Perihelion of Mercury

Paul Lee

Student Number: 990209662
Course: General Relativity
(APM426S)
Date: 21 March 2002

1. Introduction

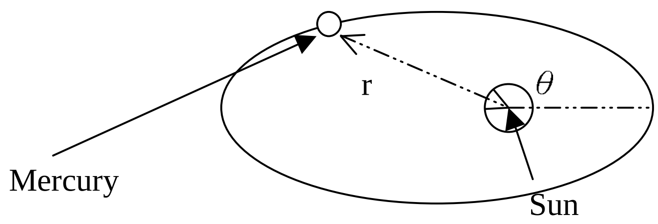
At present, there are three experimentally measurable tests on General Theory of Relativity. They are the red shift, the deflection of starlight passing the sun and the perihelion precession of Mercury. Among all the three tests, perihelion precession of Mercury is the most important one. The reason is that the explanation on red shift does not require the use of Einstein's equation. It can be done by using only conservation of energy and principle of equivalence. Similarly, the same problem for deflection of starlight passing the sun also arises. It can be explained using only special theory of relativity, principle of equivalence and classical optics. Moreover, the measurements do not agree with the one predicted by the theory. On the other hand, the same problems do not appear in perihelion precession of Mercury. As we will see later, the derivation of it requires the use of geodesic equations, which is closely related to the Einstein's equation. Also, the result from theory is in excellent agreement with the measurements. So, the aim of this paper is to look at the relativistic explanation on perihelion precession of Mercury and the failure of a classical explanation of the shift.

In this paper, we will presuppose the knowledge of differential geometry and general relativity given by the first four chapters in the book General Relativity by Robert M. Wald. This includes some basic in tensor calculus and Riemann geometry, special theory of relativity and some basics in general theory of relativity including the Einstein's equation.

We will begin our discussion by recalling the derivation of the classical Kepler problem and we will see that the two differential equations derived by the classical theory and relativistic theory are very closely related. Next, we will derive the Schwarzschild solution, which is a time-independent and radially symmetric metric for the free-space field equations. Then, we will use this solution to solve our relativistic Kepler problem and explain why there is a shift in perihelion of Mercury. In the last part, we will look at one classical explanation on this problem and its failure.

2. Classical Kepler problem

In this section, we will recall the classical solution to the Kepler problem. The problem is to prove that orbits of planets are ellipses and the Sun is located at one of the foci of the ellipse. This is also known as the Kepler's First Law.



Let us assume that the motion takes place in a plane and recall the well-known fact in Physics – Conservation of Energy. It says that if there is no net force acting on the system, then the total energy is constant. So,

$$E = \frac{1}{2}m\|v\|^2 + V(r) \tag{2.1}$$

where E is the energy of the planet which is constant, $v = (v_1, v_2, 0)$ is the velocity of the planet, m is the mass of the planet and V is the gravitational potential energy which depends only on r since we assume that the Sun is a sphere.

If we let $x = (x_1, x_2, 0)$ be the position vector, then $v = \dot{x} = (\frac{dx_1}{dt}, \frac{dx_2}{dt}, 0)$. So we can write (2.2) in polar coordinate by letting $x = (r\cos\theta, r\sin\theta, 0) = r(\cos\theta, \sin\theta, 0)$.

By Leibnitz rule, we have

$v = \dot{x} = \dot{r}(\cos\theta, \sin\theta, 0) + r(-\sin\theta, \cos\theta, 0)\dot{\theta} = (\dot{r}\cos\theta - r\dot{\theta}\sin\theta, \dot{r}\sin\theta + r\dot{\theta}\cos\theta, 0)$.

So, $$\|v\|^2 = (\dot{r}\cos\theta - r\dot{\theta}\sin\theta)^2 + (\dot{r}\sin\theta + r\dot{\theta}\cos\theta)^2 = \dot{r}^2 + r^2\dot{\theta}^2 \tag{2.2}$$

Substitute (2.2) into (2.1), we get

$$E = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) \tag{2.3}$$

Another fact that comes from Physics is Conservation of Angular Momentum. Let us look at this in more detail. Angular momentum is by definition

$$L = x \times p \tag{2.4}$$

where $p = mv$ is the linear momentum. Conservation of Angular Momentum says that if there is no net torque acting on the system, then angular momentum is constant. And torque is defined by $\tau = x \times F$, where F is the force vector. To prove this, we differentiate (2.4) with respect to time and use Leibnitz rule.

$$\frac{dL}{dt} = \dot{x} \times p + x \times \dot{p} \tag{2.5}$$

Since $\dot{x} = v$, $\dot{x} \times p = v \times mv = 0$. (2.5) becomes

$$\frac{dL}{dt} = x \times \dot{p} \tag{2.6}$$

But $\dot{p} = m\dot{v} = ma = F$, so (2.6) is just

$$\frac{dL}{dt} = x \times F = \tau = 0 \tag{2.7}$$

Here, $\tau = 0$ since we assume there is no net torque to the system. Therefore, angular momentum is conserved. Let us go back to the Solar system. In there, the angular momentum is conserved since there is no force that is not parallel to x. Next, let us rewrite (2.4) in a more familiar form. To do this, we apply polar coordinate.

$$L = r(\cos\theta, \sin\theta, 0) \times [m\dot{r}(\cos\theta, \sin\theta, 0) + mr(-\sin\theta, \cos\theta, 0)\dot{\theta}]$$
$$\Rightarrow L = r(\cos\theta, \sin\theta, 0) \times mr(-\sin\theta, \cos\theta, 0)\dot{\theta}$$
$$\Rightarrow L = mr^2\dot{\theta}(0,0,1)$$

So the magnitude of angular momentum is

$$l = mr^2\dot{\theta} \tag{2.8}$$

which is the condition for central force field.

Now let us go back to our discussion on the Kepler problem. If we differentiate (2.3) with respect to time, we get

$$0 = m\dot{r}\ddot{r} + mr\dot{r}\dot{\theta}^2 + mr^2\dot{\theta}\ddot{\theta} + V'(r)\dot{r} \tag{2.9}$$

Differentiate (2.8) with respect to time, it becomes

$$0 = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} \tag{2.10}$$

Substitute (2.10) into (2.9), we have

$$0 = m\dot{r}\ddot{r} + mr\dot{r}\dot{\theta}^2 - 2mr\dot{r}\dot{\theta}^2 + V'(r)\dot{r}$$
$$\Rightarrow m\dot{r}\ddot{r} - mr\dot{r}\dot{\theta}^2 + V'(r)\dot{r} = 0 \tag{2.11}$$

One possible solution to (2.11) is $\dot{r} = 0$ which corresponds to a circular orbit. So, to obtain the other possible solution, we cancel $\dot{r}$ in (2.11) and get

$$\ddot{r} - r\dot{\theta}^2 + \frac{V'(r)}{m} = 0 \tag{2.12}$$

To solve (2.12), we let

$$u(\theta) = \frac{1}{r(\theta)} \tag{2.13}$$

By chain rule and (2.8), $\dot{r}$ becomes

$$\dot{r} = \frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt} = r'\dot{\theta} = \frac{lr'}{mr^2} = -\frac{l}{m}u' \tag{2.14}$$

Differentiate (2.14) with respect to time and use (2.8) again, we get

$$\ddot{r} = \frac{d}{dt}\dot{r} = \frac{d\theta}{dt}\frac{d}{d\theta}(\dot{r}) = \frac{d\theta}{dt}\frac{d}{d\theta}\left(-\frac{l}{m}u'\right) = -\frac{l^2}{m^2r^2}u'' = -\frac{l^2}{m^2}u''u^2 \tag{2.15}$$

Substitute (2.8), (2.13) and (2.15) into (2.12), we have

$$-\frac{l^2}{m^2}u''u^2-\frac{l^2}{m^2}u^3+\frac{V'(r)}{m}=0$$

$$\Rightarrow u''+u=\frac{V'(r)m}{l^2u^2} \tag{2.16}$$

From Physics, we also know that $V(r)=-\frac{GMm}{r}$ for any gravitational field generated by a spherical body, where $G=6.67\times10^{-11}Nm^2/kg^2$ is called the gravitational constant and M is the mass of the Sun. So (2.16) becomes

$$u''+u=\frac{GMm^2}{l^2} \tag{2.17}$$

The right side of (2.17) is just a constant, so we can solve this ODE. Clearly, $u_p=\frac{GMm^2}{l^2}$ is a particular solution of (2.17) and $u_h=C\cos(\theta+\delta)$ is the general solution of $u''+u=0$ where C and δ are arbitrary constants. Therefore, the general solution of (2.17) is

$$u=\frac{GMm^2}{l^2}+C\cos(\theta+\delta) \tag{2.18}$$

We orient the system so that $\delta=0$, so (2.18) becomes

$$u=\frac{GMm^2}{l^2}+C\cos\theta \quad \text{or} \quad r=\frac{\frac{l^2}{GMm^2}}{1+\frac{Cl^2}{GMm^2}\cos\theta} \tag{2.19}$$

which is the equation of ellipse in polar coordinates with one of its foci at the origin. So the Sun is located at the origin and the orbit is elliptical. Therefore, this solves the classical Kepler problem.

Of course, this is not the end of the story. In 1859, French Astronomist, Le Verrier, discovered that there is a perihelion advanced of Mercury by 38” per century. Le Verrier is the one who discovered Neptune by calculating the position of it from irregularities in Uranus’s orbit. He tried to apply the same theory to this perihelion shift of Mercury by hypothesizing a planet Vulcan between Mercury and the Sun, and he spent much time on looking for this planet. Obviously, he did not succeed and this perihelion shift is eventually explained by Einstein’s General Theory of Relativity.

3. The Schwarzschild Solution

Before discussing the perihelion shift, we have to derive the Schwarzschild solution. So, imagine we have an empty space, then the metric will be given by the space-time interval $ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)$. In polar coordinates, it is

$$ds^2 = c^2dt^2 - (dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2) \tag{3.1}$$

where $c = 3.0\times10^8 m/s$ is the speed of light.

Now imagine we put a massive body at the origin, then space-time will curve down like a trough (This is how Einstein think about it). The metric corresponds to this setting with the following assumptions is called the Schwarzschild Solution.

If we consider this metric as a "space-time trough", it should be almost flat if we go very far away from the central mass. So, as $r\to\infty$, we should get back (3.1) and this is our first assumption. Next, we will assume that the central mass is stationary, so the metric is not changing with time. This assumption helps us a lot in simplifying our metric since ds^2 should be invariant if we replace dt by $-dt$. So all terms involving dt is zero except dt^2. The third assumption is that the mass of the central body is evenly distributed. This means that the metric is radially symmetric. So, if we change $d\theta$ to $-d\theta$ or change $d\phi$ to $-d\phi$, ds^2 should again not change. Therefore, we expect a diagonal solution like the following one.

$$ds^2 = A'(r')c^2dt^2 - [B'(r')dr'^2 + C'(r')r'^2(d\theta^2 + \sin^2\theta d\phi^2)] \tag{3.2}$$

where $A'(r'), B'(r')$ and $C'(r')$ are some unknown functions.

To simplify further, we choose a new coordinate $r = \sqrt{C(r')}r'$, then

$dr = \left[\frac{1}{2\sqrt{C(r')}}\frac{dC}{dr'}r' + \sqrt{C(r')}\right]dr'$ and so (3.2) becomes

$$ds^2 = A(r)c^2dt^2 - [B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)] \tag{3.3}$$

where $A(r) = A'(r')$ and $B(r)dr^2 = B'(r')dr'^2 = B'(r')\left[\frac{1}{2\sqrt{C(r')}}\frac{dC}{dr'}r' + \sqrt{C(r')}\right]^{-2} dr^2$

Next, we want to calculate the Cristoffel Symbols. From (4.2), we know that

$$g_{00} = c^2A, g_{11} = -B, g_{22} = -r^2, g_{33} = -r^2\sin^2\theta \tag{3.4}$$

and all other components of the metric vanishes. The metric is diagonal and so the inverse is easy to calculate.

$$g^{00} = \frac{1}{c^2A}, g^{11} = -\frac{1}{B}, g^{22} = -\frac{1}{r^2}, g^{33} = -\frac{1}{r^2\sin^2\theta}. \tag{3.5}$$

Now, we know all the entries in the metric, so we can use (3.1.10 inWald)

$$\Gamma^{\rho}{}_{\mu\nu} = \frac{1}{2}\sum_{\sigma} g^{\rho\sigma}\left(\frac{\partial g_{\nu\sigma}}{\partial x^{\mu}} + \frac{\partial g_{\mu\sigma}}{\partial x^{\nu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}\right) \tag{3.6}$$

to calculate the Cristoffel Symbols. Keep in mind that A and B depends only on r and substitute $x^0 = t, x^1 = r, x^2 = \theta$ and $x^3 = \phi$ into (3.6). So, we get

$$\Gamma^{0}{}_{10} = \frac{1}{2}\sum_{\sigma} g^{00}\left(\frac{\partial g_{00}}{\partial r} + \frac{\partial g_{11}}{\partial t} - \frac{\partial g_{10}}{\partial x^{\sigma}}\right) = \frac{1}{2}\left(\frac{1}{c^2 A}\right)\left(c^2\frac{\partial A}{\partial r}\right) = \frac{A'}{2A}$$

Similarly,

$$\Gamma^{0}{}_{01} = \frac{A'}{2A} \qquad \Gamma^{1}{}_{00} = \frac{c^2 A'}{2B} \qquad \Gamma^{1}{}_{11} = \frac{B'}{2B}$$

$$\Gamma^{1}{}_{22} = \frac{-r}{B} \qquad \Gamma^{1}{}_{33} = \frac{-r\sin^2\theta}{B} \qquad \Gamma^{2}{}_{21} = \frac{1}{r}$$

$$\Gamma^{2}{}_{12} = \frac{1}{r} \qquad \Gamma^{2}{}_{33} = -\sin\theta\cos\theta \qquad \Gamma^{3}{}_{23} = \cot\theta \tag{3.7}$$

$$\Gamma^{3}{}_{32} = \cot\theta \qquad \Gamma^{3}{}_{13} = \frac{1}{r} \qquad \Gamma^{3}{}_{31} = \frac{1}{r}$$

For the empty space surrounding the body, $T_{\mu\nu} = 0$. So Einstein's equation becomes

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0 \tag{3.8}$$

Multiply (3.8) on both sides by $g^{\mu\nu}$ and contract, then (3.8) will give us $R = 0$. Therefore, if we substitute $R = 0$ into (3.8), then we have

$$R_{\mu\nu} = 0 \tag{3.9}$$

Now, let us compute R_{00}, R_{11} and R_{22}. From definition of Ricci tensor (3.2.25 in Wald), we have

$$R_{ac} = R_{abc}{}^{b} \tag{3.10}$$

Also, we have formula (3.4.4 in Wald),

$$R_{\mu\nu\rho}{}^{\sigma} = \frac{\partial}{\partial x^{\nu}}\Gamma^{\sigma}{}_{\mu\rho} - \frac{\partial}{\partial x^{\mu}}\Gamma^{\sigma}{}_{\nu\rho} + \sum_{\alpha}(\Gamma^{\alpha}{}_{\mu\rho}\Gamma^{\sigma}{}_{\alpha\nu} - \Gamma^{\sigma}{}_{\nu\rho}\Gamma^{\sigma}{}_{\alpha\mu}) \tag{3.11}$$

So, we get the following equations.

$$0 = R_{00} = R_{0\nu0}{}^{\nu} = \frac{\partial}{\partial x^{\nu}}\Gamma^{\nu}{}_{00} - \frac{\partial}{\partial x^{0}}\Gamma^{\nu}{}_{\nu0} + \Gamma^{\alpha}{}_{00}\Gamma^{\nu}{}_{\alpha\nu} - \Gamma^{\alpha}{}_{\nu0}\Gamma^{\nu}{}_{\alpha0}$$

$$= \frac{\partial}{\partial r}\left(\frac{c^2 A'}{2B}\right) + \frac{c^2 A'}{2B}\left(\frac{A'}{2A}\right) + \frac{c^2 A'}{2B}\left(\frac{B'}{2B}\right) + \frac{c^2 A'}{2B}\left(\frac{2}{r}\right) - 2\left(\frac{A'}{2A}\right)\left(\frac{c^2 A'}{2B}\right)$$

$$= \frac{c^2}{2B}\left(A'' - \frac{A'B'}{2B} - \frac{A'^2}{2A} + \frac{2A'}{r}\right)$$

$$\Rightarrow A''-\frac{A'B'}{2B}-\frac{A'^2}{2A}+\frac{2A'}{r}=0 \qquad (3.12)$$

Similarly, $0=R_{11}=\frac{1}{2A}\left(-A''+\frac{A'B'}{2B}+\frac{A'^2}{2A}+\frac{2AB'}{rB}\right)$

$$\Rightarrow -A''+\frac{A'B'}{2B}+\frac{A'^2}{2A}+\frac{2AB'}{rB}=0 \qquad (3.13)$$

$$0=R_{22}=R_{2\nu 2}{}^{\nu}=-\frac{\partial}{\partial\theta}\cot\theta-\frac{\partial}{\partial r}\left(\frac{r}{B}\right)+\frac{2}{B}-\cot^2\theta-\frac{r}{B}\left(\frac{2}{r}+\frac{(AB)'}{2AB}\right) \qquad (3.14)$$

By adding (3.12) and (3.13), we get $0=\frac{2A'}{r}+\frac{2AB'}{rB}=\frac{2}{rB}(AB)'$ and so $AB=\text{constant}$. To determine this constant, we use the first assumption that we will get back the space-time interval as $r\to\infty$. So, compare (3.1) and (3.2), we get $A'(r')=B'(r')=C'(r')=1$ as $r\to\infty$. So from (3.2), we have

$$A(r)=A'(r')=1 \text{ and } B(r)=B'(r')\left[\frac{1}{2\sqrt{C(r')}}\frac{dC}{dr'}r'+\sqrt{C(r')}\right]^{-2}=1$$

Therefore, the constant is 1 and we get

$$B=\frac{1}{A} \qquad (3.15)$$

Now, consider (3.14) together with $(AB)'=0$, then

$$0=\csc^2\theta-\frac{\partial}{\partial r}\left(\frac{r}{B}\right)-\cot^2\theta$$

$$\Rightarrow\frac{\partial}{\partial r}\left(\frac{r}{B}\right)=1 \qquad (3.16)$$

Integrate (3.16) on both sides, then $\frac{r}{B}=r+K$ for some constant K. So, together (3.15), we have

$$A=\frac{K+r}{r},\, B=\frac{r}{r+K} \qquad (3.17)$$

In order to determine what K is, we consider again the case where $r\to\infty$ and we put a test body at rest relative to the central mass M. We look at the geodesic equation given by (3.3.5 inWald).

$$\frac{dx^{\mu}}{dt^2}+\sum_{\sigma,\nu}\Gamma^{\mu}{}_{\sigma\nu}T^{\sigma}T^{\nu}=0 \qquad (3.18)$$

where $x^{\mu}(t)$ is the geodesic.

In this case, the test body will move along the geodesic. Also, the test body is at rest initially, $\frac{dr}{dt} = \frac{d\theta}{dt} = \frac{d\phi}{dt} = 0$ at $t = 0$. So if we parameterize the geodesic using time,

then (3.18) becomes $\left[\frac{d^2 r}{dt^2} + \Gamma^1{}_{00}\frac{dt}{dt}\frac{dt}{dt}\right]\Bigg|_{t=0} = \left[\frac{d^2 r}{dt^2} + \Gamma^1{}_{00}\right]\Bigg|_{t=0} = 0$.

So, by (4.8), we get

$$\left.\frac{d^2 r}{dt^2}\right|_{t=0} = -\Gamma^1{}_{00} = -\frac{c^2 A'}{2B} = -\frac{c^2}{2}\left(\frac{r+K}{r}\right)\frac{d}{dr}\left(\frac{K+r}{r}\right) = \frac{c^2 K}{2r^2} + \frac{c^2 K^2}{2r^3} \tag{3.19}$$

At $r \to \infty$, $1/r^3$ decreases much faster than r^2. So (3.19) becomes

$$\left.\frac{d^2 r}{dt^2}\right|_{t=0} = \frac{c^2 K}{2r^2} + \frac{c^2 K^3}{2r^3} \to \frac{c^2 K}{2r^2} \quad \text{as} \quad r \to \infty \tag{3.20}$$

The left hand side of (3.20) is acceleration of the test body. Since the relativistic effect at $r \to \infty$ is negligible, we should get back Newton's Law of Gravitation, which is $a = -\frac{GM}{r^2}$. So, by (3.20), we have $\frac{c^2 K}{2r^2} = -\frac{GM}{r^2}$, where M is the mass of the central body at the origin and $G = 6.67 \times 10^{-11}\, Nm^2/kg^2$. Therefore, the constant is

$$K = \frac{-2GM}{c^2}. \tag{3.21}$$

Now, rewrite (3.3) using (3.17) and (3.21), we get

$$ds^2 = \left(c^2 - \frac{2GM}{r}\right)dt^2 - r^2(\sin^2\theta d\phi^2 + d\theta^2) - \frac{dr^2}{\left(1 - \frac{2GM}{c^2 r}\right)} \tag{3.22}$$

And (3.22) is the Schwarzschild metric that we are looking for.

This solution to the field equations was found in 1916 by Karl Schwarzschild. It can be used to explain neutron stars, pulsars and black holes. Now, we are going to use it to account for the perihelion shift of Mercury. (Also, perihelion shift of Mercury had already been explained in Einstein's 1915 paper. But, I can't find a book which explain this without using Schwarzschild solution.)

4. Perihelic Shift of Mercury

Consider the setting in section 4. If we replace the central mass by the Sun and put the Planet Mercury into the system. Since mass of the Sun is much bigger than mass of Mercury, we can approximate the situation by assuming that Mercury has zero mass and so the Sun will not be affected by Mercury's gravitational field. Also, assume that the Sun is not rotating. This is reasonable since rotation of the Sun is small and the Sun is huge. By theory of General Relativity, we know that Mercury will travel along the geodesic. So, let us consider the geodesic equations (3.18) together with (3.7).

$$\frac{d^2\theta}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds}-\sin\theta\cos\theta\left(\frac{d\phi}{ds}\right)^2=0 \qquad (4.1)$$

$$\frac{d^2r}{ds^2}+\left(\frac{c^2A'}{2B}\right)\left(\frac{dt}{ds}\right)^2+\left(\frac{B'}{2B}\right)\left(\frac{dr}{ds}\right)^2-\left(\frac{r}{B}\right)\left(\frac{d\theta}{ds}\right)^2-\left(\frac{r\sin^2\theta}{B}\right)\left(\frac{d\phi}{ds}\right)^2=0 \qquad (4.2)$$

$$\frac{d^2\phi}{ds^2}+2\cot\theta\frac{d\theta}{ds}\frac{d\phi}{ds}+\frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds} \qquad (4.3)$$

$$\frac{d^2t}{ds^2}+\frac{A'}{A}\frac{dr}{ds}\frac{dt}{ds}=0 \qquad (4.4)$$

To simplify these equations, consider (4.1). If we let $\theta=\frac{\pi}{2}$ and $\frac{d\theta}{ds}=0$ initially, then (4.1) tells us that $\frac{d^2\theta}{ds^2}=0$ initially. But this just means that $\frac{d\theta}{ds}$ will stay at zero for all s and so $\theta=\frac{\pi}{2}$ for all s. Therefore, by setting $\theta=\frac{\pi}{2}$ and $\frac{d\theta}{ds}=0$ initially, the planet Mercury will lie in the plane defined by the relation $\theta=\frac{\pi}{2}$. The geodesic equations (4.2) and (4.3) becomes

$$\frac{d^2r}{ds^2}+\left(\frac{c^2A'}{2B}\right)\left(\frac{dt}{ds}\right)^2+\left(\frac{B'}{2B}\right)\left(\frac{dr}{ds}\right)^2-\left(\frac{r}{B}\right)\left(\frac{d\phi}{ds}\right)^2=0 \qquad (4.5)$$

$$\frac{d^2\phi}{ds^2}+\frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds}=0 \qquad (4.6)$$

Let us solve (4.6) first. If we multiply both sides of (4.6) by r^2 and integrate with respect to s, then we get

$$0=r^2\frac{d^2\phi}{ds^2}+2r\frac{dr}{ds}\frac{d\phi}{ds}=r^2\frac{d}{ds}\left(\frac{d\phi}{ds}\right)+\frac{d(r^2)}{ds}\frac{d\phi}{ds}=\frac{d}{ds}\left(r^2\frac{d\phi}{ds}\right)$$

$$\Rightarrow r^2 \frac{d\phi}{ds} = K_1 \qquad (4.7)$$

for some constant K_1.

Similarly, we can solve (4.4) by multiplying both sides by A^2 and integrate.

$$0 = A\frac{d^2t}{ds^2} + A'\frac{dr}{ds}\frac{dt}{ds} = A\frac{d}{ds}\left(\frac{dt}{ds}\right) + \frac{dA}{ds}\frac{dt}{ds} = \frac{d}{ds}\left(A\frac{dt}{ds}\right)$$

$$\Rightarrow A\frac{dt}{ds} = K_2 \qquad (4.8)$$

for some constant K_2.

Instead of solving (4.5), let us go back to the Schwarzschild metric (3.3) and solve that first order ODE. Divide both sides of (3.3) by ds^2 and set $\theta = \frac{\pi}{2}$, we get

$$1 = A(r)c^2\left(\frac{dt}{ds}\right)^2 - B(r)\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{d\phi}{ds}\right)^2 \qquad (4.9)$$

Substitute (4.7) and (4.8) into (4.9), we get

$$1 = \frac{c^2K_2^{\ 2}}{A} - B\left(\frac{dr}{ds}\right)^2 - r^2\left(\frac{K_1}{r^2}\right)^2 \qquad (4.10)$$

Similar to the classical problem, we can also change (4.10) to an ODE depending on r and ϕ instead of r and s by writing $\frac{dr}{ds} = \frac{dr}{d\phi}\frac{d\phi}{ds}$ and apply (4.7) again.

$$1 = \frac{c^2K_2^{\ 2}}{A} - \frac{BK_1^{\ 2}}{r^4}\left(\frac{dr}{d\phi}\right)^2 - \frac{K_1^{\ 2}}{r^2} \qquad (4.11)$$

Let us substitute $u = \frac{1}{r}$ into (4.11) and apply (3.17) and (3.21),

$$1 = \frac{c^2K_2^{\ 2}}{A} - BK_1^{\ 2}\left(\frac{du}{d\phi}\right)^2 - K_1^{\ 2}u^2$$

$$\Rightarrow c^2 - 2uGM = c^4K_2^{\ 2} - c^2K_1^{\ 2}\left(\frac{du}{d\phi}\right)^2 - u^2K_1^{\ 2}(c^2 - 2uGM) \qquad (4.12)$$

Differentiate (4.12) with respect to ϕ,

$$-2GM\frac{du}{d\phi} = -2c^2K_1^{\ 2}\left(\frac{du}{d\phi}\right)\left(\frac{d^2u}{d\phi^2}\right) - 2uc^2K_1^{\ 2}\frac{du}{d\phi} + 6u^2GMK_1^{\ 2}\frac{du}{d\phi} \qquad (4.13)$$

$\frac{du}{d\phi}=0$ gives the solution $r=$constant which is the circular orbit. If $\frac{du}{d\phi}\neq 0$, then we can simplify (4.13) and get a second order differential equation.

$$\frac{d^2u}{d\phi^2}+u=\frac{GM}{K_1{}^2c^2}+\frac{3GM}{c^2}u^2 \tag{4.14}$$

Compare this equation with (2.17) corresponding to the classical problem. They look almost the same except for the term $\frac{3GM}{c^2}u^2$. Since perihelion shift is small, we should expect this term to be small and in fact $\frac{3GM}{c^2}=4.42\times10^{-21}<<1$.

This ODE is well-known problem in applied Mathematics and it can be solved using perturbation expansion. So let

$$u(\phi)=u_0(\phi)+\frac{3GM}{c^2}u_1(\phi)+O\left[\left(\frac{3GM}{c^2}\right)^2\right] \tag{4.15}$$

and we substitute (4.15) into (4.14). Ignoring the higher order terms, we get

$$\frac{d^2u_0}{d\phi^2}+u_0+\frac{3GM}{c^2}\frac{d^2u_1}{d\phi^2}+\frac{3GM}{c^2}u_1=\frac{GM}{K_1{}^2c^2}+\frac{3GM}{c^2}u_0{}^2 \tag{4.16}$$

Equate the zeroth-order terms in (4.16), we have

$$\frac{d^2u_0}{d\phi^2}+u_0=\frac{GM}{K_1{}^2c^2} \tag{4.17}$$

Equate the first-order terms in (4.16), we have

$$\frac{d^2u_1}{d\phi^2}+u_1=u_0{}^2 \tag{4.18}$$

So, if solve (4.17) similar to what we did in section 2, we get

$$u_0=\frac{GM}{K_1{}^2c^2}+C\cos(\phi+\delta) \tag{4.19}$$

Again, we choose an appropriate orientation of the axes, so that $\delta=0$. So (4.19) becomes

$$u_0=\frac{GM}{K_1{}^2c^2}+C\cos\phi \tag{4.20}$$

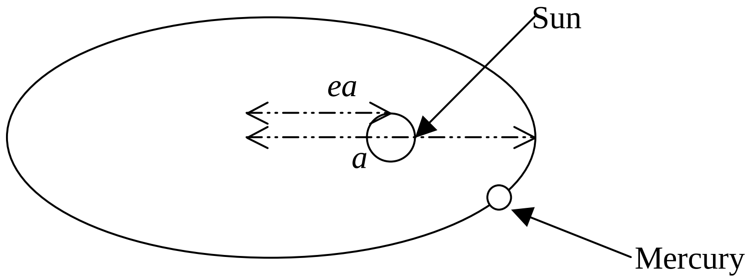
(4.17) and (2.17) are essentially the same, so we can relate C to the length of major axis, $2a$, and the eccentricity, e, of the elliptical orbit. In polar coordinates, equation of ellipse is $r=\frac{ed}{1-e\cos\phi}$, where $x=-d$ is the directrix. So rewrite (4.20) and compare, we get $\frac{1}{u_0} = \frac{1}{\frac{GM}{K_1^{\ 2}c^2} + C\cos\phi} = \frac{\frac{K_1^{\ 2}c^2}{GM}}{1 + \frac{CK_1^{\ 2}c^2}{GM}\cos\phi}$. So the eccentricity is

$e = -\frac{CK_1^{\ 2}c^2}{GM}$ and we have

$$C = -\frac{eGM}{K_1^{\ 2}c^2} \qquad (4.21)$$

Substitute this into (4.20) we have

$$u_0 = \frac{GM}{K_1^{\ 2}c^2}(1 - e\cos\phi) \qquad (4.22)$$



Now we have to determine K_1 in terms of a and e. From (4.22), we know that u_0 is minimized at $\phi = 0$ and is maximized at $\phi = \pi$. That means $r \approx \frac{1}{u_0}$ is maximized at $\phi = 0$ and minimized at $\phi = \pi$. So

$$\left.\frac{1}{u_0}\right|_{\theta=\pi} + \left.\frac{1}{u_0}\right|_{\theta=0} = 2a \qquad (4.23)$$

By (4.22), (4.23) can be rewritten as

$$\frac{1}{\frac{GM}{K_1^{\ 2}c^2}(1+e)} + \frac{1}{\frac{GM}{K_1^{\ 2}c^2}(1-e)} = 2a$$

$$\Rightarrow \frac{1}{\frac{GM}{K_1^{\ 2}c^2}(1-e^2)} = a$$

$$\Rightarrow K_1^{\ 2} = \frac{aGM(1-e^2)}{c^2} \qquad (4.24)$$

So (4.22) becomes

$$u_0 = \frac{1}{a(1-e^2)}(1 - e\cos\phi) \qquad (4.25)$$

Next, we have to solve (4.18). So, substitute (4.25) into (4.18), we have

$$\frac{d^2u_1}{d\phi^2}+u_1=\frac{1}{a^2(1-e^2)^2}(1-2e\cos\phi+4e^2\cos^2\phi)$$

$$\Rightarrow \frac{d^2u_1}{d\phi^2}+u_1=\frac{1}{a^2(1-e^2)^2}(1+2e^2-2e\cos\phi+2e^2\cos 2\phi) \qquad (4.26)$$

The homogeneous solution to (4.26) is periodic and will be small comparing to (4.25) (since it will have $\frac{3GM}{c^2}$ before it), so we only need the non-homogeneous solution which will increase as θ increases. To find the non-homogeneous solution to (4.26), we can separate (4.26) into three equations.

$$\frac{d^2v_1}{d\phi^2}+v_1=\frac{1+2e^2}{a^2(1-e^2)^2}, \quad \frac{d^2v_2}{d\phi^2}+v_2=-\frac{2e\cos\phi}{a^2(1-e^2)^2}, \quad \frac{d^2v_3}{d\phi^2}+v_3=\frac{2e^2\cos 2\phi}{a^2(1-e^2)^2} \qquad (4.27)$$

Then $u_1=v_1+v_2+v_3$ is a non-homogeneous solution to (4.26) where v_1,v_2,v_3 are particular solutions to (4.27). And they are

$$v_1=\frac{1+2e^2}{a^2(1-e^2)^2}, \quad v_2=-\frac{e}{a^2(1-e^2)^2}\phi\sin\phi, \quad v_3=-\frac{2e^2}{3a^2(1-e^2)^2}\cos 2\phi \qquad (4.28)$$

where v_1 and v_3 can be easily found using method of undetermined coefficient and v_2 can be found using variation of parameters.

Combining (4.28) and (4.25), (4.15) becomes

$$u(\phi)\approx\frac{1}{a(1-e^2)}(1-e\cos\phi) + \frac{3GM}{c^2}\left[\frac{1+2e^2}{a^2(1-e^2)^2}-\frac{e}{a^2(1-e^2)^2}\phi\sin\phi-\frac{2e^2}{3a^2(1-e^2)^2}\cos 2\phi\right] \qquad (4.29)$$

Clearly, $-\frac{e}{a^2(1-e^2)^2}\phi\sin\phi$ is the non-periodic term which account for the perihelion shift. To determine what exactly the perihelion shift is, we rewrite (4.29) and make some approximation again.

$$u(\phi)\approx\frac{1}{a(1-e^2)}-\frac{e}{a(1-e^2)}\cos\phi-\frac{3GM}{c^2}\frac{e}{a^2(1-e^2)^2}\phi\sin\phi + \frac{3GM}{c^2}\left[\frac{1+2e^2}{a^2(1-e^2)^2}-\frac{2e^2}{3a^2(1-e^2)^2}\cos 2\phi\right] \qquad (4.30)$$

Again, since $\frac{3GM}{c^2}\left[\frac{1+2e^2}{a^2(1-e^2)^2}-\frac{2e^2}{3a^2(1-e^2)^2}\cos 2\phi\right]$ is small and periodic, it will not influence the orbit much for large ϕ. Therefore, we ignore this term and look at

$$\frac{1}{a(1-e^2)}-\frac{e}{a(1-e^2)}\cos\phi-\frac{3GM}{c^2}\frac{e}{a^2(1-e^2)^2}\phi\sin\phi \qquad (4.31)$$

To simplify (4.31), consider $\cos(\phi-\varepsilon\phi)=\cos\varepsilon\phi\cos\phi+\sin\varepsilon\phi\sin\phi$. For ε small, we can approximate this by

$$\cos(\phi-\varepsilon\phi)\approx\cos\phi+\varepsilon\phi\sin\phi \qquad (4.32)$$

Replace ε in (4.32) by $\dfrac{3GM}{ac^2(1-e^2)}$ and apply it on (4.31), then we have

$$u(\phi)\approx\frac{1}{a(1-e^2)^2}-\frac{e}{a(1-e^2)}\cos\left[\phi-\frac{3GM}{ac^2(1-e^2)}\phi\right] \qquad (4.33)$$

Compare (4.33) with the classical solution (2.19), we see that (4.33) described an elliptical orbit for ϕ small and we will see the effect of perihelion shift for ϕ large. Perihelion occurs when r is minimized or when u is maximized, so differentiate (4.33) we will get $u'(\phi)\approx\dfrac{e}{a(1-e^2)}\left[1-\dfrac{3GM}{ac^2(1-e^2)}\right]\sin\left[\phi-\dfrac{3GM}{ac^2(1-e^2)}\phi\right]$. So,

$$\left[1-\frac{3GM}{ac^2(1-e^2)}\right]\phi=2\pi n \qquad (4.34)$$

will give maximum u or approximately

$$\phi=2\pi n\left[1+\frac{3GM}{ac^2(1-e^2)}\right] \qquad (4.35)$$

Here, we use again the fact that $\dfrac{3GM}{ac^2(1-e^2)}$ is small. So, instead of 2π, the successive perihelia will occur at intervals of

$$\Delta\phi=2\pi\left[1+\frac{3GM}{ac^2(1-e^2)}\right] \qquad (4.36)$$

Therefore, for each revolution the perihelion will advance an amount

$$\delta\phi=\frac{6\pi GM}{ac^2(1-e^2)} \qquad (4.37)$$

Using (4.37), one can calculate the perihelion shift of Mercury to be 42.89 sec per century and the observational result gives 42.6 ± 1.0 sec per century. This agreement with the experimental result and the fact that this shift cannot be explained using classical theory serve as an important verification of the General Theory of Relativity.

5. A Classical Explanation on Perihelic Shift

We mentioned at the end of Section 1 that Le Verrier tried to explain the perihelic shift of Mercury by hypothesizing a planet Vulcan. But such an explanation was given up since such planet is never found. However, if we assumed that the Sun is a slightly flattened sphere into an ellipsoid, then the planetary orbits will be shift. Here is a very brief (because time is running out) discussion of this explanation.

In our new system with a slightly flattened sphere, conservation of energy and conservation of angular momentum still hold. The only thing that changes is the potential. So, let us calculate the potential created by a sphere which is widened by a bulge around its equator. Develop the potential in spherical harmonics (which I don't know the detail), the leading two terms will be of the form

$$V(r,\theta) = -\frac{GM}{r}\left[1 + D\frac{3\cos^2\theta - 1}{r^2}\right] + O\left(\frac{1}{r^4}\right) \tag{5.1}$$

where D is some constant depends on the deformation.

Since r is huge and the deformation from the sphere is small, we ignore the terms $O\left(\frac{1}{r^4}\right)$. Assume the motion of the planets is restricted to the plane described by the relation $\theta = \frac{\pi}{2}$, then (5.1) becomes

$$V(r) = -\frac{GM}{r} - \frac{B}{r^3} \tag{5.2}$$

where $B = 2DGM$.

By equation (2.16), we have

$$u''+u = \frac{m}{l^2u^2}\left(\frac{GM}{r^2} + \frac{3B}{r^4}\right) = \frac{m}{l^2u^2}(GMu^2 + 3Bu^4) = \frac{GMm}{l^2} + \frac{3Bm}{l^2}u^2 \tag{5.3}$$

The constant B depends on deformation of the sphere which is assumed to be small, so we can apply perturbation theory. (5.3) is very similar to (4.14). If we do a similar calculation as what we did in Section 4, we will get

$$\delta\phi = \frac{6\pi GMm^4}{l^4}B \tag{5.4}$$

Let T be the period of revolution expressed in units of centuries, then the perihelic shift per century is given by

$$S = \frac{\delta\phi}{T} = \frac{6\pi GMm^4}{l^4 T}B \tag{5.5}$$

Next, we will expressed (5.5) in terms of R, the mean distance of the planet from the Sun. Then by (2.8), we have

$$l \approx mR^2\dot{\theta} \tag{5.6}$$

Integrate (5.6) with respect to t over one revolution, we get

$$lt\Big|_{t=0}^{t=T} \approx mR^2\theta\Big|_{\theta=0}^{\theta=2\pi}$$

$$\Rightarrow lT \approx 2\pi mR^2 \tag{5.7}$$

Also, by Kepler's third law which can be easily proven using Newton's Law of Gravitation, we have

$$\frac{T^2}{R^3} = \mathrm{K} \tag{5.8}$$

where K is the Kepler constant which is the same for all planets. Using (5.7) and (5.8), (5.5) becomes

$$S = \frac{3GMBK^{3/2}}{(2\pi)^3}R^{-7/2} \tag{5.9}$$

Let us do the same thing using the result we found in Section 4. By (4.37), we have $\delta\phi = \dfrac{6\pi GM}{ac^2(1-e^2)}$, so

$$S = \frac{6\pi GM}{ac^2(1-e^2)T} \approx \frac{6\pi GM}{c^2T}R^{-1} \tag{5.10}$$

Substitute (5.8) into (5.10), we have

$$S = \frac{6\pi GM}{c^2}R^{-5/2} \tag{5.11}$$

Compare (5.5) and (5.11), we see that the R dependence of the two equations is the main difference. So, let us take the logarithm on both sides of (5.5) and (5.11), then $\log S$ and $\log R$ will give a linear relation. The classical theory will predict a linear relation with slope $-7/2 = -3.5$ while relativity theory will give a linear relation with slope $-5/2 = -2.5$.

Observed data from reference (3)	Mercury	Venus	Earth
Mean distance from the sun R	58×10^{11}	108×10^{11}	149×10^{11}
Observed shift S	43.11 ± 0.45	8.4 ± 4.8	5.0 ± 1.2

If we plot $\log S$ against $\log R$ using the observed data and draw the best-fit line, we will get a straight line with slope -2.30 ± 0.26. This shows that General Theory of Relativity gives a much better explanation on perihelic shift than this classical explanation. In fact, there is no classical theory that can explain this perihelic shift, so Mercury Perihelion remains to be the verification of the General Theory of Relativity.

References:

(1) Robert M. Wald. General Relativity. The University of Chicago Press, 1984.

(2) J. Weber. General Relativity and Gravitational waves. Interscience Publishers, inc., New York, 1961.

(3) Adler Bazin and Schiffer. Introduction to General Relativity. McGraw-Hill, Inc., 1965.

(4) J. David Logan. Applied Mathematics, 2nd Edition. John Wiley & Sons, Inc. 1997.

(5) G. ‘t Hooft. Introduction to General Relativity. 2001.

(6) Sean M. Carroll. Lecture Notes on General Relativity. 1997.