The Kepler Problem & Root Finding

Eric Angle

August 15, 2013

1 The Two Body Problem

The classical lagrangian L for a system of two particles with masses m_1, m_2 and positions $\mathbf{r}_1, \mathbf{r}_2$ interacting through a central potential $U(|\mathbf{r}_2 - \mathbf{r}_1|)$ is

$$L = \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{r}}_2^2 - U(|\mathbf{r}_2 - \mathbf{r}_1|).$$

Expressed in terms of the relative position $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and the center of mass $\mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/M$, where the total mass $M = m_1 + m_2$, this becomes

$$L = \frac{1}{2}\mu\dot{\mathbf{r}}^2 + \frac{1}{2}M\dot{\mathbf{R}}^2 - U(r),$$

with the reduced mass $\mu = m_1 m_2/M$. Throughout, we will work in the center of mass frame, where $\mathbf{R} = \mathbf{0}$, and

$$L = \frac{1}{2}\mu\dot{\mathbf{r}}^2 - U(r),$$

which yields the equation of motion

$$\frac{d}{dt}(\mu\dot{\mathbf{r}}) = \frac{dU}{dr}\hat{\mathbf{r}} \quad \Rightarrow \quad \ddot{\mathbf{r}} = \frac{1}{\mu}\frac{dU}{dr}\hat{\mathbf{r}}.$$

Using the equation of motion,

$$\frac{d}{dt}(\mathbf{r}\times\dot{\mathbf{r}}) = \mathbf{r}\times\ddot{\mathbf{r}} = 0,$$

which is simply angular momentum conservation, with the angular momentum $\mathbf{G} = \mu\mathbf{r}\times\dot{\mathbf{r}}$. Since $\mathbf{r}\bullet\mathbf{G} = 0$, and $\mathbf{G}$ is constant, $\mathbf{r}$ must lie in the plane perpendicular to $\mathbf{G}$. Therefore, we can express $\mathbf{r}$ in two dimensions−we choose polar coordinates (r, ϕ)−and the Lagrangian becomes

$$L = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - U(r),$$

with corresponding conserved $(\partial L/\partial t = 0)$ energy (the hamiltonian is equal to the energy in this case)

$$E = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\phi}^2\right) + U(r).$$

We can eliminate $\dot{\phi}$ in place of r using angular momentum conservation, since[1]

$$\mathbf{G} = \mu r\hat{\mathbf{r}}\times\left(\dot{r}\hat{\mathbf{r}} + r\dot{\phi}\hat{\boldsymbol{\phi}}\right) = \mu r^2\dot{\phi}\hat{\mathbf{z}} = G\hat{\mathbf{z}} \quad \Rightarrow \quad \dot{\phi} = \frac{G}{\mu r^2}, \tag{1}$$

which gives

$$E = \frac{1}{2}\mu\dot{r}^2 + \frac{G^2}{2\mu r^2} + U(r).$$

Solving for $\dot{r}$, we obtain

$$\dot{r} = \pm\left\{\frac{2}{\mu}[E - U(r)] - \frac{G^2}{\mu^2 r^2}\right\}^{1/2},$$

or

$$dt = \pm dr\left\{\frac{2}{\mu}[E - U(r)] - \frac{G^2}{\mu^2 r^2}\right\}^{-1/2} \quad \Rightarrow \quad t(r) - t_0 = \int dr\left\{\frac{2}{\mu}[E - U(r)] - \frac{G^2}{\mu^2 r^2}\right\}^{-1/2}, \tag{2}$$

where we've chosen the positive sign, which amounts to taking t positive since the integrand is positive. Equation **??** gives a relationship between r and t. To obtain a corresponding relationship between ϕ and r, we can use $d\phi = Ldt/\mu r^2$:

$$d\phi = \pm dr\frac{G}{r^2}\left\{2\mu[E - U(r)] - \frac{G^2}{r^2}\right\}^{-1/2} \quad \Rightarrow \quad \phi(r) - \phi_0 = \int dr\frac{G}{r^2}\left\{2\mu[E - U(r)] - \frac{G^2}{r^2}\right\}^{-1/2}, \tag{3}$$

where we've chosen the positive sign, which amounts to taking ϕ positive since the integrand is positive. Equations **??** and **??** give the general solution to the two body problem.

[1]Alternatively, $0 = \partial L/\partial\phi = \dot{p_\phi} \Rightarrow p_\phi = \partial L/\partial\dot{\phi} = \mu r^2\dot{\phi} \Rightarrow \dot{\phi} = p_\phi/\mu r^2$.

2 The Kepler Problem

The two body problem for the case $U \propto 1/r$ is Kepler's problem. We'll consider only an attractive potential

$$U(r) = -\frac{k}{r}, \tag{4}$$

with $k > 0$, applicable to the gravitational and electrostatic (for opposite charges) interactions. Substituting this potential into equation **??** yields

$$\phi = \int dr \frac{G}{r^2}\left(2\mu E + \frac{2\mu k}{r} - \frac{G^2}{r^2}\right)^{-1/2} = \cos^{-1}\left(\frac{G/r - \mu k/G}{\sqrt{2\mu E + \mu^2 k^2/G^2}}\right), \tag{5}$$

or

$$\frac{p}{r} = 1 + \epsilon \cos\phi, \tag{6}$$

where $p = G^2/\mu k$, $\epsilon = \sqrt{1 + 2EG^2/\mu k^2}$, and we've taken $\phi_0 = 0$. Equation **??** describes a conic section with one focus at the origin (the center of mass), latus rectum $2p$, eccentricity ϵ, and perihelion (point on the orbit nearest to the origin) at $\phi = 0$.

We'll specialize to bound states, $E < 0$, whose orbits are ellipses ($0 \le \epsilon < 1$) with major axis

$$a = \frac{p}{1-\epsilon^2} = \frac{k}{2|E|}$$

and minor axis

$$b = \frac{p}{\sqrt{1-\epsilon^2}} = \frac{G}{\sqrt{2\mu|E|}}.$$

For a segment of the path with angle $d\phi$, an area $dA = r^2 d\phi/2$ is swept out. Using equation **??**, this becomes $dA = G dt/2\mu$, or $dt = 2\mu dA/G$, which we can integrate to give the period T of the orbit:

$$T = \frac{2\mu A}{G} = \pi k \sqrt{\frac{\mu}{2|E|^3}} = 2\pi\sqrt{\frac{\mu a^3}{k}},$$

where we used $A = \pi ab$ for an ellipse.

Substituting the potential **??** into equation **??** yields

$$t = \int dr \left(-\frac{2|E|}{\mu} + \frac{2k}{\mu r} - \frac{G^2}{\mu^2 r^2}\right)^{-1/2} = \frac{T}{2\pi}(x - \epsilon \sin x),$$

or

$$y = x - \epsilon \sin x, \tag{7}$$

where

$$r = a(1 - \epsilon\cos x), \tag{8}$$

x is the *eccentric anomaly*[2], $y = 2\pi t/T$ is the *mean anomaly*[3], and we've taken $t_0 = 0$, so that the perihelion is at $t = 0$.

The solution of Kepler's problem for bound states in an attractive potential is thus reduced to solving Kepler's equation **??**, whose solution x for a value of y gives r (equation **??**) and ϕ (equation **??**) as a function of time ($t = yT/2\pi$).

3 Kepler's Equation

Since $r(nT - t) = r(t)$ and $\phi(nT - t) = 2\pi n - \phi(t)$, with n integer, we need only determine the system's time evolution from $t = 0$ to $T/2$, a half period. At $t = 0$, $x = y = 0$, and at $t = T/2$, $x = y = \pi$, so we may focus on solving Kepler's equation for y ranging linearly from 0 to π, with x in the range 0 to π for each y value. The following sections aim to numerically solve Kepler's equation.

[2] `http://en.wikipedia.org/wiki/Eccentric_anomaly`
[3] `http://en.wikipedia.org/wiki/Mean_anomaly`

3.1 Iterative Method

Kepler's equation may be rewritten

$$x = y + \epsilon \sin x,$$

which motivates us to employ an iterative scheme

$$x_{n+1} = g\left(x_n\right), \tag{9}$$

where

$$g\left(x_n\right) = y + \epsilon \sin x_n \tag{10}$$

and $x_0 = y$.

For a general iterative scheme **??**, the error in x_n, defined as $\Delta_n = x^* - x_n$, where $x^* = g\left(x^*\right)$, is:

$$x^* - \Delta_{n+1} = g\left(x^* - \Delta_n\right) = x^* - g'\left(x^*\right)\Delta_n + \frac{1}{2}g''\left(x^*\right)\Delta_n^2 + O\left(\Delta_n^3\right),$$

which simplifies to

$$\Delta_{n+1} = g'\left(x^*\right)\Delta_n - \frac{1}{2}g''\left(x^*\right)\Delta_n^2 + O\left(\Delta_n^3\right). \tag{11}$$

Neglecting $O\left(\Delta_n^2\right)$ terms,

$$\Delta_{n+1} \approx g'\left(x^*\right)\Delta_n \approx \left[g'\left(x^*\right)\right]^{n+1}\Delta_0,$$

which shows that the iterative scheme **??** converges if $\left|g'\left(x^*\right)\right| < 1$ and diverges if $\left|g'\left(x^*\right)\right| > 1$. If $\left|g'\left(x^*\right)\right| = 0$ or 1, $O\left(\Delta_n^2\right)$ terms must be considered.

For Kepler's equation, $g\left(x_n\right)$ is given by equation **??**, and

$$g'\left(x_n\right) = \epsilon \cos x_n.$$

Since $\left|g'\left(x^*\right)\right| = |\epsilon|\left|\cos x^*\right| \leq |\epsilon| < 1$, the iteration **??** will theoretically converge.
IMPLEMENTATION

3.2 Aitken's Acceleration Method

If an iteration scheme is converging, $\Delta_{n+1} = C_n\Delta_n$, where $|C_n| < 1$. Near convergence, C will be approximately constant, so that $\Delta_{n+1} \approx C\Delta_n$. Eliminating C yields

$$\frac{\Delta_{n+1}}{\Delta_n} \approx \frac{\Delta_{n+2}}{\Delta_{n+1}}.$$

Solving for x^* yields

$$x^* \approx \frac{x_n x_{n+2} - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n}. \tag{12}$$

Replacing x_n by x^*, given by equation **??**, can accelerate the convergence of some iterative schemes.
Convergence.
Implementation

3.3 Root-Finding Methods

The problem of solving Kepler's equation may be recast into one of finding solutions x^* of $f\left(x^*\right) = 0$, where

$$f(x) = x - \epsilon \sin x - y.$$

Note that f has a single root, since its derivative $f'(x) = 1 - \epsilon \cos x > 0$ for $\epsilon < 1$. The following are a few methods of root-finding.

3.3.1 Interval Bisection Method

A simple root-finding method is bisection, which after developing lower and upper bounds x_L and x_U, respectively, for a root of a function f, bisects the interval at $x_M = (x_L + x_U)/2$ and determines in which half of the interval the root must lie. If $f(x_L) f(x_M) \leq 0$, the root is between x_L and x_M; otherwise, it's between x_M and x_U. In the former case, the method takes $x_R \to x_M$, and in the latter case, $x_L \to x_M$, and the process iterates. The convergence of the bisection method is given simply by

$$\Delta_{n+1} = \frac{1}{2}\Delta_n.$$

If a root is desired to a tolerance δ, then the number of necessary iterations n may be bounded:

$$\delta > \Delta_n = \frac{1}{2^{n-1}}\Delta_1 = \frac{x_U - x_L}{2^{n-1}} \quad \Rightarrow \quad n > \log_2\left(\frac{x_U - x_L}{\delta}\right).$$

For solving Kepler's equation, we choose $x_U = \pi$, $x_L = 0$, and $\delta = 5 \times 10^{-15}$, so that $n > 49$. In running kepler.f, we find $n = 50 > 49$.

3.3.2 Newton's Method

Newton's method approximates the roots of a function f by linearly expanding it about x_n, evaluating the expansion at x_{n+1}, and setting the result to zero,

$$0 = f(x_{n+1}) \approx f(x_n) + (x_{n+1} - x_n) f'(x_n) \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \tag{13}$$

so that the iterative equation **??** takes the specific form

$$g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{14}$$

For Kepler's equation,

$$g(x_n) = x_n + \frac{y - x_n + \epsilon \sin x_n}{1 - \epsilon \cos x_n}.$$

To analyze Newton's method's convergence, the first derivative of g is

$$g' = \frac{f f''}{f'^2},$$

which vanishes at $x = x^*$ since $f(x^*) = 0$. The second derivative is

$$g'' = \frac{f'^2 f'' + f f' f''' - 2 f f''^2}{f'^3}.$$

Evaluated at $x = x^*$, this becomes

$$g''(x^*) = \frac{f''(x^*)}{f'(x^*)},$$

again since $f(x^*) = 0$. Referring to equation **??**,

$$\Delta_{n+1} \approx \left[-\frac{f''(x^*)}{2 f'(x^*)}\right] \Delta_n^2.$$

For Kepler's equation,

$$\Delta_{n+1} \approx \frac{1}{2}\left(\cot x^* - \csc x^*/\epsilon\right)^{-1} \Delta_n^2.$$

3.3.3 Secant Method

If the derivative of f is unavailable or cumbersome, one can replace the derivative in **??** with a secant line approximation, which depends only on the function itself:

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \quad \Rightarrow \quad x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n).$$

This iteration scheme depends on the two previous values and not just one:

$$x_{n+1} = h(x_{n-1}, x_n),$$

where

$$h(x_{n-1}, x_n) = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n).$$

For Kepler's equation,

$$h(x_{n-1}, x_n) = x_n - \left[1 + \frac{\epsilon(\sin x_{n+1} - \sin x_n)}{x_n - x_{n+1}}\right]^{-1} (x_n - \epsilon \sin x_n - y).$$