Physics 305: Ch. 1

Orbital Motion

In last week's homework, we used the second-order Runge-Kutta integration scheme to solve the equations of motion of particles falling near the surface of the earth. This week, we will examine the motion of falling bodies again — this time planets falling towards the Sun.

For this physical situation, second-order equations are not always accurate enough to properly model the interactions of bodies in orbit. For possibly the first few orbits, the motion determined using the Runga-Kutta equation is very close to the "real" motion, but gradually diverges from the correct motion so after many orbits the motion is no longer related to the initial conditions. We will continue to use this method with a suitably small time step. However, you should keep in mind a higher order integration scheme may be necessary to achieve a solution of high accuracy.

1.1 A review of classical mechanics

Let's review a few of the most basic facts:

For almost any motion

$$m\vec{a} \equiv m\dot{\vec{v}} = \vec{f} \tag{1.1}$$

"force changes velocity, not position"

Often the force is minus the gradient of some potential energy.

$$\vec{f} = -\nabla\Phi(\vec{x}) \tag{1.2}$$

When this is true, (and sometimes when it isn't), energy is conserved:
For forces which are the gradient of a potential

$$\dot{E} = \frac{d}{dt}\left(\frac{1}{2}mv^2 + \Phi(\vec{x})\right) = 0 \tag{1.3}$$

The gravitational force (here from a body with mass M fixed at the origin) is of this type:

$$\frac{GMm}{r^2}(-\hat{r}) = -\nabla\left(\frac{-GMm}{r}\right) \tag{1.4}$$

Here r is the magnitude of the position vector $\vec{r}$, and $\hat{r}$ is a unit vector pointing in the direction of $\vec{r}$. Notice that the potential energy, $-GMm/r$, is negative, and we have chosen the zero so that the potential goes to zero as $r \to \infty$.. Obviously the kinetic energy, $\frac{1}{2}mv^2$, is positive. If the total energy is negative, the object can never escape to infinity — it is in a bound orbit. If the energy is positive, it can keep on going out forever — bye-bye.

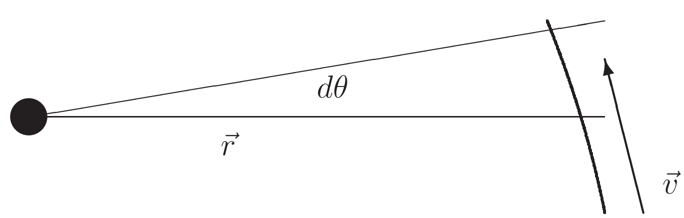
In addition to being the gradient of a potential, this gravitational force is a "central force" — in the direction of $\hat{r}$. For central forces, angular momentum is conserved

$$\dot{\vec{L}} = \frac{d}{dt} m\,\vec{r} \times \vec{v} = \vec{0} \tag{1.5}$$

For motion in a plane (which we take to be the xy plane), $\vec{L}$ is in the $\hat{z}$ direction, and we can treat it as a scalar:
In two dimensions

$$L = m\,(xv_y - yv_x) \tag{1.6}$$

We will often use conservation of angular momentum rephrased as follows:
The rate at which the planet's orbit "sweeps out" area is constant



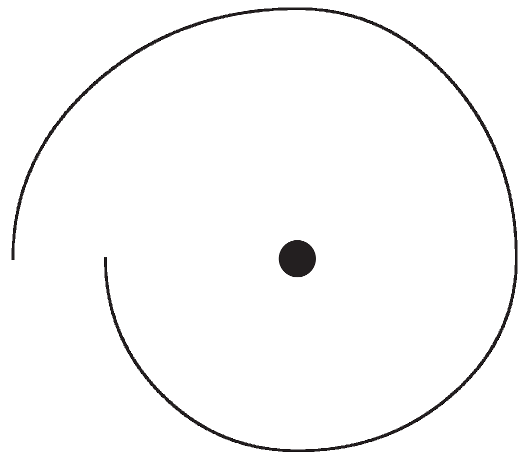
For infinitesimal $d\theta$, the length of the segment of the orbit on the right side is $v\,dt$, and the area of the segment is one half the radius times the component of $v\,dt$ perpendicular to $\vec{r}$, which is just the angular momentum.

$$\frac{d}{dt}(\text{area swept out}) = \frac{1}{2m}\vec{r} \times \vec{v} = \frac{L}{2m} \tag{1.7}$$

Now let us apply this to the motion of a planet in the sun's gravitational field.

First, bound orbits in a $1/r$ potential are closed. If the orbit ever returns to the same angle in polar coordinates, it will return at the same radius and same velocity.

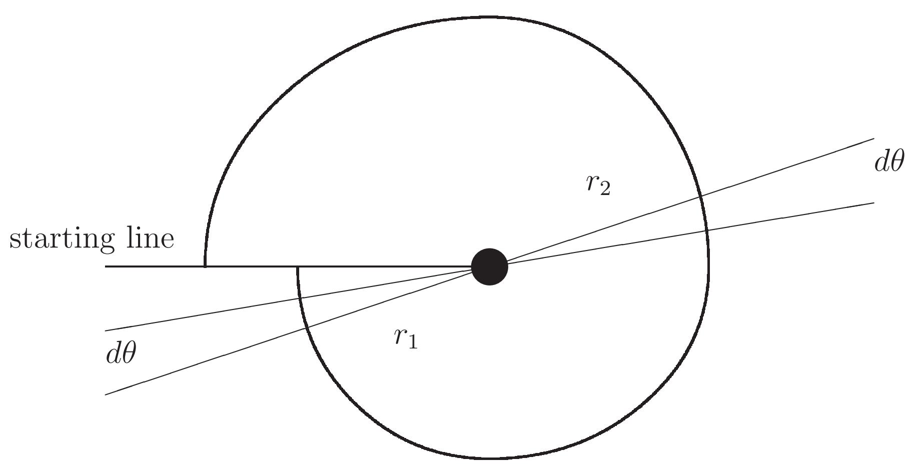
In other words, you can't have an orbit like this:



This is actually a remarkable feature; only a few potentials have this property. There are deep connections between this property and other things, for example the fact that the quantum mechanical energy levels of the hydrogen atom are given by a simple formula.

Perhaps even more remarkably, it is possible to understand this fact using only the tools of introductory level mechanics. This discussion is condensed from H. Abelson, A. diSessa and L. Rudolph, *Am. J. Phys.* **43**, 579 (1975).

To see that orbits in a $1/r$ potential are closed, consider two infinitesimal segments of the orbit on opposite sides, covering the same angle $d\theta$.



From conservation of angular momentum, the time spent in each of the two segments is proportional to the area of the segment. Since they subtend the same angle, the area of each segment is proportional to the square of the radius:

$$\frac{t_1}{t_2} = \frac{r_1^2}{r_2^2} \tag{1.8}$$

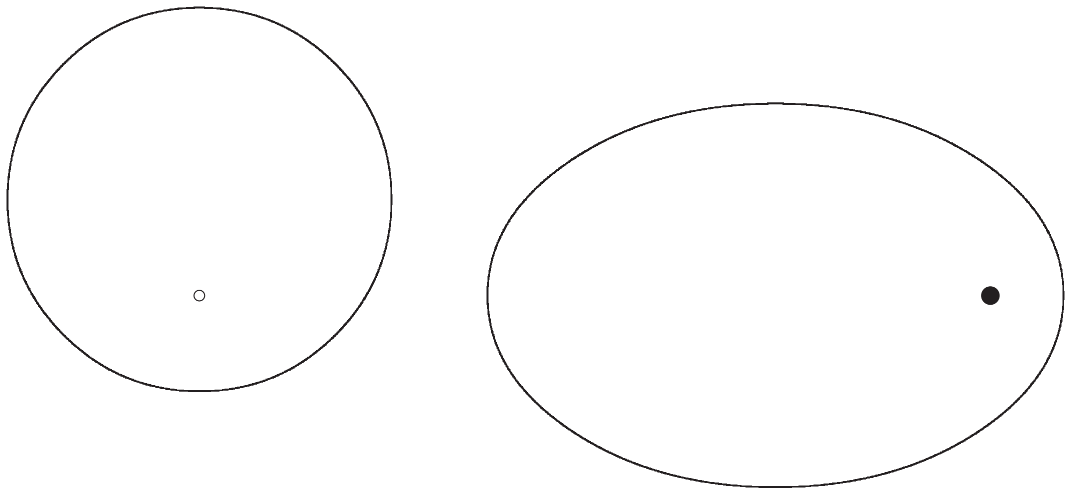
The force is in opposite directions in the two segments, and its magnitude is proportional to $1/r^2$. Therefore the **impulse**, or change in momentum, that the planet gets in segment 1 is exactly opposite to the impulse that it gets in segment 2, $\vec{\Delta}p = \vec{f}\,dt$. Therefore, when the planet returns to the same polar angle from the sun (the end of the orbit drawn in the picture above), it must have exactly the same velocity (both x and y components!) as it did when it started. The total change in velocity as it returns to the starting line is zero.

But now energy conservation tells us that it must have exactly the same radius. If $mv^2/2 - GMm/r$ is constant, if you come back across the starting line with the same $\vec{v}$, you must also come back at the same r.

Once again, **if you start off at some position with some velocity in a $1/r$ potential, unless the orbit is unbound you will come back to the same position with the same velocity.** This is a special property of the $1/r$ potential.

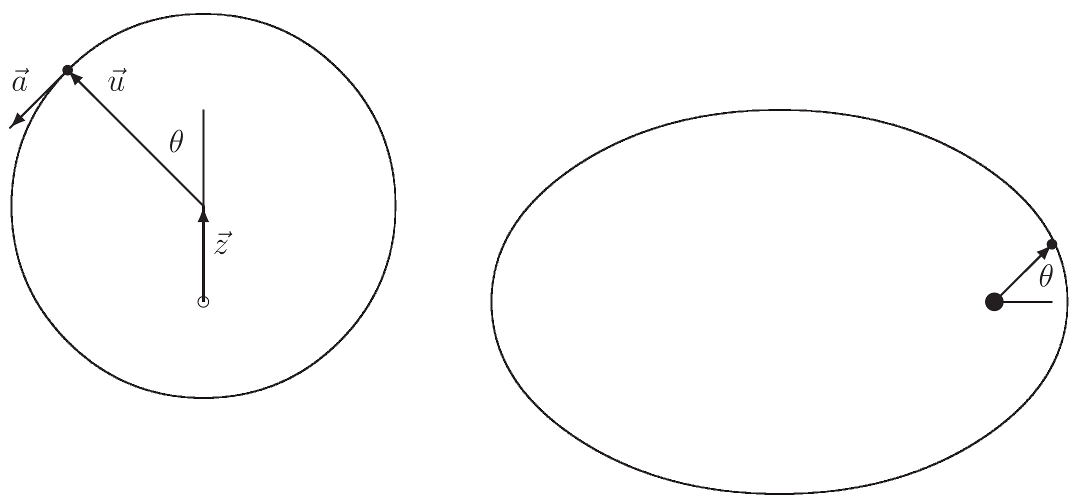
Now you can make a remarkable observation about the velocities that the planet has as it goes around its orbit. Imagine dividing up the whole orbit into segments of equal angle $d\theta$. In each of these segments, the magnitude of the impulse, or change in velocity, will be the same — the argument above just depended on the observation that the time to traverse the segment was proportional to the area of the segment, which is proportional to r^2. At the same time, the acceleration, or change in velocity per unit time, is proportional to $1/r^2$. In each successive segment, the direction of the impulse changes by $d\theta$. Now, this impulse is just m times the change in velocity. So let's integrate the impulse to see what the velocity does. You start at some velocity $\vec{v}_0$. Then, each time the orbit in position space passes through an angle increment $d\theta$, you add an impulse to the velocity vector. Each time, the direction is rotated by $d\theta$, and the magnitude of each little kick is the same. So the velocity vector just moves around in a circle starting at v_0. **Be careful:** we didn't say that the velocity vector moves around this circle at a constant speed — it doesn't. Remember that the time required to sweep out each of these little $d\theta$ wedges is proportional to r^2.

This next figure illustrates the velocity space and position space orbits. I have made a choice here — I oriented the position space orbit with its closest approach to the sun along the $+x$ axis. Clearly, at this point its velocity is in the $+\hat{y}$ direction, and this is the point where it has its maximum velocity. Therefore, the center of the velocity space orbit must be along the $+y$ axis. (The dot in the postion space orbit represents the position of the sun. The small circle in the velocity space orbit is at zero velocity.)



velocity space orbit position space orbit

Now let's put some labels in the velocity space orbit. Let $\vec{z}$ be the displacement of the center of this orbit from the origin, and $\vec{u}$ be the vector from the center to some point of the orbit. Let θ be the angle between these two vectors.

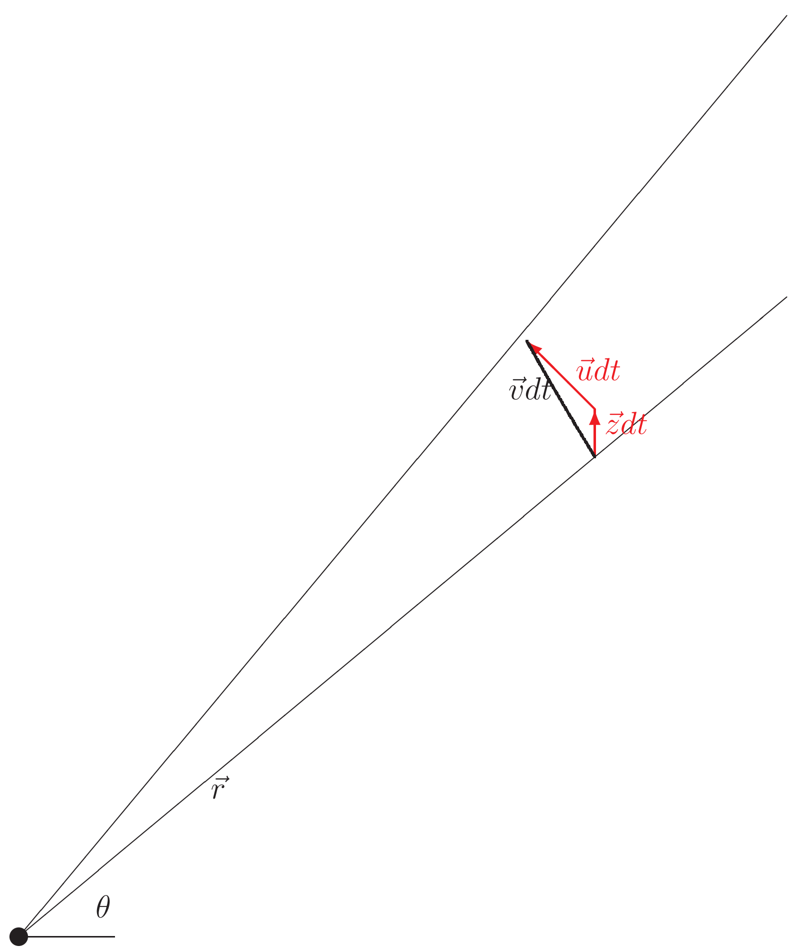


velocity space orbit position space orbit

Now, and this takes a moment's thought, $\vec{u}$ is perpendicular to the change in velocity, since the velocity is moving around the velocity space circle. In other words, $\vec{u}$ is perpendicular to the acceleration $\vec{a}$. But the acceleration is in the radial direction. So the point in the position space orbit corresponding to the marked point in the velocity space orbit is at a displacement perpendicular to $\vec{u}$. In other words, it is at an angle θ relative to the x axis in position space, when the corresponding point was at an angle θ relative to the y axis in velocity space.

1.2 Orbits in $1/r$ potentials are ellipses

Now we are finally ready to derive an equation for the shape of the orbit. Consider a small segment of the position space orbit, which is traversed in time dt. Magnify a portion of the position space orbit in the preceeding picture.



In this picture I show the segment of the orbit, which is the vector $\vec{v}\,dt$. From the velocity space diagram, this is equal to $\vec{z}\,dt + \vec{u}\,dt$. Remember that $\vec{u}$ is perpendicular to $\vec{r}$. Also remember that θ is both the polar angle in the position space orbit, and the angle between $\vec{z}$ and $\vec{u}$ in the velocity space orbit. Now, the area of this wedge is proportional to the component of $\vec{v}\,dt$ that is perpendicular to $\vec{r}$. With $\vec{u} \perp \vec{r}$ and θ

the angle between $\vec{u}$ and $\vec{z}$, this is just $dt\,(u + z\cos(\theta))$.

$$\text{area} = \frac{1}{2} r\,(u + z\cos(\theta))\,dt \tag{1.9}$$

(Again u is the magnitude of $\vec{u}$, etc.) But the area swept out is also just $L\,dt/2m$, and L is constant. So we must have

$$r\,(u + z\cos(\theta)) = \text{constant} \tag{1.10}$$

which is the equation for an ellipse in polar coordinates.

1.3 Conserved quantities

An excellent way to check whether your program is working, and whether your step size is small enough, is to monitor the conserved quantities of the orbit, namely energy and angular momentum.

For the energy, you will need both the potential and kinetic energies. Notice that the potential energy is negative.

For angular momentum, life is simple in two dimensions. In general, $\vec{L} = \vec{r} \times \vec{p}$, where $\vec{r}$ is the position and $\vec{p}$ is the (linear) momentum. Since $\vec{L}$ is perpendicular to both $\vec{r}$ and $\vec{p}$, if $\vec{r}$ and $\vec{p}$ are in the xy plane, $\vec{L}$ must be in the $\hat{z}$ direction. So, in two dimensions, it is effectively a scalar. From the formula for the z component of the cross product in Cartesian coordinates,

$$L = m(x\,v_y - y\,v_x) = constant \tag{1.11}$$

1.3.1 Choosing a good time step

If you were to make a first guess about an appropriate time step for integrating the Keplerian orbit, you would probably say that the time step should be a small fraction of the orbital period. And indeed this would work for a circular orbit.

However, when one has a very eccentric orbit, the planet spends more of the time moving slowly far from the Sun, and then briefly zooms in to small distance at high speed. One needs a much smaller time step to track that motion, something comparable to the time step you would use for a circular orbit with the same perihelion distance! This can be far smaller than the time step you would derive based on the orbital period, and the object will crawl around the outer parts of its orbit with a time step that is much smaller than needed.

Another way to assess your time step is by monitoring another conserved quantity, namely the orientation angle of the major axis of the elliptical orbit. This is called the angle of perihelion. Note that the eccentricity itself is determined by the energy and angular momentum, but the angle is not (a rotated ellipse has the same energy and angular momentum). You can detect that the object is at perihelion by looking to see when the radial velocity goes from being negative to positive.

The clever solution to this is to change the time step as we go around the orbit. This is called an "adaptive step size" and is briefly described in the supplementary material. For this homework, you will use a constant step size, but you'll have to inspect the solutions to ensure that you are using a small enough step for the most difficult parts of the orbit!

1.4 Orbital Mechanics

Following are listed several useful equations pertaining to orbital motion. We will not derive these equations, but you can find their derivations in your favorite undergraduate mechanics book.

- The perihelion distance, r_p is the distance of closest approach to the Sun in an elliptical orbit.
- The aphelion distance, r_a, is the furthest distance from the Sun in an elliptical orbit.
- The semi-major axis of orbit a is $2a = r_a + r_p$.
- The eccentricity e is the measure of non-circularity of an orbit.
  - $e < 1$: ellipse
  - $e = 0$: circle (special case of an ellipse)
  - $e = 1$: parabola
  - $e > 1$: hyperbola
- The relation between perihelion and aphelion distance is related to the eccentricity via:

  $$r_a = r_p \frac{1+e}{1-e}, \quad r_a = a(1+e), \quad r_o = a(1-e)$$

- For a circular orbit, the velocity is $v_c = \sqrt{\frac{GM}{a}}$, where M is the mass of the Sun.
- The velocity at perihelion for an orbit with given eccentricity is $v_p = \sqrt{\frac{GM(1+e)}{r_p}}$
- Kepler's laws say that the period of an orbit is independent of the eccentricity and depends only on the semi-major axis and the mass of the Sun. Using a circular orbit, you can quickly show that the orbital period is $T = 2\pi a^{3/2}/\sqrt{GM}$.

- The semi-major axis and eccentricity do determine the energy and angular momentum of the orbit. The relations are

$$a = \frac{GMm}{-2E}, \tag{1.12}$$

where M is the mass of the Sun and m is the (much smaller) mass of the object, and

$$e = \sqrt{1 + \frac{2EL^2}{(GMm)^2 m}} \tag{1.13}$$

1.5 A note about units

In principle you can work a problem like this week's homework in any units you want: meters and seconds, miles and months, furlongs and fortnights, or whatever. Sometimes, though, it is easiest to take a set of units particularly appropriate for your problem, and this week's assignment is an example. We would like to study the orbit of the earth around the sun, or variations on this orbit. As most of you know, the earth travels around the sun once per year. This suggests that a **year** is a reasonable unit of time. What about a unit of length? In the same spirit, we can use an *astronomical unit*, or "AU", which is the average distance from the earth to the sun.

Practice questions: These questions **must** be done before you start coding.

1. Neglecting the eccentricity of the earth's orbit (that is, assume it is circular), what is the speed of the earth in A.U./year? (This is important, because in your simulation this will be the initial velocity needed to get a circular orbit, if you do things right.

2. Newton's laws (both of motion and gravity — Newton did almost everything here) tell us that the acceleration of the earth is $a = GM/r^2$, where r is the earth-sun distance, M is the mass of the sun, and G is the gravitational constant. What is GM in $\text{AU}^3/\text{year}^2$?
   Don't go plugging in all the constants; just figure out what the acceleration is for an earth going around once per year. For an object in circular motion, the acceleration is $\omega^2 r$, where ω is the angular velocity. If the period is one year, the $\omega = 2\pi$ radians per year. So if $r = 1$ AU and the acceleration is $4\pi^2$, the $GM =?$.

1.6 A note about graphing ellipses

In this weeks assignment you are asked to make plots of ellipses and circles. These won't look right unless the scales on the `x` and `y` axes are the same. Therefore you

will have to specify the `x` and `y` ranges for the `graph` command:
`graph -T X -x -6 2 y -4 4 ...`
In this example, both axes cover a range of eight. In using Philsplot, make sure that when you draw the box the two ranges are the same.

1.7 Components of a vector

In this week's assignment, when you are writing equations of motion you will need to pick out components of a vector. Please see the "Coding Notes" in last week's class notes for simple ways to do this.

In your code, you will probably need a variable, perhaps "`r`", for the magnitude of the radius vector. Note that this also means you will need an "`r_mid`" at the midpoint of the time interval in your Runge-Kutta algorithm. Of course, `r_mid` can't be computed until both components of the position at the midpoint of the time interval have been computed.

1.8 Relation between period and radius

One of your homework problems asks you to verify that the periods of your orbits have the expected dependence on the radius. Just to be clear, the expected dependence is one of Kepler's laws: the period is proportional to the three halves power of the radius, or $T = CR^{3/2}$. (For an ellipse, replace "radius" by "semimajor axis".)

How do you test this? Well, the slope of a log-log plot would do the trick. Or you could calculate $T/R^{3/2}$ and see if it is a constant.

1.9 Repeated nagging about testing your program

A hint about testing your programs — before you turn the thing loose to run for a full orbit, take a careful look at the first step. Print out everything during and at the end of the step, and see if it makes sense. You can put a "`exit(0);`" command in your C code if you want it to just stop at a particular place.

1.10 Supplement: Adaptive Step-size Control

It would be useful if we could at least estimate the error we are making while numerically integrating an ODE. One way of doing this is to compare the results of integrations using two different step sizes. For example, using the fourth-order Runge-Kutta scheme above, if we take a single step of length $2h$, we will have the solution

$$\phi(x + 2h) = y_2 + (2h)^5 K f^{(5)}(\xi) + O(h^6), \tag{1.14}$$

where $\phi(x)$ is the exact solution (which doesn't depend on the step-size!), y_2 is the estimated value of the solution after one $2h$-long step, and K is a numerical constant. If we take two successive steps of length h, we have

$$\phi(x+2h) = y_{1+1} + 2(h)^5 K f^{(5)}(\xi) + O(h^6), \tag{1.15}$$

where y_{1+1} again is the estimated value, this time from the two-step solution. To $O(h^5)$, the $Kf^{(5)}(\xi)$ will be the same in both steps so the difference in the solutions at the end of the interval, $\Delta = y_2 - y_{1+1}$ will be a measure of the truncation error we have introduced.

Let us say that we will tolerate an error of δ. What step-size h_{new} would have given us that error? Since we expect the error to scale as h^5, we can write

$$h_{\text{new}} = h \left| \frac{\delta}{\Delta} \right|^{1/5}. \tag{1.16}$$

If the error we have just made is smaller than that we desire, $|\Delta| < |\delta|$, we can accept this solution and safely increase the step-size for the next step according to this equation. If our solution was not sufficiently accurate ($|\Delta| > |\delta|$), however, we had best throw it out and begin the step again, again with the (now smaller) step-size h_{new}. Such a monitoring of the solution and changing the step-size accordingly is known as *adaptive step-size control* and is a powerful technique for minimizing the computational effort for a given required accuracy.

The trouble with this is that it requires doing *two* simultaneous solutions to obtain Δ. A clever trick is to employ two difference schemes of differing orders to estimate Δ. The Runge-Kutta method lends itself well to such a procedure since it is possible to use the same intermediate values (similar to the k_i in equation **??** above) in a fourth- and a fifth-order method. The difference between these two solutions then gives the measure of the truncation error we need.

The design of such a system is beyond the level of this course, and implementing it in a program, while not difficult, is somewhat complex. See *Numerical Recipes*, section 16.2, for an example which is quite useful in practice (as well as a much more thorough discussion of integrating ODEs in general). We will be happy to provide this code for anyone who is interested. It is a perfect example of how, once one knows the basics, one can make use of a repository of already-written routines (in this case the library which comes with the book).