Lecture 6: Kepler's Problem

Newton once said that he stood on the shoulders of giants. Although it is said he was alluding to Hooke negatively, we can understand it in a positive way. Scientifically, the giants for Newton are Galileo and Kepler (1571–1630 AD). Galileo's contribution was already presented in previous lectures, and now we proceed to study Kepler's problem. The solution to Kepler's problem was accomplished by Newton, based on which the concept of universal gravity came into being. This is one of the most influential achievements of the human mind.

6.1 Kepler's story

Kepler summarized Tycho Brahe's observational data and proposed three laws of planetary motion, based on which Newton identified the inverse–square law of gravity. This was actually a quite complicated and interesting story.

At Kepler's time, people only knew five planets besides the Earth: Mercury, Venus, Mars, Jupiter, and Saturn. Kepler was inspired by the fact that there are five convex regular polyhedra (Platonic solids): tetrahedron, cube, octahedron, dodecahedron, and icosahedron. He proposed a one-to-one correspondence between the five planets and the five Platonic solids:

> Mercury $\leftrightarrow$ octahedron, Venus $\leftrightarrow$ icosahedron, Mars $\leftrightarrow$ dodecahedron, Jupiter $\leftrightarrow$ tetrahedron, Saturn $\leftrightarrow$ cube.

As shown in Fig. 6.1, the sphere of the Earth's orbit is set as a reference. The Earth's orbital sphere is circumscribed to a dodecahedron whose circumscribed sphere is the Mars orbit. The Mars orbital sphere is circumscribed to a tetrahedron whose circumscribed sphere is the Jupiter orbit. Furthermore, the Jupiter orbital sphere is circumscribed to a cube whose circumscribed sphere is the Saturn orbital sphere. On the other hand, the Earth's orbital sphere has an inscribed icosahedron whose inscribed sphere is the Venus orbital sphere. The Venus orbital sphere has an inscribed octahedron whose inscribed sphere is the Mercury orbital sphere.

Kepler wrote his theory in a book and sent it to Tycho Brahe, who spent his life observing planetary motion and accumulated enormous data. (He also sent it to Galileo, but Galileo did not respond.) Tycho Brahe welcomed and hired Kepler as his assistant. After Tycho Brahe's passing away, Kepler spent 20 years analyzing Tycho's data and hoped to verify his model. To his disappointment, Kepler failed to fit Mars's orbit by a circle. Finally, in 1605, he reluctantly recognized that planetary orbits are ellipses, which is Kepler's first law. The largest eccentricity of planetary orbits is that of Mercury, which is 0.2. Eccentricities for others are not large: $e_{\text{mars}} = 0.09$, $e_{\text{jupiter}} = 0.05$, $e_{\text{saturn}} = 0.05$, $e_{\text{uranus}} = 0.05$, $e_{\text{earth}} = 0.02$, $e_{\text{neptune}} = 0.008$, $e_{\text{venus}} = 0.007$, and $e_{\text{moon}} = 0.05$, which are good approximations to circles. After further studies, he published Kepler's second law, i.e., the area law, and the third law which relates the radii and periods of different orbits. Nevertheless, Kepler did not feel pride in these discoveries since an ellipse is not as "perfect" as a circle.

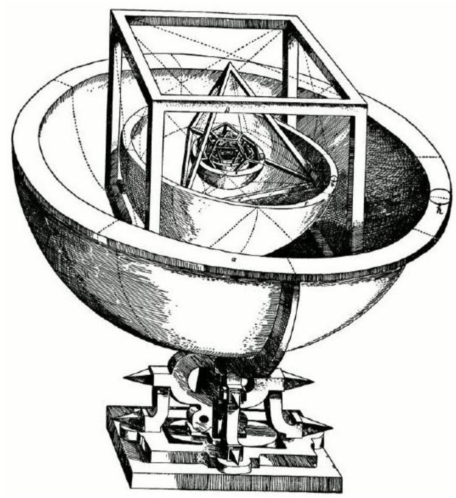


Figure 6.1 (from Wikipedia). Kepler's model of the solar system with suggested correspondences: Mercury $\leftrightarrow$ octahedron, Venus $\leftrightarrow$ icosahedron, Mars $\leftrightarrow$ dodecahedron, Jupiter $\leftrightarrow$ tetrahedron, Saturn $\leftrightarrow$ cube.

Hence, Kepler discovered the laws of planetary motion in a dramatic way. He was studying the wrong problem but arrived at the correct answer. Kepler's laws served as the motivation and foundation for Newton's theory of gravity.

6.2 Kepler's three laws

6.2.1 The first law

The planet's orbit is a planar ellipse, and the Sun lies at one of the foci of the ellipse.

Kepler's first law is actually a very strong statement. Generally speaking, a planet moves in three-dimensional (3D) space, but this law states that it is planar. A planar motion does not always form a closed orbit, but Kepler's first law asserts that it is closed and periodic. Copernicus, influenced by the aesthetic philosophy of the Greeks, thought that a planet's orbit should be a circle. Nevertheless, Kepler figured out that, in general, a planetary orbit is an ellipse. A circular orbit is a special case in which the Sun lies at the center.

6.2.2 The second law

Since the general orbit of a planet is elliptical rather than circular, the motion at each point on the orbit is different. To connect the motion along the orbit is precisely what Kepler's second law tells us:

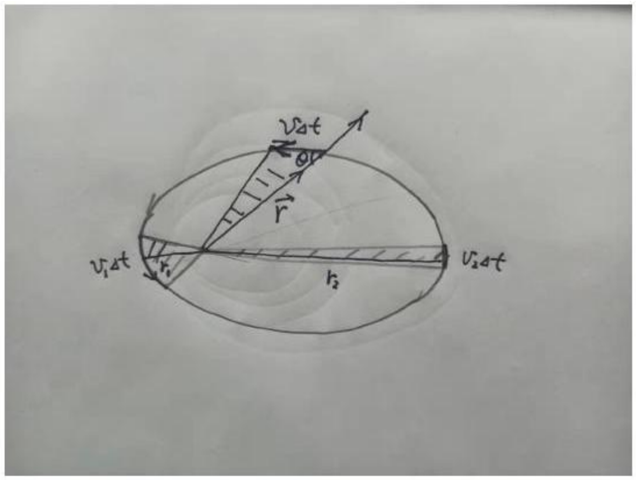


Figure 6.2 Kepler's second law. The area that the Earth–Sun line sweeps in a unit time Δt is proportional to the Earth's angular momentum.

The areas swept by the line connecting the Sun and a planet are equal in equal time intervals.

Assume that within a time interval Δt, the line running from the Sun to a planet sweeps an area ΔS. Kepler's second law states that $\Delta S/\Delta t$ is a constant.

Consider two special positions in a planetary orbit: the apogee (the farthest point from the Sun) and the perigee (the nearest point from the Sun). Denote the displacement vectors relative to the Sun at the apogee and perigee as $\mathbf{r}_a$ and $\mathbf{r}_b$, respectively. Correspondingly, the velocities are $\mathbf{v}_a$ and $\mathbf{v}_b$, then $\mathbf{v}_a \perp \mathbf{r}_a$ and $\mathbf{v}_b \perp \mathbf{r}_b$.

The arc length traveled around the apogee during time Δt is $\Delta s = v_1 \Delta t$, and the area swept at the apogee is $\Delta S = \frac{1}{2} r_1 \Delta s = \frac{1}{2} r_1 v_1 \Delta t$. Similarly, the same area should be swept during Δt around the perigee: $\Delta S = \frac{1}{2} r_2 v_2 \Delta t$. Then

$$m r_1 v_1 = m r_2 v_2, \tag{6.2}$$

where the planet mass m is multiplied. The product of linear momentum and displacement is actually the angular momentum, whose precise definition will be given later. It means that the angular momentum L_1 at the apogee equals L_2 at the perigee. Since $r_1 > r_2$, we have $v_1 < v_2$.

If the planet is at a general point in the orbit, then $\mathbf{v}$ and $\mathbf{r}$ are no longer perpendicular. Their relative angle is denoted by θ. Then the area swept during a small time interval Δt is

$$\Delta S = \tfrac{1}{2}\, r v \sin\theta, \tag{6.3}$$

and, with an area direction,

$$\boldsymbol{\Delta S} = \tfrac{1}{2}\, \mathbf{r} \times \mathbf{v}.$$

Hence it means angular momentum conservation,

$$\mathbf{L} = m\,\mathbf{r} \times \mathbf{v}, \tag{6.4}$$

which does not change with time.
Angular momentum conservation is a fundamental law of nature as a consequence of spatial isotropy, which will be explained later. Simply put, because the gravitational force passes through the Sun's center, it does not generate torque to change the angular momentum.

6.2.3 The third law

Different initial conditions can lead to different orbits. Kepler's third law connects different orbits:
For different orbits, the ratio between the cube of the semi-major axis and the square of the period is a constant.

For an elliptic orbit (Fig. 6.3), the origin is set at the focus. The y-axis intersects the ellipse and cuts a chord (the latus rectum), whose half-length is denoted as h. The x-axis is the major axis intersecting the ellipse, whose half-length is a.

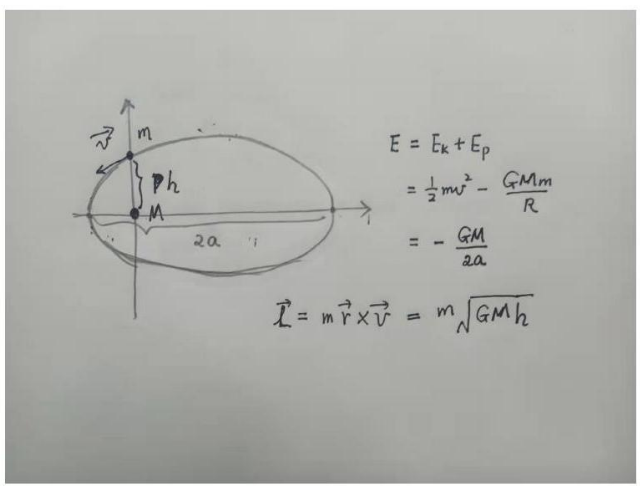


Figure 6.3 For an elliptic orbit, the total energy is determined by the semi-major axis a as $E = -\dfrac{GMm}{2a}$. The angular momentum is determined by the half-length of the latus rectum h as $l = m\sqrt{GMh} = mh\sqrt{GM/h}$.

Kepler's third law implies the inverse–square law: Consider the special case of a circular orbit, then $a = R$. A simple dimensional (scaling) analysis is as follows. Due to the nature of periodic motion, the acceleration scales as

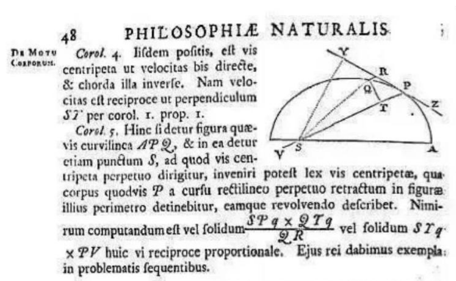
$F/m \sim v/T \sim R/T^2$. According to Kepler's third law $T^2 \sim R^3$, we arrive at

$$F \sim \frac{1}{R^2}. \tag{6.5}$$

Certainly, the above argument should not be viewed as a proof—rather as a motivation for further exploration.

6.3 Solution to Kepler's problem by the geometric method

Kepler's laws are phenomenological laws based on astronomical observations, whose simplicity and beauty are already impressive. Such beautiful laws cannot be a coincidence, which stimulated physicists including Isaac Newton to explore the underlying law of gravity. By assuming the inverse–square law of gravity, Newton derived Kepler's three laws of planetary motion. Although in modern formalism this can be done concisely via calculus—and indeed solving motion under forces was a main motivation for Newton to invent his fluxional calculus—



48 PHILOSOPHIÆ NATURALIS

De Motu Corporum.

Corol. 4. Iisdem positis, est vis centripeta ut velocitas bis directe, & chorda illa inverse. Nam velocitas est reciproce ut perpendiculum *ST* per corol. 1. prop. 1.

Corol. 5. Hinc si detur figura quævis curvilinea *APQ*, & in ea detur etiam punctum *S*, ad quod vis centripeta perpetuo dirigitur, inveniri potest lex vis centripetæ, qua corpus quodvis *P* a cursu rectilineo perpetuo retractum in figuræ illius perimetro detinebitur, eamque revolvendo describet. Nimirum computandum est vel solidum $\frac{SPq \times QTq}{QR}$ vel solidum $STq \times PV$ huic vi reciproce proportionale. Ejus rei dabimus exempla in problematis sequentibus.

Figure 6.4 A page of Newton's *Principia*.

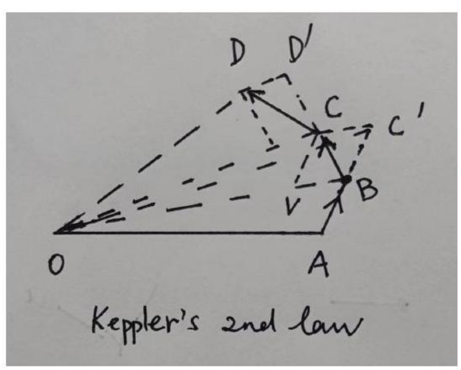


Figure 6.5 The geometric proof of Kepler's second law.

interestingly, in the *Principia* Newton did not use calculus. Rather, he adopted the style of Euclid's *Elements* using geometric methods. At that time, the mathematical foundation of calculus was not rigorously established (it would be in the 19th century). Newton wanted to avoid criticism of infinitesimals that might hinder recognition of his theory of gravity.

6.3.1 Proof of Kepler's second law

Assume gravity is a central force field, which is sufficient for the proof of Kepler's second law. Historically, it was important to ask: what drives the planets to move around? Some believed that planets were propelled by invisible angels' beating wings, implying a tangential driving force. However, we will show the force is centripetal rather than tangential.

The geometric picture to prove Kepler's second law is presented in Fig. (6.5). Suppose the Sun is located at O and the planet starts from A. Within a small time interval Δt (first-order infinitesimal), it moves to B. We use a short line segment AB to approximate its trajectory, with an error of second order. The velocity is $\mathbf{v}_{AB} = \dfrac{\mathbf{AB}}{\Delta t}$. If there were no gravity, in the next time interval Δt, the planet would continue straight toward C', such that $\mathbf{AB} = \mathbf{BC'}$. It is obvious that

$$S_{\triangle OAB} = S_{\triangle OBC'}. \tag{6.6}$$

However, gravity pulls the planet back. The attraction is along the direction of $\mathbf{OB}$, such that the planet is pulled from C' to C, hence $\mathbf{C'C} \parallel OB$, and $\mathbf{v}_{BC} = \dfrac{\mathbf{BC}}{\Delta t}$. It is easy to show that

$$S_{\triangle OBC} = S_{\triangle OBC'}, \tag{6.7}$$

since they share the same base and height. Hence within the same time interval Δt, the areas swept by the Sun–planet line are equal,

$$S_{\triangle OAB} = S_{\triangle OBC}. \tag{6.8}$$

Repeating this process—after consecutive intervals Δt the planet arrives at $D, E, \ldots$—gives

$$S_{\triangle OAB} = S_{\triangle OBC} = S_{\triangle OCD} = S_{\triangle ODE} = \ldots \tag{6.9}$$

Moreover, one can see that the trajectory $ABCD\ldots$ is a planar curve. This completes the proof of Kepler's second law.

The area rate is $\dfrac{\Delta S}{\Delta t} = \dfrac{1}{2m}\dfrac{mr^2\Delta\theta}{\Delta t} = \dfrac{L}{2m}$, where L is the magnitude of the orbital angular momentum. If adding direction to the area, we arrive at

$$2m\,\frac{\Delta \mathbf{S}}{\Delta t} = \frac{mr^2\Delta\theta}{\Delta t} = \mathbf{L}. \tag{6.10}$$

6.3.2 Proof of Kepler's first law

Next we prove Kepler's first law—the trajectory is generally an ellipse—following the method presented in "Feynman's Lost Lecture: The Motion of Planets Around the Sun."

Kepler's second law alone does not ensure a closed orbit. For simplicity, we temporarily assume closedness and justify it later. Assume gravity is a central inverse–square law,

$$\mathbf{F} = -\frac{GMm}{r^2}\,\mathbf{e}_r, \tag{6.11}$$

where M is the Sun's mass and m is the planet's mass.

Newton realized the inverse–square law by comparing the Moon's orbital motion with free fall on the ground. Ancient astronomers measured the Earth–Moon distance compared to the Earth's radius via parallax. The ratios d_{em}/r_e by Ptolemy, Huygens, and Tycho Brahe are close: $59, 60, 60.5$. The falling distance in one second on Earth's surface is about 5 m. How much is the Moon's "falling distance" in one second? Without Earth's gravity, the Moon would follow the tangent with distance $\Delta x = v\Delta t$. To bend back to a circle, the Moon must "fall" a distance s satisfying

$$\begin{aligned} \frac{v\Delta t}{s} &= \frac{2d_{em}}{v\Delta t}, \\ s &= \frac{1}{2}\frac{v^2}{d_{em}}\,t^2 = \frac{1}{2}\omega^2 d_{em}\,t^2 = \frac{2\pi^2}{T^2}\,d_{em}\,t^2. \end{aligned} \tag{6.12}$$

Plugging in $T = 27.3$ days and $d_{em} = 60\,r_e$, we arrive at $s \approx 1.36\,\mathrm{mm}$, which is about $1/3676 \approx 1/60^2$ of the terrestrial 5 m. This crude estimate accurately suggests the inverse–square relation, building Newton's confidence that satellites and planets obey the same law. Thus the concept of *universal gravity* was born.

Now prove the elliptic orbit. In Fig. (6.6A), the Sun is at F. Start from the perigee A on the orbit, and mark points $B, C, \ldots$ such that $\angle BFA = \angle CFB = \cdots = \Delta\theta$ with small $\Delta\theta$. The radii are $r_A, r_B, r_C, \ldots$ and the velocities $\mathbf{v}_A, \mathbf{v}_B, \mathbf{v}_C, \ldots$.

Compare $\Delta\mathbf{v}_{BA} = \mathbf{v}_B - \mathbf{v}_A$ and $\Delta\mathbf{v}_{CB} = \mathbf{v}_C - \mathbf{v}_B$. The time to sweep $\triangle BOA$ is

$$\Delta t = \frac{S_{\triangle BOA}}{\Delta S/\Delta t} = \frac{\frac{1}{2}r_A^2\Delta\theta}{L/(2m)} = \frac{m r_A^2 \Delta\theta}{L}. \tag{6.13}$$

Hence, by Newton's second law,

$$\Delta\mathbf{v}_{BA} = \frac{\mathbf{F}\,\Delta t}{m} = -\frac{GMm\,\Delta\theta}{L}\,\mathbf{e}_A, \tag{6.14}$$

which is opposite the radial direction. The radius dependence cancels. Since $\Delta\theta$ is fixed for each small triangle, $\Delta\mathbf{v}$ has fixed magnitude while its direction changes by $\Delta\theta$ each step; thus the endpoints of the velocity vectors lie on a circle in velocity space. From Eq. (6.14), the tangential direction of the velocity circle reflects the displacement direction in real space, i.e., the velocity-space motion is dual to the real-space motion.

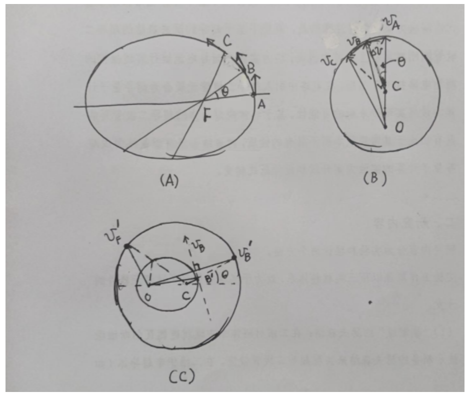


Figure 6.6 The geometric proof of the elliptical orbit of planet motion. A) F is the location of the Sun and A is the perigee. The angle between two neighboring radii is fixed at $\Delta\theta$. B) The velocity vectors $\mathbf{v}_A, \mathbf{v}_B, \mathbf{v}_C \ldots$ are plotted from the origin. Their ending points span a circle with the radius $v = \frac{GMm}{L}$, whose center is denoted as C. C) Rotate the velocity circle by 90°. Use the locations of O and C as the two foci, and v as the major axis length to construct an ellipse. Such an ellipse is similar to the planet orbit.

The circle's center is generally not at the origin O; denote it C. By planar geometry, the angle of $\Delta\mathbf{v}$ about C equals its incremental direction change each step. The circle radius is

$$v_r = \frac{\Delta v}{\Delta\theta} = \frac{GMm}{L}. \tag{6.15}$$

The time evolution of $\mathbf{v}$ is periodic; does that imply periodic real-space motion? Not always. However, by the geometric correspondence between real and velocity spaces, the answer here is yes: when the velocity vector returns to $\mathbf{v}_A$ with the same tangent direction, the real-space direction is again FA. To reconstruct the trajectory from the velocity circle, Feynman's trick is to rotate the velocity diagram by 90° (Fig. 6.6C) and reconstruct the real-space curve in the same figure. Then $\mathbf{v}'_B$ is perpendicular to $\mathbf{v}_B$ (the tangent to the real path). To locate the tangent point, construct the bisector of Ov'_B (parallel to $\mathbf{v}_B$). As $\mathbf{v}'_B$ runs around the circle, the envelope of these bisectors forms an ellipse with foci at O and C, and major axis equal to the circle radius.

To see this, recall the optical property of ellipses: a ray from one focus reflects to the other. Connect Cv'_B. The bisector line of $\mathbf{v}_B$ intersects Cv'_B at B'. Connect OB' and CB'. Since O and v'_B are mirror images about the bisector, the ray OB' reflects to C, and $B'v'_B$ is the image of OB'. Hence

$$|CB'| + |OB'| = |Cv'_B| = \frac{GMm}{L}, \tag{6.16}$$

so B' lies on the ellipse. Moreover the bisector is tangent to the ellipse: for any other point P on this bisector,

$$|CP| + |OP| = |CP| + |v'_B P| > |Cv'_B|. \tag{6.17}$$

We now build a point-to-point mapping between the constructed ellipse (Fig. 6.6C) and the planetary trajectory by using polar coordinates. For the real-space trajectory, set the origin at F; for the constructed ellipse, set the origin at the focus C. At each polar angle θ, the tangent lines on the two curves are parallel. For a curve $\mathbf{r}(\theta) = r\,\mathbf{e}_r$,

$$\frac{d\mathbf{r}}{dt} = \frac{dr}{d\theta}\,\mathbf{e}_r - r\,\mathbf{e}_\theta. \tag{6.18}$$

For the two curves,

$$\mathbf{r}_1 = r_1(\theta)\,\mathbf{e}_r, \qquad \mathbf{r}_2 = r_2(\theta)\,\mathbf{e}_r, \tag{6.19}$$

parallel tangents at the same θ imply

$$\frac{1}{r_1}\frac{dr_1}{d\theta} = \frac{1}{r_2}\frac{dr_2}{d\theta}, \tag{6.20}$$

hence

$$\mathbf{r}_1(\theta) = C_0\,\mathbf{r}_2(\theta), \tag{6.21}$$

for a constant C_0. Therefore the real trajectory is similar to the constructed ellipse—i.e., it is an ellipse.

6.4 Comment on Kepler's third law

The geometric proof is beautiful and reveals the nature of planetary motion. Next we use calculus for further exploration.

Kepler's third law can be shown by a scaling method. Suppose $\mathbf{r}(t)$ solves

$$\frac{d^2\mathbf{r}(t)}{dt^2} = -\frac{GM}{r^2}\,\mathbf{e}_r. \tag{6.22}$$

Perform the scaling

$$\mathbf{r}^s(t) = \lambda_1\,\mathbf{r}(\lambda_2 t). \tag{6.23}$$

It is easy to show

$$\frac{d^2\mathbf{r}^s(t)}{dt^2} + \frac{GM}{(r^s)^2}\,\mathbf{e}_r = \lambda_1\lambda_2^2\frac{d^2\mathbf{r}(t)}{dt^2} + \lambda_1^{-2}\frac{GM}{(r)^2}\,\mathbf{e}_r = 0, \tag{6.24}$$

provided

$$\lambda_2^2\lambda_1^3 = 1. \tag{6.25}$$

Thus the spatial size L of the orbit and the period T obey

$$L^3/T^2 = \text{const}. \tag{6.26}$$

Actually, Kepler's third law is stronger: the length scale relevant for the energy is the semi-major axis alone, independent of the minor axis.