first law is satisfied. Well, as we saw above the area of the triangle with vertices (0,0), $(x(\theta), y(\theta))$ and $(x(\theta) + \dot{x}(\theta)\cdot\Delta\theta, y(\theta) + \dot{y}(\theta)\cdot\Delta\theta)$ is just $|x\dot{y} - y\dot{x}|\cdot\Delta\theta$. Substituting, this works out as $b(a + c\cos(\theta))\cdot\Delta\theta$. So we have to make equal units Δt of time correspond to equal increments $(a + c\cos(\theta))\cdot\Delta\theta$ of θ. (We drop the factor b because it is a constant.) This means that the true *time* derivatives of x and y should not be their θ-derivatives as above but rather:

$$\dot{x} = \frac{-a.\sin(\theta)}{(a+c\cos(\theta))}, \dot{y} = \frac{b.\cos(\theta)}{(a+c\cos(\theta))}, \text{ (time rate of change)}$$

Now the final step: compute the acceleration of the planet and, hopefully, it will look like the law we want. This means we differentiate the above expressions with respect to θ and then divide again by $(a+c\cos(\theta))$ to get the true time derivative. This is a mess and either trust me or get a big piece of white paper and work it out yourself. It works out that:

$$\ddot{x} = \frac{-ax}{(a+c\cos(\theta))^3} = \frac{-ax}{r^3}, \ddot{y} = \frac{-ay}{(a+c\cos(\theta))^3} = \frac{-ay}{r^3}$$

So we finally have all three of Kepler's laws for planetary motion, assuming the sun attracts each planet by the inverse square law of gravity.

But, hold on, you say: the law of gravity says that the sun is attracted by the planets too, and the planets should attract each other, etc, etc. And here's another problem: each part of the sun is attracting the earth and all these parts, as seen from the earth for example, cover half a degree of sky. So were we right to assume the sun was a *point* mass attracting the earth? All these should complicate things.

Well, yes and no. I'm not going to go into detail but here is what emerges. If there are two bodies *A* and *B* attracting each other by the inverse square law, then they simultaneously attract each other but the effect is that they both move in ellipses with their center of gravity at one of its foci. In the case of where one body is much more massive than the other, then the heavy one hardly moves and we are back in the case studied above. The second point really bothered Newton, but he was able, after considerable work, to show that any body of uniform density attracts other bodies in exactly the same way as if all its was concentrated at its center. That's the good news.

But if one looks at *three* (or more) bodies, all attracting each of the others with a force proportional to the product of their masses and inversely as the square of their distance apart – then lots of unexpected and really complex things can happen and no simple formulas exist for the resulting paths. It's a remarkable fact that with *two* bodies, the effect of gravity is so simple and elegant while with *three* it's so complex. So, when dealing, for example, with the motion of the moon around the earth, the sun's attraction as well as the earth's must be taken into account and Kepler's laws don't hold very well. And the outer planets, especially Jupiter and Saturn are very massive so they affect the motion of all the other planets in a measurable way. Working out the effect of Newton's laws and taking into account all the attractions in the solar system is very challenging, especially before computers were devised. But it did turn out that they predicted everything correctly up to the accuracy of contemporary measurements, except (as mentioned above) for a small glitch in Mercury's orbit whose explanation waited for Einstein's total reformulation of the theory.

With some caveats, it is not difficult to follow three such bodies by computer simulation and so we have worked out a problem on this. First of all, we consider the *restricted three body problem*: the case where two of the bodies have very large mass, hence move in ellipses around each other while the third has very small mass, so it doesn't effect the

motion of the two massive bodies while being attracted by them. To choose an interesting case, the problem below deals with something that might well have happened to us, here on earth. There are lots of binary star systems in the heavens – pairs of stars circling around each other. And there is no obvious reason why such a binary pair shouldn't have planets moving in some orbits around the pair of them. So we want to compute exactly such a situation, assuming two heavy suns of equal mass and a third body which is relatively light planet going around them both. But now the attraction of the binary pair is *not* at all the same as being attracted to a single virtual mass between the pair, so the planet is not going to follow an ellipse. Here's the problem:

Problem: Let's work out another counterfactual universe! In this one, we suppose we are on a planet orbiting a binary star system, say a green star and an orange star of equal mass. Life is fun: there are two sunrises and two sunsets, very colorful. But the earth's motion is now much less predictable. In fact, it is dangerously unstable. Let's set this up like this:

1. Let the two suns have mass M, let them be at a distance $2d$, moving in a circle of radius d around each other:

$$(x_{S1}, y_{S1}) = d(\cos(at), \sin(at)), \quad (x_{S2}, y_{S2}) = -d(\cos(at), \sin(at)),$$

This makes both of them have constant velocity da and constant acceleration, each towards the other, of size da^2, so this will be produced by the law of gravity provided that:

$$\frac{GM^2}{4d^2} = \text{force} = \text{mass} \times \text{accel.} = Mda^2, \text{ or } GM = 4d^3a^2$$

2. Next, suppose that a planet of much smaller mass m circles both of them, starting at a distance e. Let's give it the same velocity it would have if it were circling a single sun of mass $2M$ at distance e. If it were so simple, we'd have

$$(x_P, y_P) = e(\cos(bt), \sin(bt)),$$

making the velocity eb and the acceleration eb^2. Again, the law of gravity puts a constraint on this:

$$\frac{GMm}{e^2} = \text{force} = \text{mass} \times \text{accel.} = meb^2, \text{ so } GM = e^3b^2.$$

So let's require that a,b,d,e satisfy $4d^3a^2 = e^3b^2$. But because there are two suns, the planet won't move so simply.

3. Set this up in Excel like this: make the columns represent
    A. *time,*
    B. *x position of sun 1,*
    C. *y position of sun 1,*
    D. *x position of planet,*
    E. *y position of planet,*
    F. *distance from sun 1 to planet and*
    G. *distance of sun 2 to planet.*

Then we can calculate columns *A,B,C* using the formulas above. I found it convenient to take $a = 2\pi$ so that the sun's period is $t = 1$ and a time sampling of $\Delta t = 1/100$, giving

100 samples around each orbit of the two suns. We can fix the units of space by taking d = 1 (so now $GM = 16\pi^2$). Note that columns F,G can be calculated explicitly from columns B,C,D,E, and to propagate columns D,E, we use Newton's law:

$$\ddot{x} = GM\left(\frac{x_{S1} - x}{r_1^3} + \frac{x_{S2} - x}{r_2^3}\right),$$

$$\ddot{y} = GM\left(\frac{y_{S1} - y}{r_1^3} + \frac{y_{S2} - y}{r_2^3}\right),$$

$$r_1 = \sqrt{(x_{S1} - x)^2 + (y_{S1} - y)^2},$$

$$r_2 = \sqrt{(x_{S2} - x)^2 + (y_{S2} - y)^2}$$

We can discretize using the same method as above, with second differences allowing us to move ahead one step in time.

4. As a warmup, take e very large and start the planet off at $(e,0)$ with velocity $(0,be)$ and check that the planet orbits nicely in a circle. You'd better take $\Delta t = 1/100$ or it will take forever for the planet to get around the distant suns. Then increase the velocity and you should find the planet going in a large ellipse. A good way to track the results is to plot x and y against each other and also $\|(x,y)\|$ and atan2(x,y) against time (the function 'atan2' coverts a point (x,y) in the plane to its angle in polar coordinates).

5. Then, try e = 4, $\Delta t = 1/100$ and compute the orbit for $0 \le t \le 8$. This is the fun case. What do you find? If a 'year' is the time for the planet to circle both suns, how many times each year will the 2 suns line up – causing consternation among the inhabitants of this planet. There will be orange eclipses and green ones when the orange or green sun comes in front of the other. You may also want to plot the angle between the 2 suns as seen from the planet: how far apart do they get in the sky and what does this mean for twilight? (To work out this angle ϕ, look at the triangle formed by the planet and 2 suns. Its area can be expressed by the formula above in terms of the positions of the planet and the suns and can also be written as $(1/2)r_1 r_2 \sin\phi$.) The ancient astronomers of this planet would have been confounded by the complexity of the motion of the two suns.

6. Finally try e = 3.5, $\Delta t = 1/100$. Try fiddling with the initial velocity – can you make the planet behave and stay near the suns? I found huge instability. The simulations tend to be inaccurate for this time step but the overall effect seems the same even if the time step is much smaller.