

Mercury's Perihelion

Chris Pollock

March 31, 2003

Introduction In 1846, the French Astronomer Le Verriere, doing calculations based on Newton's theory of gravitation, pin-pointed the position of a mass that was perturbing Uranus' orbit. When fellow astronomers aimed their telescopes at his location, they recognized, for the first time, the eighth planet. Newton's theory had reached its zenith. Shortly after, however, it became clear to Le Verriere that additional mass, nearer to the sun than Mercury, was needed to explain the strange advance of Mercury's orbit. When no such mass was observed, astronomers began to doubt Newton's theory. Then, along came Albert Einstein, whose theory nearly perfectly explained Mercury's erstwhile mysterious motion. This essay is a history of Newton's theory of gravity, the enigma of Mercury, and Einstein's convincing solution. It will blend together mathematical and physical theories with a narrative about the brilliant scientists who chose to tackle the problem of gravitation. In particular, this paper will show the following computations:

1. A derivation of Kepler's First Law concerning elliptical orbits from Newton's Law of Gravitation and Newton's Second Law.
2. An outline of a method to approximate non-relativistic perturbations on Mercury's orbit by assuming external planets are heliocentric circles of uniform linear mass density.
3. A calculation of relativistic perihelion shift using Einstein's theory of relativity and the Schwarzschild solution.

Pre-Newtonian Theories and Ideas The problem I set out to explain, the advance of Mercury's perihelion, was of tantamount importance to the astronomical community. This seemingly miniscule enigma stood glaringly in the way of humanity's understanding of the solar system. Let's begin by reviewing the early evolution of celestial theories.

In the second century CE, the Greek astronomer Claudius Ptolemy hypothesized that the sun, moon, and planets orbited the earth along circles called deferents. On a smaller scale, they travelled along smaller circles, called epicycles, whose centres moved along the deferents. This theory was sufficiently accurate to explain observations of the time. It dominated until around 1543, when Copernicus published his book, *de Revolutionibus Orbium Caelestium*. His solar system was heliocentric but retained the orbital deferents and epicycles of Ptolemy's theory. It is interesting to note that Copernicus had difficulty explaining Mercury's motion and once commented that "this planet has . . . influenced many perplexities and labours on us in our investigation of its wanderings." (from Baum, 11)

More exact naked-eye planetary observations were taken by Tycho Brahe, who, I discovered while writing this paper, *did not* cut off his own nose. I had always believed, courtesy of my eleventh grade science teacher, that the great Danish observational astronomer 'removed' his nose to allow his face a better seal with his telescope. My teacher, it turns out, was doubly dishonest since

Brahe lost part of his nose in a duel with a student and telescopes had not been invented in 1600, when Brahe made his observations.

A young mathematician who worked with Brahe examined the elder's notes and tried to calculate the orbit of Mars. Although an apparently simple task, this was disconcertingly complex since planets appear to change speed and direction based on the motion of not only the planet but also Earth.

When Kepler finished his calculations, he determined that Mars moved in an ellipse with the sun at one focus. Kepler became quite proficient at predicting Mercury's passes between the Sun and Earth, called transits. By hand, he calculated Mercury's 7 November 1631 transit time, accurate to within five hours (Baum, 15), although he died before witnessing the event he had predicted. A subsequent transit, on 23 October 1651, was predicted, using corrections of Kepler's calculations, with an accuracy of a few minutes.

Based on his planetary observations, Kepler made the following statements, known collectively as Kepler's Laws (Stewart, 897):

1. A planet revolves around the Sun in an elliptical orbit with the Sun at one focus
2. The line joining the Sun to a planet sweeps out equal areas in equal times.
3. The square of the period of revolution of a planet is proportional to the cube of the length of the major axis of its orbit.

And so, by the middle of the 17th century, kinematic data on planetary paths were fairly accurately known. Hence, all of the pieces were on the table for a brilliant theorist (read: Newton) to assemble, explaining the dynamics behind planetar motion.

Newton's Law Issac Newton was born on Christmas Day, 1642. Since he was a failure at farming, his mother sent him to university. In the middle of his study at Cambridge, the plague broke out, and the school was closed to students for 1665-1666. Newton returned home, and in this marvellously creative period, wrote about both gravitation and calculus. On the former question, he considered a rock twirling around on the end of a string. The rock, he knew, tended to launch, but the string provided a counteracting force. Newton wondered, then, what provided a counteracting force in the case of planetary motion. Could it be gravity, the force that held people on the Earth's surface?

In his magnum opus, 1687's *Philosophiae Naturalis Principia Mathematicae*, Newton calculated that the path of planets would be elliptical if they were subjected to a force of gravitation that varied with the inverse square of the separation between the planets and the Sun. Starting from Newton's Second Law and the Newton's Law of Gravitation, it is possible to prove each of Kepler's Laws. Since Kepler's First Law deals with the elliptical shape of planetary orbits, a derivation is included below. The discussion that follows is based on Chapter 11.4 of Stewart's *Calculus*. This is the so-called "one-body problem" since the system under consideration contains a test particle (the planet) moving

under the attraction of a massive body (the Sun). The planet's mass is assumed to be so small relative to that of the Sun that the Sun remains fixed in space. Also, the effects of all other planets are neglected.

Let's begin by stating Newton's Second Law:

$$\mathbf{F} = m\mathbf{a} \quad (1)$$

where \mathbf{F} is the force experienced by the planet, m is its mass, and \mathbf{a} is its acceleration.

And Newton's Law of Gravitation:

$$\mathbf{F} = -\frac{GMm}{r^3}\mathbf{r} = -\frac{GMm}{r^2}\mathbf{u} \quad (2)$$

The coordinate system is radial. \mathbf{F} is the force experienced by the planet, G denotes the gravitational constant, $\mathbf{r} = \mathbf{r}(t)$ is the planet's position vector, \mathbf{u} is a unit vector in the direction of \mathbf{r} , and M and m are the masses of the Sun and planet, respectively. Further, $r = |\mathbf{r}|$, $\mathbf{v} = \mathbf{r}'$, and $\mathbf{a} = \mathbf{r}''$.

First, we will show that the planet moves in a plane. Equating the \mathbf{F} s from (1) and (2) gives

$$\mathbf{a} = \frac{-GM}{r^3}\mathbf{r} \quad (3)$$

hence \mathbf{a} and \mathbf{r} are parallel, so

$$\mathbf{r} \times \mathbf{a} = \mathbf{0}$$

From the properties of cross products, we know that

$$\begin{aligned} \frac{d}{dt}(\mathbf{r} \times \mathbf{v}) &= \mathbf{r}' \times \mathbf{v} + \mathbf{r} \times \mathbf{v}' \\ &= \mathbf{v} \times \mathbf{v} + \mathbf{r} \times \mathbf{a} \\ &= \mathbf{0} + \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

So, we can conclude that $\mathbf{r} \times \mathbf{v}$ is a constant vector, say \mathbf{h} . We can assume that \mathbf{r} and \mathbf{v} are not parallel. This implies that $\mathbf{h} \neq \mathbf{0}$. We can now conclude that \mathbf{r} and \mathbf{h} are perpendicular, so the planet always lies in a plane through the origin perpendicular to \mathbf{h} . Planetary motion, according to Newton's Laws, is planar.

Now, we can prove examine the shape of the orbit within this plane. Let's start by rewriting \mathbf{h} :

$$\begin{aligned} \mathbf{h} &= \mathbf{r} \times \mathbf{v} \\ &= \mathbf{r} \times \mathbf{r}' \\ &= r\mathbf{u} \times (r\mathbf{u})' \\ &= r\mathbf{u} \times (r\mathbf{u}' + r'\mathbf{u}) \\ &= r^2(\mathbf{u} \times \mathbf{u}') + rr'(\mathbf{u} \times \mathbf{u}) \\ &= r^2(\mathbf{u} \times \mathbf{u}') \end{aligned}$$

So, combining the above result with (3) gives

$$\begin{aligned}\mathbf{a} \times \mathbf{h} &= -\frac{GM}{r^2} \mathbf{u} \times r^2(\mathbf{u} \times \mathbf{u}') \\ &= -GM \mathbf{u} \times (\mathbf{u} \times \mathbf{u}') \\ &= -GM[(\mathbf{u} \cdot \mathbf{u}')\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\mathbf{u}']\end{aligned}$$

Since \mathbf{u} is a unit vector, $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 = 1$, so

$$\mathbf{a} \times \mathbf{h} = -GM[(\mathbf{u} \cdot \mathbf{u}')\mathbf{u} - \mathbf{u}'] \quad (4)$$

Now, let's consider the $(\mathbf{u} \cdot \mathbf{u}')$ term in (4). Since $\mathbf{u} \cdot \mathbf{u}$ is constant,

$$\begin{aligned}\frac{d}{dt}(\mathbf{u} \cdot \mathbf{u}) &= 0 \\ \rightarrow \mathbf{u}' \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{u}' &= 0 \\ \rightarrow \mathbf{u} \cdot \mathbf{u}' &= 0\end{aligned}$$

Substituting into (4) we get

$$\mathbf{a} \times \mathbf{h} = GM\mathbf{u}'$$

But \mathbf{h} is a constant vector, so

$$\begin{aligned}(\mathbf{v} \times \mathbf{h})' &= \mathbf{v}' \times \mathbf{h} \\ &= \mathbf{a} \times \mathbf{h} \\ \rightarrow (\mathbf{v} \times \mathbf{h})' &= GM\mathbf{u}'\end{aligned}$$

Integrating both sides yields

$$\mathbf{v} \times \mathbf{h} = GM\mathbf{u} + \mathbf{c} \quad (5)$$

where \mathbf{c} is a constant vector.

Since the planet's motion is confined to an arbitrary plane, we can, without loss of generality, call this the xy -plane. Recalling that \mathbf{h} is perpendicular to this plane, let's define a standard basis vector \mathbf{k} in the direction of \mathbf{h} . Looking at (5) and recalling that $(\mathbf{v} \times \mathbf{h})$ and \mathbf{u} both lie in the xy -plane, we can conclude that \mathbf{c} is also in the xy -plane. Let's choose the x - and y -axes such that the standard basis vector \mathbf{i} is in the direction of \mathbf{c} . There is really nothing special about this choice. Any set of axes would do just fine. Let's go one step further and call θ the angle between \mathbf{c} and \mathbf{r} . Hence (r, θ) is the planet's position in polar coordinates. If, at this point, it is unclear where we are heading, recall that we set out to show that the planet's path is elliptical. So, ultimately, we're seeking

an expression for $r(\theta)$. Let's look at an expression for $\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h})$, substitute in (5), and turn the crank . . .

$$\begin{aligned}\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h}) &= \mathbf{r} \cdot (GM\mathbf{u} + \mathbf{c}) \\ &= GM\mathbf{r} \cdot \mathbf{u} + \mathbf{r} \cdot \mathbf{c} \\ &= GMr\mathbf{u} \cdot \mathbf{u} + |\mathbf{r}||\mathbf{c}| \cos \theta \\ &= GMr + rc \cos \theta\end{aligned}$$

where $c = |\mathbf{c}|$.

Now, let's solve for r in the previous equation.

$$r = \frac{\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h})}{GM + c \cos \theta}$$

Define $e = \frac{c}{GM}$, and substitute into the above equation:

$$r = \frac{\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h})}{GM[1 + e \cos \theta]} \quad (6)$$

So, we now have the desired equation for r . Let's simplify the numerator of (6).

$$\begin{aligned}\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h}) &= (\mathbf{r} \times \mathbf{v}) \cdot \mathbf{h} \\ &= \mathbf{h} \times \mathbf{h} \\ &= h^2\end{aligned}$$

where $h = |\mathbf{h}|$. So, (6) becomes

$$\begin{aligned}r &= \frac{h^2}{GM[1 + e \cos \theta]} \\ &= \frac{eh^2}{c[1 + e \cos \theta]}\end{aligned}$$

And, for one final simplification, let's rename $\frac{h^2}{c}$ as d . So

$$r = \frac{ed}{1 + e \cos \theta} \quad (7)$$

This is the equation of a conic section in polar coordinates. The value of e in the denominator determines the type of conic section (7) represents. Since planets have closed orbits, $0 < e < 1$ and r represents an ellipse. Further, e is the eccentricity of Mercury's orbit, and one of the ellipse's foci is located at the origin. So, beginning with only Newton's Second Law and Law of Gravitation,

we have verified Kepler's First Law, that planets moves in an elliptical orbit with the sun at one of the foci.

Of course, the actual solar system is far more complex than this model. Newton's central idea, that planetary motions are governed by an inverse square force of gravity, gave mathematicians and astronomers the tools to explain the clockwork motion of the planets, as long as precise masses and positions were known for all planets. Shortly after *Principia's* publication, scientists began to use these laws to make successively more accurate predictions.

Applying Newton's Amazing Law of Gravitation Comet Halley had been predicted to return, by Halley himself, in winter 1758-59. In 1757, Alexis-Claude Clairault, of France, and two assistants rushed to find a more precise return date. Using Newton's Law of Gravitation and considering perturbations caused by Jupiter and Saturn, they arrived at a date of mid-April 1759. Comet Halley's perihelion, the point in its orbit nearest to the sun, was only 33 days earlier than they had predicted. This error was due mainly to computational shortcuts they made to ensure that they finished their calculations *before* Comet Halley arrived.

William Herschel, who was born in Hanover in 1738 but moved to England in 1757, made a shocking discovery in 1781, making use of the telescope. He found a seventh planet orbiting the Sun. Uranus became the first planet discovered since antiquity. As predicted by Newton's theory, Uranus followed an elliptical path. The astronomers of the world excitedly turned their attention to this new planet.

There is one figure who played so important a role in verifying and employing Newton's amazing theory that it is necessary to look at his life in more detail. That man is Urbain Jean Joseph Le Verriere, born in Normandy in 1811. He attended Ecole Polytechnique and graduated with a degree in chemistry. He received an appealing job offer that required him to move away from Paris and his girlfriend. He declined the job, chose marriage, gave up chemistry, and ended up in astronomy through a fortuitous job offer.

In 1839, he began calculating the stability of the solar system and planetary orbits. In 1841, then, he applied his "unique analytic skills and almost superhuman endurance for calculation to master the motion of Mercury." (Baum, 71) This planet orbits nearest the sun and has a large eccentricity and short period. A detailed description of its motion had, so far, eluded theorists and stood out as a critical test for Newton's Gravitation. In 1845, Le Verriere gained reknown for calculating Mercury's 1845 transit of the Sun to within 16 seconds (Baum, 73). Unsatisfied by this error, however, he did not publish his tables of Mercury's motion. At this point, he left Mercury (!) temporarily and turned to another problem.

Since the discovery of Uranus, its orbit had been lagging behind predictions. Le Verriere, always sure of Newton, set to calculating the position and mass of an object that could bring about the observed lag in Uranus' orbit. He began with an assumption that the unknown mass must be exterior to Uranus's orbit, since a perturbation had not been observed in Saturn's orbit. He then tried to calculate the location, mass, and orbital parameters of such an object.

In mid-1845, a young Cambridge mathematician, John Couch Adams, was also considering the same problem. Adams was, in fact, the first to theorize the new planet's position, but his results were not well disseminated. By 1846, the famous Le Verriere too issued a prediction as to the new planet's position, and a team at the Berlin Observatory found it almost immediately, only 55 arc minutes from Le Verriere's calculation, and 2.5 degrees from Adams'. The director of the observatory exclaimed that this was "the most outstanding conceivable proof of the validity of universal gravitation" (Baum, 117). Indeed, Newton's theory had reached its peak. Its power, in the hands of a talented mathematician or astronomer like Le Verriere of Adams, seemed without bounds. With another success behind him, Le Verriere then returned to the problem of Mercury.

The Hunt for an Intermercurial Planet In 1854, Le Verriere was appointed director of the Paris Observatory. He undertook, with the assistance of many human 'computers,' a study of the whole solar system. His dream was to explain every minute movement of the "jewelled clockwork that seemed, for an elusive moment, to be stable, self-adjusting, and eternal" (Baum, 2).

In 1859, he began examining observations of Mercury's motion in detail, relying mainly on 14 very accurate solar transit times, recorded between 1697 and 1848. He concluded that the ellipse of Mercury's orbit was precessing slowly, which was expected, since the other planets of the solar system were expected to Mercury's perihelion to precess. The observed precession amounted to a perihelion advance of approximately 565 seconds of arc per Earth century. Le Verriere then calculated an expected precession by considering the effects force each planet exerted on Mercury. This was an accurate, though computationally demanding technique. The outer planets, Le Verrier predicted, should cause an advance of 527 seconds of arc per century, leaving a residual 38 seconds that he was not able to explain using Newton's theory (Baum, 136). In the next section, we will calculate an approximation of this Newtonian perturbation.

An Approximation of the Newtonian Precession of Mercury's Orbit It is possible to model, very precisely, the Newtonian force each planet was expected to exert on Mercury. An elegant approximation, however, was described by Price and Rush in their paper entitled *Non-relativistic contribution to Mercury's perihelion precession*. They replaced each of the outer planets by a ring of uniform linear mass density. Since Mercury's precession is slow compared to the orbits of even Uranus, their approach yields a fairly accurate time-averaged effect of the moving planets' effects. An sketch of their calculations follows.

First, let's replace each planet by a ring with linear mass density given by the following equation:

$$\lambda_i = \frac{M_i}{2\pi R_i}$$

where λ_i is the linear mass density along the orbit of the i th planet from the sun, M_i is the mass of the i th planet, and R_i is the radius of the i th planets orbit, which is assumed to be circular.

We can calculate the gravitational pull exerted on Mercury by each of the exterior planets by considering the radial components of the force exerted by

differential angular elements on Mercury using Newton's Law of Gravitation and integrating over the circle. Please see page 532 of Price and Rush's paper for further details on this integration.

This yields

$$\mathbf{F}_i = \left[\frac{G\pi\lambda_i m a}{R_i^2 - a^2} \right] \mathbf{u} \quad (8)$$

where \mathbf{F}_i is the radial force exerted on Mercury by the i th planet, a is Mercury's distance from the Sun, and \mathbf{u} is a unit vector for Mercury's position. Note that since $R_i > a$, \mathbf{F}_i in (8) is positive, so the force exerted on Mercury by each external planet is directed outward, opposite to the force exerted by the Sun.

Now, we will refer to some results from Classical Mechanics. From the requirements that planetary orbits are stable and closed, the following differential equation relates gravitational force and planetary motion:

$$\Phi(r) = m(\ddot{r} - \dot{\theta}^2 r)$$

where $\Phi(r)$ is the total applied force, and "." represents differentiation with respect to time.

Also, the planet's angular momentum, $mr^2\dot{\theta}$, is a constant, say h . Substituting this value into the above equation yields

$$\Phi(r) = m\left(\ddot{r} - \frac{h^2}{m^2 r^3}\right)$$

Now, let's consider an orbit for Mercury that oscillates about a circular orbit of radius a . Using a Taylor series approximation and solving the resulting differential equation gives an equation for apsidal angle, which is defined as the angle between the perihelion and aphelion:

$$\Psi = \frac{\pi}{\sqrt{3 + a \left[\frac{\Phi'(a)}{\Phi(a)} \right]}} \quad (9)$$

where Ψ is the apsidal angle, $\Phi(a)$ is the net central force, and a is the radius of the circular orbit around which we are perturbing.

Now, let's use (8) to find the sum of the forces of all of the external planetary rings on Mercury:

$$F(a) = G\pi m \sum_{i=2}^9 \frac{\lambda_i a}{R_i^2 - a^2} = 7.587 \times 10^{15} N \quad (10)$$

The numerical value was obtained by substituting masses and orbital radii for all external planets into (10).

The Gravitational force exerted by the Sun on Mercury is $F_\circ = -1.318 \times 10^{22} N$, and the net force that Mercury experiences is given by $\Phi(a) = F_\circ + F(a)$. Let's differentiate this expression and multiply by a .

$$a\Phi'(a) = aF_\circ' + aF'(a) \quad (11)$$

Let's now examine the terms on the right-hand side of (11). Differentiating (10) yields:

$$F'(a) = G\pi m \sum_{i=2}^9 \frac{\lambda_i(R_i^2 + a^2)}{R_i^2 - a^2} \quad (12)$$

And F_\circ , the magnitude of the force exerted on Mercury by the Sun, is given by Newton's Law of Gravity

$$F_\circ = -\frac{GM_s m}{a^2}$$

where M_s is the mass of the Sun. Differentiating,

$$\begin{aligned} F'_\circ &= \frac{2GM_s}{a^3} = \frac{-2}{a} F_\circ \\ \rightarrow aF'_\circ &= -2F_\circ \end{aligned} \quad (13)$$

Substituting (12) and (13) into (11) gives the following for Mercury's apsidal equation:

$$\Psi = \frac{\pi}{\sqrt{3 + \frac{-2F_\circ + Gm\pi a \sum_{i=2}^9 \lambda_i \frac{R_i^2 + a^2}{(R_i^2 - a^2)^2}}{F_\circ + F(a)}}} \quad (14)$$

We now apply a binomial expansion on both the numerator and denominator of (14) and neglect all terms higher than first power of $\frac{F(a)}{F_\circ}$:

$$\Psi = \pi \left(1 - \frac{F(a)}{F_\circ} - \frac{Gm\pi a \sum_{i=2}^9 \lambda_i \frac{R_i^2 + a^2}{(R_i^2 - a^2)^2}}{2F_\circ} \right) \quad (15)$$

Now, compute the numerical value for Ψ by substituting the value of $F(a)$ calculated in (10), F_\circ , and mass and orbital radius for the exterior planets into (15):

$$\Psi = \pi(1 + 9.884 \times 10^{-7})$$

where Ψ is the angle between perihelion and aphelion. Please see p533 of Price and Rush's article for the numerical valuse used in calculating Ψ .

The precession of Mercury's orbit per revolution, then, is

$$\frac{2\Psi - 2\pi}{T}$$

where T is the period of Mercury's orbit, 87.969 days. So, in more conventional units,

$$precession = \frac{531.9 \text{ arcsec}}{\text{century}} \quad (16)$$

Le Verriere predicted a perihelion advance of 527 arcseconds per century, and this approximation yields 531.9 arcseconds. Still, Newton's theory was not able to explain the observed perihelion precession using *known masses in the*

solar system. Le Verrier, naturally, tried to apply the same technique that had led him to the discovery of Neptune and began looking for a missing body in the solar system.

The Hunt for an Intermercurial Planet An increase of approximately 10 percent in Venus' mass would explain Mercury's perihelion advance, but it would also affect Earth's orbit in a way that had not been observed. Since the missing mass must not affect Earth, Le Verriere decided that it must be nearer to the Sun than Mercury's orbit. And so began the hunt for Le Verriere's ghost planet, or rather planets. He quickly realized that a single planet so near the Sun would have been bright and, hence, visible during solar eclipses. Since no such planet had been observed during past eclipses, Le Verriere hypothesized, instead, that the mass was in the form of many small bodies.

In 1860, he visited a small-town French doctor and part-time astronomer who had recorded observations during what he believed to be a transit of an intermercurial planet. Le Verriere, convinced by the man's story, released the news of the new planet, which was quickly dubbed Vulcan. He had, again, delighted the French scientific community. The team of Newton and Le Verriere had again triumphed. Or so it seemed.

Based on the doctor's observations, Le Verriere calculated the planet's distance from the sun, 0.147AU, and period, 19 days 17 hours (Baum, 156). The astronomical community tried again and again to observe the elusive planet Vulcan, but, as time passed with no further sightings, doubts as to the planet's existence began to mount. With so much attention directed at observing the area around the sun during subsequent eclipses, and with no, or at least very few, credible sightings, most astronomers lost faith in Vulcan. Until his death in 1877, Le Verriere remained utterly convinced that the missing mass existed and would eventually be found, verifying the supremacy of Newton's Law of Gravitation.

By 1890, however, almost nobody believed a sufficient amount of matter would be found inside Mercury's orbit to explain the perihelion advance. In 1895, Newcomb corrected some inconsistencies in planetary mass and repeated Le Verriere's calculations. He found an extra perihelion shift of 43" per century, slightly larger than Le Verriere's result. Since hidden mass was, by this time, out of the question, perhaps, he thought, the problem lay with Newton's Law of Gravitation. If, for example, the exponent in the denominator were 2.00000016 (Baum, 233), then the motion of Mercury could be more accurately explained.

This represented a true paradigm shift. While formerly observations were questioned and Newton's theory was unassailable, now credible scientists began to question the foundations of Newton Law of Gravitation. The observational problem was, essentially, closed. The theory problem, in contrast, was wide open.

Einstein's General Theory of Relativity In November 1915, Albert Einstein, sitting at a desk in Berlin, wrote his General Theory of Relativity. It is ironic to note that Newton's theory had peaked in the same town, only 70 years earlier, when Neptune was discovered at the Berlin Observatory.

Einstein did not set out to solve the problem of Mercury's perihelion shift.

In fact, the answer fell out as a neat consequence of his theory. In Newton's theory, as demonstrated previously, the motion of planets are ellipses with the Sun at one focus. Here, I will show that Einstein's Theory of General Relativity changes this conclusion, although this change is minor in most cases.

The general relativistic calculation that follows will consider the motion of a test particle in the gravitational field of a massive body. The test particle's mass is assumed to be so small that it has no effect on the massive body. Fortunately, the Schwarzschild solution to Einstein's field equations describes precisely this case.

Here, I will assume that the reader is familiar with the Schwarzschild solution. A good derivation and discussion of the Schwarzschild solution can be found in Hans Stephani's *General Relativity*. Schwarzschild's spherically symmetric vacuum solution has a line element

$$ds^2 = \left(1 - \frac{2m}{r}\right)dt^2 - \frac{1}{1 - \frac{2m}{r}}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (17)$$

where m , in our case, is the mass of the sun in relativistic units.

Starting from the Schwarzschild line element, it is possible to deduce the motion of a test mass (a planet). Please refer to section 15.3 of Ray D'Inverno's *Introducing Einstein's Relativity* for further details. Before we begin, though, it is interesting to note that Einstein reached the same conclusion about Mercury's perihelion precession in 1915, without the use of Schwarzschild's solution.

Since the test mass (Mercury) moves along a timelike geodesic, the Lagrangian is identical to kinetic energy, and $g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 1$, where $\alpha, \beta = 0, 1, 2, 3$.

So, the Lagrangian, L is as follows:

$$\begin{aligned} L &= \frac{m}{2}v^2 \\ &= \frac{m}{2}g_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} \\ &= \frac{m}{2}g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta \end{aligned}$$

where $\dot{x} = \frac{dx}{d\tau}$ and τ is proper time.

So, from (17), the Lagrangian for Mercury's force-free motion is given by the following equation:

$$L = \frac{m}{2}\left[\left(1 - \frac{2m}{r}\right)\dot{t}^2 - \frac{1}{1 - \frac{2m}{r}}\dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\dot{\phi}^2\right] = \frac{m}{2} \quad (18)$$

Let $x^0 = t$, $x^1 = r$, $x^2 = \theta$, and $x^3 = \phi$. Recall the Euler-Lagrange equation:

$$\frac{\partial L}{\partial x^a} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^a}\right) = 0 \quad (19)$$

Let's now apply the Euler-Lagrange equation to (18). For the $a = 0$ case:

$$\frac{d}{d\tau}\left[\left(1 - \frac{2m}{r}\right)\dot{t}\right] = 0 \quad (20)$$

For the $a = 2$ case:

$$\frac{d}{d\tau}(r^2\dot{\phi}) - r^2 \sin\theta \cos\theta \dot{\phi}^2 = 0 \quad (21)$$

And for the $a = 3$ case:

$$\frac{d}{d\tau}(r^2 \sin^2\theta \dot{\phi}) = 0 \quad (22)$$

It is not necessary to take the $a = 1$ case, which corresponds to r . Since Mercury's motion involves four equations, $t = t(\tau)$, $r = r(\tau)$, $\theta = \theta(\tau)$, and $\phi = \phi(\tau)$, equations (18), (20), (21), and (22) provide sufficient information. We have chosen to omit the $a = 1$ case since differentiating (18) with respect to $x^1 = r$ will clearly make the biggest mess.

As with the Newtonian case, we first must establish whether planar motion is possible. Let's consider motion in the equatorial plane and check that it is stable. So, at $\tau = 0$, let's assume that $\theta = \frac{\pi}{2}$ and $\dot{\theta} = 0$. Since $\theta = \frac{\pi}{2}$, (21) gives

$$\begin{aligned} \frac{d}{d\tau}(r^2\dot{\theta}) &= 0 \\ \rightarrow r^2\dot{\theta} &= \text{constant} \end{aligned}$$

At $\tau = 0$, $r^2\dot{\theta} = 0$, so $r^2\dot{\theta} = 0$ for all τ . If we assume that $r \neq 0$, then $\dot{\theta} = 0$ for all τ . Hence $\theta = \frac{\pi}{2}$ for all τ , and planar motion is possible. Let's consider motion in the xy -plane.

Integrating (22) gives

$$r^2\dot{\phi} = h \quad (23)$$

where $h = \text{constant}$. Hence, h represents angular momentum. Next, integrate (20) to give

$$\left(1 - \frac{2m}{r}\right)\dot{t} = k \quad (24)$$

where $k = \text{constant}$.

Now, substitute (24) and $\theta = \frac{\pi}{2}$ into (18):

$$\frac{k^2}{1 - \frac{2m}{r}} - \frac{\dot{r}^2}{1 - \frac{2m}{r}} - r^2\dot{\phi}^2 = 1 \quad (25)$$

Let $u = \frac{1}{r}$. So

$$\begin{aligned} \dot{r} &= \frac{dr}{d\tau} \\ &= \frac{d}{d\tau}\left(\frac{1}{u}\right) \\ &= -\frac{1}{u^2}\left(\frac{du}{d\phi}\right)\left(\frac{d\phi}{d\tau}\right) \\ &= -\frac{1}{u^2}\left(\frac{du}{d\phi}\right)hu^2 \end{aligned}$$

$$\rightarrow \dot{r} = -h\left(\frac{du}{d\phi}\right) \quad (26)$$

Pretty soon, we'll be at a tidy differential equation for Mercury's motion. Substitute (23) and (26) and $u = \frac{1}{r}$ into (25)

$$\frac{k^2}{1-2mu} - \frac{h^2 \frac{du}{d\phi}}{1-2mu} - h^2 u^2 = 1$$

Multiply through by $\frac{1-2mu}{h^2}$

$$\frac{k^2}{h^2} - \left(\frac{du}{d\phi}\right)^2 - u^2(1-2mu) = \frac{1-2mu}{h^2}$$

And, rearranging this expression will give a first order differential equation for Mercury's motion:

$$\left(\frac{du}{d\phi}\right)^2 + u^2 = \frac{k^2-1}{h^2} + \frac{2m}{h^2}u + 2mu^3 \quad (27)$$

By differentiating (27) and dividing through by 2, we get a second order differential equation for Mercury's motion:

$$\frac{d^2u}{d\phi^2} + u = \frac{m}{h^2} + 3mu^2 \quad (28)$$

This term very closely resembles the differential encountered in Newtonian motion, except there is an additional $3mu^2$ term in (28). This term, however, is very small. In the case of Mercury, for example, the ratio of the terms on the right-hand side of (28) is on the order of 10^{-7} . We can, then, solve (28) using a perturbation method. Let's introduce a parameter $\epsilon = \frac{3m^2}{h^2}$. Also, let's express differentiation with respect to ϕ with a 'prime' and rewrite (28):

$$u'' + u = \frac{m}{h^2} + \epsilon\left(\frac{h^2 u^2}{m}\right) \quad (29)$$

All that remains is solving this differential equation. Let's assume the solution has a form $u = u_o + \epsilon u_1 + O(\epsilon^2)$. Now, we will differentiate this solution twice, substitute it into (29), and rearrange:

$$u''_o + u_o - \frac{m}{h^2} + \epsilon(u''_1 + u_1 - \frac{h^2 u_o^2}{m}) + O(\epsilon^2) = 0 \quad (30)$$

For the first approximation of a solution, let's equate the coefficients of ϵ , ϵ^2 , . . . to zero. Then $u_o = \frac{m}{h^2}(1 + \epsilon \cos\phi)$ is the zeroth order solution to (30). (Actually, the solution is $u_o = \frac{m}{h^2}(1 + \epsilon \cos(\phi - \phi_o))$, but we can set ϕ_o to zero for simplicity) This solution can be verified by differentiating twice and substituting into (30). Remembering that $u = \frac{1}{r}$, this solution agrees with the Newtonian solution.

Now, let's examine the coefficient of ϵ in (30):

$$\begin{aligned} u_1'' + u_1 &= \frac{h^2 u_o^2}{m} \\ &= \frac{m}{h^2} (1 + e \cos \phi)^2 \\ &= \frac{m}{h^2} (1 + 2e \cos \phi + e^2 \cos^2 \phi) \end{aligned}$$

$$\rightarrow u_1'' + u_1 = \frac{m}{h^2} (1 + \frac{1}{2} e^2) + \frac{2me}{h^2} \cos \phi + \frac{me^2}{2h^2} \cos 2\phi \quad (31)$$

Let's try a general solution $u_1 = A + B\phi \sin \phi + C \cos 2\phi$ and solve for its coefficients.

$$\begin{aligned} u_1' &= B \sin \phi + B\phi \cos \phi - 2C \sin 2\phi \\ u_1'' &= 2B \cos 2\phi - B\phi \sin \phi - 4C \cos 2\phi \end{aligned}$$

So,

$$u_1'' + u_1 = (A) + (2B) \cos \phi + (-3C) \cos 2\phi \quad (32)$$

Comparing (32) to (31), we arrive at the following values for A, B, and C:

$$\begin{aligned} A &= \frac{m}{h^2} (1 + \frac{1}{2} e^2) \\ B &= \frac{me}{h^2} \\ C &= -\frac{me^2}{6h^2} \end{aligned}$$

Hence,

$$u_1 = \frac{m}{h^2} (1 + \frac{1}{2} e^2) + \frac{me}{h^2} \phi \sin \phi - \frac{me^2}{6h^2} \cos 2\phi$$

And, at last, the general solution to first order is

$$u \approx u_o + \epsilon u_1$$

$$\rightarrow u \approx u_o + \frac{\epsilon m}{h^2} [1 + e\phi \sin \phi + e^2 (\frac{1}{2} - \frac{1}{6} \cos 2\phi)] \quad (33)$$

Examining the correction term in (33), we see that the $e\phi \sin \phi$ term increases after each revolution, and hence becomes dominant. Let's substitute our solution for u_o , neglect the other terms in the correction, and obtain a simplified version of (33):

$$u \approx \frac{m}{h^2} (1 + e \cos \phi + \epsilon e \phi \sin \phi)$$

$$\rightarrow u \approx \frac{m}{h^2} [1 + e \cos[\phi(1 - \epsilon)]] \quad (34)$$

It is straightforward to check that (34) satisfies (30) to first order by differentiating and substituting. Examining (34) we see that Mercury's orbit is no longer an ellipse. It is still periodic, but the period is now given by the following equation:

$$period = \frac{2\pi}{1 - \epsilon} \approx 2\pi(1 + \epsilon) \quad (35)$$

And Mercury's perihelion precession per orbit, in relativistic units, is given by subtracting 2π from its period

$$precession \approx 2\pi\epsilon = \frac{6\pi m_r^2}{h_r^2} \quad (36)$$

where ϵ is dimensionless, m_r is the mass of the sun in relativistic units, and h_r is Mercury's angular momentum in relativistic units. Converting m_r and h_r into non-relativistic units gives a more useful form of the perihelion shift equation:

$$precession \approx \frac{6\pi G^2 m^2}{c^2 h^2} \quad (37)$$

where G is the Gravitational constant, m is the sun's mass, c is the speed of light, and h is Mercury's angular momentum per unit mass.

Lastly, let's use Kepler's second and third laws to rewrite an approximation of Mercury's perihelion precession in a more convenient form.

From Kepler's second law,

$$\frac{dA}{dt} = \frac{L}{2\mu}$$

where A is the area swept out by the orbit, L is angular momentum, and μ is reduced mass. Integrating this equation over one elliptical orbit,

$$\begin{aligned} \pi ab &= \frac{L}{2\mu} T \\ &\approx \frac{hT}{2} \end{aligned}$$

For an ellipse, $b^2 = a^2(1 - e^2)$, so

$$T^2 = \frac{4\pi^2(1 - e^2)a^4}{h^2} \quad (38)$$

From Kepler's third law,

$$T^2 = \frac{4\pi^2 a^3}{G(m + m_m)}$$

where m_m is Mercury's mass. Since m_m is very small compared to m , it can be neglected:

$$T^2 \approx \frac{4\pi^2 a^3}{Gm} \quad (39)$$

From (39), we can solve for G^2m^2

$$G^2m^2 = \frac{16\pi^4 a^6}{T^2} \left(\frac{1}{T^2} \right) \quad (40)$$

And now, substituting (38) into (40) gives

$$G^2m^2 = \frac{4\pi^2 a^2 h^2}{T^2(1 - e^2)} \quad (41)$$

Finally, combining (38) and (37) gives an equation for Mercury's relativistic perihelion precession per orbit:

$$precession = \frac{24\pi^3 a^2}{cT^2(1 - e^2)} \quad (42)$$

where a is the semimajor axis of Mercury's orbit, c is the speed of light, T is the period of Mercury's orbit, and e is the eccentricity of Mercury's orbit.

Using values from NASA's *Solar System Bodies: Mercury* website, this gives a perihelion shift of 42.9 seconds of arc per century. That Einstein's theory of General Relativity so elegantly explained Mercury's perihelion shift served as an early verification of the theory. Shortly after, during the eclipse of 21 September 1922, light deflection was observed, verifying another prediction of Einstein's theory.

A more striking example of perihelion shift than Mercury's has been observed. In 1975, binary pulsar PSR 1913+16 was discovered. The two massive sources, in very tight orbit, display a perihelion advance of 4.23 degrees per year!

Conclusions

The perihelion shift was initially observed in only Mercury's orbit since it is nearest to the sun, and hence experiences the greatest perihelion drift. Since Mercury's orbit is fairly eccentric, this perihelion shift is easily detectable (Freundlich, 40). If it were more circular, on the other hand, it would have been much more difficult to observe.

More recently, significant perihelion drift has been observed in other planets. Einstein's calculations predict a shift for Venus of 8.6 seconds of arc per century, and 8.4 seconds has been observed. Similarly, relativity predicts a shift for Earth of 3.8 seconds of arc per century, while 5.0 seconds has been observed (D'Inverno, 198).

Einstein's paper elegantly and convincingly explained the observed drift in perihelion without the need for ghost planets or asteroid belts. From Einstein's original 1916 paper, entitled *The Foundations of the General Theory of Relativity*: "Calculation gives for the planet Mercury a rotation of the orbit of 43" per century, corresponding exactly to astronomical observation (Le Verriere); for the astronomers have discovered in the motion of the perihelion of this planet, after allowing for all disturbances by other planets, an inexplicable remainder of this magnitude." (Lorentz et al., 164)

References

- Baum, Richard and Sheehan, William. *In Search of Planet Vulcan: The Ghost in Newton's Clockwork Universe*. Plenum Trade, New York. 1997.
- Callahan, James J. *The Geometry of Spacetime: An Introduction to Special and General Relativity*. Springer, New York. 1991.
- Davis, Phil (Web curator). *Solar System Bodies: Mercury*. <http://solarsystem.nasa.gov/features/planets/mercury/mercury.html>. Updated 18 March 2003.
- D'Inverno, Ray. *Introducing Einstein's Relativity*. Oxford University Press, Oxford. 1993.
- Freundlich, Erwin. *The Foundations of Einstein's Theory of Gravitation*. Translated from German by Henry L. Brose. Cambridge University Press, Cambridge. 1920.
- Price, Michael P., and Rush, William F. *Non relativistic contribution to Mercury's perihelion*. in *American Journal of Physics* 47(6). 531-534. June 1979.
- Lorentz, H. A., Einstein, A., Minkowski, H., and Weyl, H. *The principle of relativity: a collection of original memoirs on the special and general theory of relativity*. contained "The foundation of general relativity," by A. Einstein. Dover, New York. 1952.
- Roseveare, N. T. *Mercury's Perihelion from Leverriere to Einstein*. Caledon Press, Oxford. 1982.
- Stephani, Hans. *General Relativity: An introduction to the theory of the gravitational field*. Cambridge University Press, Cambridge. 1996.
- Stewart, John. *Calculus: Fourth Edition*. Brooks/Cole, Pacific Grove. 1999.
- Torretti, Roberto. *Relativity and Geometry*. Permagon Press, Oxford. 1983.