Chapter 2
General Orbit Background

Unlike our first topic, attitude, *orbit* is a word for which most people probably could conjure up a reasonable visual image. They might envision the Earth and the other planets circling the Sun, or the Moon or a spacecraft circling the Earth. But beyond that, their perceptions can get kind of fuzzy, and sometimes pretty inaccurate. For example, our experiences on the surface of the Earth lead us to assume that if you want to catch up with something you should move faster. For example, if you are trailing another car and want to overtake it, your natural reaction is to lean on the accelerator to increase your speed. But when orbiting the Earth, this would be exactly the wrong strategy. As Earth-orbiting objects move from perigee (their point of closest approach to the Earth) to apogee (their point of furthest separation from the Earth), they exchange kinetic energy for potential energy, and as a result fly slower at higher altitudes than at lower altitudes. This counterintuitive aspect of orbital motion is also relevant to astronauts performing rendezvous operations with other spacecraft. In orbit, the effect of the equivalent action (relative to the car chase scenario), firing a thruster in the direction of your motion, will be to raise your orbital altitude, causing you to move slower, and thereby losing ground with your rendezvous partner. What you need to do instead is to fire your thruster opposite to your direction of motion, lowering your orbital altitude, causing you to move faster, and resulting in an eventual overtaking of your quarry. So even though we think of the classical dynamics engendered by Newton's law of gravitation over 300 hundred years ago as being sensible and down to Earth (sorry), the motion of objects in the presence of gravitational fields is not always what you'd expect based on your interactions with the Earth's gravitational field in your daily lives.

2.1 Historical Perspective

For everyday life spent at or near the surface of the Earth, dropped objects appear to fall at a constant acceleration of about $9.8\,\text{m/s}^2$ (or $32\,\text{feet/s}^2$ if you prefer English units). This value is accurate to 1% over a spectrum of locations ranging from the

H.L. Hallock et al., *ACS Without an Attitude*, NASA Monographs
in Systems and Software Engineering, DOI 10.1007/978-1-4471-7325-0_2

Earth's poles, to its equatorial regions, to points located up to 20 miles above the equator. So Galileo's 400 year old observation (that, in the absence of air resistance, two dropped objects of different mass experience the same acceleration independent of altitude) has been something you could count on pretty reliably (Aristotelian ivory tower philosophizing to the contrary) for most of human existence. However, when people started looking up at the heavens in a serious, quantitative fashion, a handful of very smart people began to realize that if things worked upstairs (in the heavens) the way things appeared to work downstairs (on the Earth), Galileo's law of falling bodies could not be a reliable guide to explaining how celestial bodies interacted with each other, even though they had an instinctive feeling that the force that caused the Sun and planets to follow predictable paths in the sky relative to the Earth was the same force that caused a dropped hammer to fall and hit your big toe. The force that causes your hammer to hit your thumbnail instead of a steel nail is a different one (see poltergeists or gremlins).

Actually, the idea that things should work upstairs the same way they work downstairs would have been considered downright silly, possibly even depraved, 1,600 years ago. Even someone like Aristotle, whose work in the biological sciences was based soundly on observation and measurement, took at times what we would view today as almost a theologically based view of the physical sciences. To Aristotle, and nearly everybody else from his age (Aristarchus of Samos, the birthplace of Pythagoras, being an exception), the real world here on the Earth's surface was (probably with good reason) seen as messy, complicated, and inelegant. By contrast, the heavens were perfect, simple, and elegant. Since the circle was for the Greeks an expression of perfection, it followed that objects that lived in the perfect heavens must follow paths composed of perfect circles.

The Greeks' obstinate commitment to heavenly bodies only being allowed to move in immutable circles (because that's the way things are intended to be as opposed to the way we find them to be, also a not unfamiliar modern day phenomenon), coupled with an unshakable belief that the Earth was the center of the Universe, shackled them to a cumbersome system that required that the known planets of the time follow paths composed of circles within circles. The coming of the Renaissance began a revolution (yes, that pun was intended) from a conceptual standpoint that promoted the Sun as the central body (supported by Copernicus' measurements of planetary motions made about 500 years ago), displacing Aristotle and Company's Earth-centered model. Galileo further raised the discussion (about 400 years ago) from metaphysics to actual physics when he aimed his custom-built telescope at Jupiter and observed its moons in orbit, suggesting that just as that planet's moons orbited the planet, the planets themselves could be orbiting the Sun. However, the key to burying the Earth-centered model was provided by an Odd Couple that together embodied the fundamentals on which modern science is based, namely careful quantitative measurements, elegant mathematics and models, and uncompromising intellectual honesty.

The Odd Couple in question, Tycho Brahe and Johannes Kepler, made their contributions around the same time as Galileo, but worked in a politically safer location, Denmark. Brahe was an expert at making positional measurements (providing major improvements over those produced by Copernicus) and accurately recording them, but lacked theoretical vision and was awful at mathematics. By contrast, Kepler's optical vision was terrible, but he could see meaning in Brahe's mountain of data and had the extraordinary mathematical skills required to organize the data within a working model. After several years of empirical efforts with Brahe's Mars positions, Kepler concluded that an ellipse was the simplest curve to which Brahe's data could be fit to the level of accuracy demanded by Brahe's highly precise observations. Kepler summarized the impact of his discovery in his Three Laws,[1] namely,

1. Each planet follows an elliptical orbit with the Sun at one focus of the ellipse.
2. A line connecting the Sun and a planet sweeps out equal areas in equal times.
3. The square of a planet's orbital period is proportional to the cube of the major axis of the ellipse.

Kepler's laws and associated model finally ended the science controversy (though not the religious or political ones) regarding who orbited whom, and made quantitative confirmed predictions regarding the relative positions and rates of the objects. On the downside, Kepler's work in no way explained why any of these neat facts should be true. That's where Isaac Newton enters the story. When plague closed down Cambridge University for two years (1665–1667), Newton spent his free time creating the basis for differential calculus and developing his laws of motion and gravitation (not bad for a vacation). But because he had some initial problems explaining details of the Moon's motion, Newton did not publish his *Principia*[2] (describing Newton's Laws of Motion and Gravitation) until 1687. In the *Principia*, Newton's Law of Gravitation (see Eq. 2.1) appears in print for the first time. As applied to the gravitational interaction between two idealized point masses, it states that

1. The magnitude of the force experienced by one of the objects is proportional to the product of the two masses and inversely proportional to the square of their separation distance.
2. The direction of the force on one of the objects is towards the other object and is along the line connecting the centers of the two objects.

$$\mathbf{F}_{12} = \frac{G m_1 m_2}{|\mathbf{R}_{12}|^3}\mathbf{R}_{12} \tag{2.1}$$

where

$\mathbf{F}_{12}$ = the force exerted on object 1 by object 2 (newtons)
G = gravitational constant = $6.67428 \times 10^{-11}\ \mathrm{m^3kg^{-1}s^{-2}}$
m_1 = mass of object 1 (kg)
m_2 = mass of object 2 (kg)

[1] See, e.g., http://www-spof.gsfc.nasa.gov/stargaze/Kep3laws.htm.

[2] I. Newton, Philosophiae Naturalis Principia Mathematica (Mathematical Principles of Natural Philosophy), 1687; cf. the translation by I. Cohen and A. Whitman, University of California Press, Berkeley, Los Angeles, London, 1999.

$\mathbf{R}_{12}$ = position vector from object 1 to object 2 (m)
$|\mathbf{R}_{12}|$ = magnitude of position vector from object 1 to object 2 (m)

Newton further showed that the external gravitational force exerted by an object with a spherically symmetric mass distribution is equivalent to that which would be exerted if all of the mass were concentrated at the object's center. (The demonstration is actually fairly straightforward, but we'll leave it to bored readers who enjoy integrations involving spherical coordinates.) In real world applications, celestial objects are never such perfect spheres, and the problem to be solved usually involves more than two objects (often with non-gravitational perturbing forces thrown into the mix as well). But Newton's Law of Gravitation still is the key to solving the classic problem of orbital motion. To deal with non-spherical masses, one can integrate over the object's mass distribution to obtain the influence of the entire celestial object. Similarly, if more than two objects are interacting gravitationally with each other, just sum the effects of all the other bodies on the one you're looking at. However, because the other bodies will interact with each other, no closed-form solution exists for the N-body problem once N is 3 or larger, hence Newton's problems with the Moon, primarily influenced by the Earth and Sun. For such problems, approximations must be used and the analytical approach used is to integrate the equations of motion–an easy thing to do in our computer age, but a far more difficult problem in Newton's time. Such approximation methods are particularly appropriate once the problem is further complicated by including non-gravitational perturbing influences such as atmospheric drag, solar radiation pressure, and propulsive forces along with gravity when determining the orbital motion of a spacecraft (as opposed to the motion of planets or moons).

2.2 Orbital Shapes

For the time being, let's not worry about the details regarding how you predict an object's orbit; that will be covered in a later chapter (or as we often say in bullet slide presentations, "Great question! Dave will talk about that later today."). Instead, let's consider the best way to describe an object's orbit assuming you have all the information required to define it. First, you have to know where the object is at a given instant in time, say right now. An easy way to define an object's position is by specifying the three-dimensional vector (magnitude and direction) that connects the object to the origin of a three-dimensional coordinate frame. But when discussing an object's orbit, you need to know both where the object is right now as well as where the object is going to be in the future (or where it came from in the past if your interests lean more towards history than current events). In general, if you want to predict an object's future position (relative to its current position) you at least need to know in what direction it's heading and how fast it's going (i.e., its velocity vector relative to its current position). So if you know an object's position and velocity vector relative to the coordinate frame's origin at a given instant in time, you can guess at what its position will be at some later (or earlier) point in time. If the new time is very close to the original time, you can compute the new position pretty accurately just by multiplying the velocity vector by the time difference and adding it to the starting position (a crude integration over time).

The problem with this solution is that if the object is following a wildly varying path, the error in the position prediction will rise quickly as the time difference increases, or to put it another way, the original information will go stale very rapidly. However, when examining the orbital behavior of two (and only two) spherically symmetric masses acting under the influence of an inverse square law force like gravity, the orbital motion of one of the objects relative to the other will always be a conic section (i.e., circle, ellipse, parabola, or hyperbola). In such a case, the predictive power of those seven pieces of information (time, three position vector coordinates, and three velocity vector coordinates) will be extremely good. For all such trajectories, knowing an object's position and velocity vectors at a given instant in time is sufficient information to predict its motion at all other instants in time, future or past, along that curve.

As a specific example of a conic section, let's imagine an object in an elliptical orbit (recall that Kepler demonstrated that the planetary orbits are ellipses with the Sun at a focus) and ask how many pieces of information you need to define all points on the ellipse as a function of time. First, you need to define the size and shape of the ellipse. Figure 2.1 illustrates how this is done. Given a pair of points, the foci of the ellipse ($F1$ and $F2$ in the figure), and a distance d, an ellipse is defined as the set of points $\{X\}$ for which the sum of lengths $F1$-to-X plus $F2$-to-X is equal to d. The major axis is the longest line segment connecting two points on the ellipse; in Fig. 2.1 it connects points A and B, and passes through both foci. Half the major axis is called the semi-major axis, symbolized by a. The ellipse minor axis bisects the major axis at right angles, connecting points C and D in the figure. Half the minor axis is called the semi-minor axis, symbolized by b. The semi-major and semi-minor axes of an ellipse define its contours in the same manner that the radius of a circle defines its contours. Indeed, a circle is the special case of an ellipse in which the values of the semi-major and semi-minor axes are identical (equal to the circle radius), with the distance d being the circle diameter.

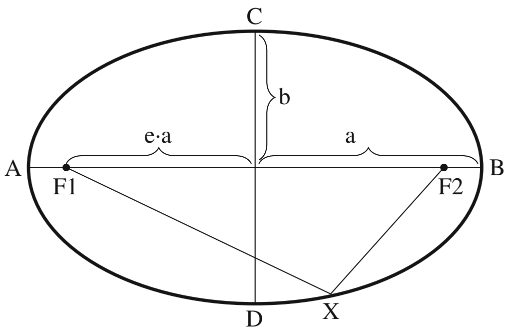

Fig. 2.1 Parameters that shape an elliptical orbit

Another way of describing the shape of the ellipse is via the eccentricity, a unitless parameter that is related to the semi-major and semi-minor axes via Eq. 2.2:

$$e = \frac{\sqrt{a^2 - b^2}}{a} \tag{2.2}$$

where

e = eccentricity (unitless)
a = semimajor axis
b = semiminor axis

Before moving on (in the next section) with parameters that characterize the orientation of an ellipse in space, let's explore some physics associated with various elliptical orbits. Let's consider a sequence of ellipses defined by a small satellite orbiting a massive central body (at $F1$), with the satellite at some initial time located a distance R_0 from the massive body and having velocity $\mathbf{V}_0$ perpendicular to the line segment back to the central body. It is important for this discussion that $\mathbf{V}_0$ is perpendicular to the line back to the central body; it implies that at our initial time the satellite is at either a point of closest approach to or farthest retreat from the central body (perigee or apogee, if the central body is Earth). As it turns out, if $V_0^2 = GM/R_0$, the orbit is a circle. If $V_0^2 < GM/R_0$ (but not zero), then the orbit is an ellipse with $0 < e < 1$, and the initial position is one of furthest retreat. The satellite falls in towards the central body, speeding up as it does so. Because the satellite's initial velocity was perpendicular to the line back to the central body, the satellite at first moves away from that line. Eventually, however, gravity pulls it back towards the line (after it passes the minor axis line), causing the satellite to whip around the central body. After passing the point of closest approach, the satellite rises back up to its initial position. If $V_0^2 = 0$, then we're just dropping the object from rest, and it falls into the central body. Although this orbit is still formally an ellipse, in this limit it is a simple line segment connecting the two foci ($F1$ being the location of the central body, and $F2$ the initial position of the satellite). The minor axis has shrunk to 0, and the eccentricity for the line segment is 1. (Well, strictly speaking, this discussion for the "$V_0^2 = 0$" case should really be taken as the limiting case as V_0^2 approaches infinitesimally close to 0. Since we're dealing with a point mass approximation for the central body, we really can't deal with the infinite forces that would apply as the orbiting body encounters the central point.)

Now let's go in the other direction, one of increasing V_0^2. For $GM/R_0 < V_0^2 < 2GM/R_0$, we again have an ellipse with e increasing from 0 to just under 1. Our initial position now is one of closest approach, the initial velocity being as large as it will ever be. As the body moves away from its point of closest approach, the gravitational force from the central body slows it down until it reaches its point of furthest retreat, after which it falls back to its initial position. For $V_0^2 = 2GM/R_0$, the satellite never reaches its point of further retreat. Indeed, for our initial conditions, the velocity $(2GM/R_0)^{1/2}$ is just enough for the satellite to escape from the central body with zero velocity at infinite distance. Although e was heading towards 1 in this limit, Eq. 2.2 is only well defined for finite a and b. Both a and b become infinite at $V_0^2 = 2GM/R_0$. Suffice it to say that the orbit is no longer an ellipse, but rather has become a parabola with arms that continue to extend outwards forever. (You can imagine that the location of $F2$, as well as the lengths a and b, have been pulled out to infinity.)

If V_0^2 is increased further, the arms of the orbit pull apart yet further and the orbit becomes that other conic section, the hyperbola. "Hyperbola" comes from the same Greek root as "hyperbole", meaning excess; satellites in hyperbolic orbits may be considered as having excess (greater than zero) energy at infinite distance. Geometrically, with appropriate redefinitions of parameters a, b, and e that we won't go into here, hyperbolas have eccentricity $e > 1$. They accurately model the orbits of uninvited (and usually undesired) guests to our Solar System, as they (hopefully) follow near-miss trajectories relative to the permanent residents. And for departing residents, parabolas and hyperbolas describe the paths followed by objects that have reached escape velocity and are departing from the central body to which they had previously been bound.

2.3 Specifying the Orbit's Orientation in Inertial Space

Getting back to ellipses and how to specify their orientation in space, we now need to specify how the plane defined by the ellipse is oriented relative to a coordinate frame, e.g., GCI for an Earth orbiting satellite (which we'll assume for the remainder of this section). Imagine the ellipse superimposed on the GCI frame's equatorial plane with the minor axis aligned with the vernal equinox. Now tilt the ellipse about a line that runs through the center of the Earth and parallel to the minor axis. The line about which the tilt is done is called the *line of nodes* (only coincidentally parallel to the minor axis in this example). The tilt angle is called the *inclination* and is a right-handed rotation about the line of nodes. The inclination's range of rotation is 0–180°, where 90° will yield a polar orbit and inclinations greater than 90° yield retrograde orbits. But placing the ellipse's minor axis at a special orientation relative to the vernal equinox was, well special. We can generalize further by rotating the ellipse's plane about the GCI z-axis. This angle is called the *right ascension of the orbit's ascending node*, where the *ascending node* is the point on the orbital ellipse where the object rises from the GCI frame's southern hemisphere and enters the northern hemisphere. (The point where the object moves from the northern to southern hemisphere is called the *descending node*, hence *line of nodes* as the term for the line connecting the nodes.) The right ascension of the ascending node is measured from the vernal equinox (the GCI x-axis) and has a range from 0 to 360°. Note that for an equatorial orbit (inclination equal to zero) there is no orbital ascending node or descending node, so the right ascension of the ascending node is undefined. This is an orbital equivalent of the singularity situation we talked about in Chap. 1 when mentioning that the right ascension Euler angle is undefined when the declination is ±90°. Figure 2.2 illustrates how these two angles (right ascension of ascending node

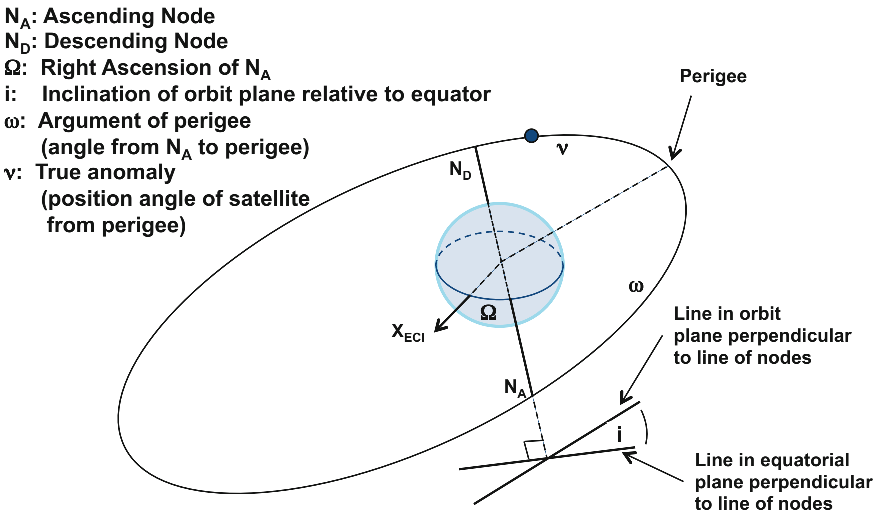

Fig. 2.2 Parameters that orient an elliptical orbit

and inclination) define the orientation of orbital plane in space. It also illustrates two other important angles, to be described in the next paragraph and next section.

Another parameter is still needed for full generality, because there's no requirement that the ellipse's major axis be aligned so that those points on the ellipse (at the ends of the major axis) take on the largest positive and negative Z-component values. Instead, we can envision a rotation about the orbital plane normal vector that places the major axis in a more general configuration within the (now) fixed orbital plane. For example, a 90° rotation about orbit normal would interchange major axis with minor axis. This angle of rotation within the orbit plane is called the *argument of perigee* for Earth orbits, or in general is called the *argument of periapsis* (again, see Fig. 2.2). By convention, it is the angle measured in the orbital plane from the orbit ascending node to (for Earth orbits) perigee, the object's closest approach to the Earth during an orbital period. The range of the argument of perigee is 0–360°. As with the right ascension of the ascending node, the argument of perigee is undefined for equatorial orbits (i.e., inclination equal to zero) because the ascending node (from which it is measured) is undefined. Although the two angles are undefined when inclination is exactly zero, you can consider any given such orbit as the limiting case for a family of orbits with non-zero but infinitesimal inclination, all with different ascending node right ascension and argument of perigee, but all having the same value for the sum of the two angles, as long as the orbit is not circular.

2.4 The Location of the Spacecraft in the Orbit

So now with five parameters (semimajor axis, eccentricity, inclination, right ascension of the ascending node, and argument of perigee), we're able to fix the orbit ellipse's size and shape, orbit plane orientation, and orbit orientation within the orbit plane. But there's one thing we need to know that has not yet been mentioned, namely the position of the orbiting object within the orbit at the current time, or some other time of interest. That information is supplied via the *true anomaly* at a given epoch time, which is the actual angular position of the object along the orbit relative to perigee. So this geometric model for describing an orbit as a whole requires exactly six parameters and a time to define an object's orbit, just as the previously described "vectorial' perspective (where we looked at what the object was doing while in its orbit) required exactly six parameters (position and velocity vectors) and a time. The *epoch time* is usually the time associated with the start time of an ephemeris file or the beginning of a new orbit, for example as a result of an orbit modification maneuver.

However, the true anomaly is difficult to calculate (it is the root of a transcendental equation), so in practice a more straightforwardly determined parameter, the *mean anomaly*, is provided instead. The mean anomaly is also an angular displacement from perigee, but is determined simply by dividing the time since perigee passage by the orbital period, then multiplying by 2π radians:

$$M = \frac{2\pi}{T}(t - T_P) \tag{2.3}$$

where

M = mean anomaly at time t (radians)
T = orbital period
T_P = time of perigee passage
(t, T, and T_P are all measured in the same time units)

In other words, it is the fraction of the period of the orbit covered since perigee passage. The mean anomaly does not have the true anomaly's concrete geometric meaning (except for circular orbits, where they have the same value), but the two can be related via the following two equations:

$$M = E - e\sin(E) \qquad \text{“Kepler's Equation”} \tag{2.4}$$

$$\tan(\nu/2) = ((1+e)/(1-e))^{1/2}\tan(E/2) \qquad \text{“Gauss's Equation”} \tag{2.5}$$

where

ν = true anomaly
M = mean anomaly
E = eccentric anomaly
e = eccentricity

Since this book is supposed to be the Kinder, Gentler ACS, we'll avoid getting caught up too much in these different anomaly flavors; it is sufficient to consider E an intermediate variable to compute ν from M. (It would have been nice if E could be expressed as an analytic function of M, as opposed to the other way around, but we'll take what we can get.) Let's concentrate now on the mean anomaly at epoch and see under what conditions it becomes undefined, as we did with right ascension of the ascending node and argument of perigee. First, if the orbit is circular, perigee is undefined and so is mean anomaly, but again we have a situation where the orbit may be considered a limiting case for a family of orbits defined by the sum of two angles, in this case it is the sum of the argument of perigee (ω) and the mean anomaly (M) that is well-behaved (for a non-equatorial orbit). This bit of hocus-pocus works because the angle from the ascending node (a well-defined point for non-equatorial orbits) and the spacecraft location (always a well-defined point) is equal to the sum of ω and M. Similarly, if the inclination is zero and the orbit is circular, then the mean anomaly, argument of perigee, and the right ascension of the ascending node will all individually be ill-defined, but the quantity representing their geometric sum (i.e., the angle from the the GCI x-axis to the spacecraft) will be constant for a family of Kepler parameter sets all representing the same orbit.

Although this focus (no pun intended) on singularities seems somewhat esoteric, utilization of these relationships between orbital parameters for "degenerate" orbits can be of great value when transforming an orbital (position, velocity) pair into equivalent Keplerian elements. For example, these singularity relationships were utilized repeatedly when developing the ground system algorithms for computing the uplink parameters required to refresh HST's onboard ephemeris models.

2.5 Keplerian Element Types

Putting all six of these parameters (called Keplerian elements) and an epoch time together, we have all the information required to define the current position of an object in a perfect elliptical orbit and, in the absence of orbit perturbations, have all the information required to determine the orbital position of the object at any other time as well, future or past. It also can be shown (but not here) that specifying the six components of two object position vectors on the ellipse at two different times (as long as the vectors are not parallel) provides the same information as the other two approaches (i.e., Keplerian elements and (position, velocity) vector pair, both with associated times) although, if the angle between the two position vectors is near 0 or 180°, the results of the computation of new position vectors may be numerically unstable. These relationships were also utilized when developing the ground system ephemeris uplink parameter generation algorithms for HST. So we have a similar situation with orbit formulations as we had with the attitude formulations discussed in Chap. 1. There are lots of ways to describe an orbit, each having its own advantages and disadvantages.

The Keplerian elements provide for orbits the sort of visualization advantages that the Euler angles provided for attitude. You can easily grasp the size and shape of the orbit, how the orbital plane is oriented, where perigee is located, and where the spacecraft is located on the ellipse at the epoch time. Also, in the absence of orbit perturbations, only the mean anomaly changes with time. On the downside, we have seen that three out of six Keplerian elements become ill-defined for special geometries, circular and equatorial orbits that tend to be very "popular" for Earth orbiting science missions. And although the (position, velocity) vector formulation is pretty user-hostile from a visualization standpoint, it is ideal for numerical calculations of "real" orbits in the presence of perturbative effects, especially when an ephemeris file with many vector pairs is provided, enabling interpolation between points. So in practice, you pick your formulation for its convenience relative to the problem you're trying to solve or the function you're trying to perform.

There are even two types of Keplerian elements, each having their own uses. Implicitly, the type we've been talking about are Keplerian *osculating elements*, so called because they will generate an ideal orbit that will "kiss" the physical orbit (when orbit perturbation effects are included) at a single point, the point corresponding to the epoch time. However, in the presence of perturbations, the ideal osculating orbit and real physical orbit will diverge over time. A more accurate fit to the physical orbit can be achieved through the use of *Brouwer mean elements*. Brouwer mean elements effectively provide averaged elements over several orbits, thereby averaging out the influence of periodic perturbative effects rather than (as would the osculating elements) incorporating the value of the perturbation value at the epoch time as if it was constant in time. So osculating elements may propagate and amplify the effects of periodic perturbations with time, while mean elements will bound their effects. But without inclusion of explicit time dependence (i.e., adding derivative terms to the basic six-term Keplerian set), neither approach will deal effectively with secular (i.e., ramping) perturbations over long propagation periods. Probably by now the term "perturbative effects" is sounding like some sort of physics *deus ex machina* invoked to intimidate an audience from asking orbit questions. ("Quaternion" is another good term for that purpose when presenting attitude material to a review board.) Just to put that suspicion to rest, we'll spend the rest of this chapter discussing the physical phenomena (and manmade machinery) that perturb orbits and the orbital geometries particularly susceptible to those influences.

2.6 Orbit Perturbations - Oblate Earth

Perturbative effects on Earth orbits are caused by forces other than those associated with a "point Earth" (i.e., spherically symmetric Earth) attracting the object. An analytical approach in which you treat all other influences as perturbations will only be successful numerically if any additional forces are much smaller in magnitude than that generated by the point Earth, but fortunately (for overworked flight dynamics analysts) that is usually the case for spacecraft orbiting the Earth. For the missions we've worked with here at GSFC so far, there are five forces that exert particularly significant influence on spacecraft orbits, namely non-point Earth gravitational forces, non-two-body gravitational forces, aerodynamic drag, solar radiation pressure, and thrusters.

Back in elementary school you probably learned that the Earth is shaped more like a pear than an apple (disregarding flat bottoms and pointy tops). This is still true today, and this aspect of the Earth's structure, referred to as Earth oblateness, is the key enabler of two major mission classes. If you model the Earth as an oblate spheroid (a good approximation to the slightly pear-shaped distortion mentioned by your teacher), great circles along a meridian on the Earth's surface are replaced by ellipses whose semi-major axes are equal to the Earth's equatorial radius (about 6,378 km) while their semi-minor axes are equal to the Earth's polar radius (about 6,357 km), yielding an ellipse eccentricity of about 0.08. Small circles at constant latitudes will remain small circles, but their radii will be slightly modified (from the spherical Earth model) because of the equatorial bulge relative to the poles.

To get an idea how oblateness influences a spacecraft's orbit, let's start with the spherical Earth. Assume the radius of the sphere is the "real" Earth's polar radius. In the absence of any perturbing influences, the orbit sketched out previously in Fig. 2.2 will be constant in time. Next, replace the spherically symmetric Earth model with one in which some of the mass is moved into an equatorial bulge extending about 20 km out from the immediate vicinity of the equator. When the spacecraft is at its orbital nodes, the force arising from the mass in the bulge exerts an in-plane force on the spacecraft in the direction of the bulge's (and Earth's) center. Visualizing the spacecraft's orbit with orbit normals sticking out of it as if it were a rotating top (where the spacecraft as it orbits is sitting on the external edge of the top), the effect of the bulge at the nodes is just a force applied to the center-of-mass of the top. By contrast, at the points 90° from the nodes, the bulge mass exerts an out-of-plane force that is trying to tip the top down (i.e., reduce the orbit inclination). As we'll see in Chap. 3, that's the same thing as saying the bulge is exerting a torque on the top, whose effect will cause the top's spin axis (i.e., the orbit normal) to precess westwards (when inclination is less than 90°) about the Earth's spin axis. The direction of rotation is determined by the cross product of the spacecraft position vector with the direction of the force. Expressed in terms of Keplerian elements, that's equivalent to saying that the equatorial bulge will cause the orbit's right ascension of the ascending node to rotate.

The rate of the nodal rotation depends mostly on the altitude and inclination of the spacecraft's orbit. As you get farther and farther away from the Earth, the Earth acts more and more like a single point of gravitational attraction. Mathematically, you can show (but we won't for fear of losing what little readership we have left) that the nodal precession rate varies inversely with the 3.5th power of the semi-major axis, so the rate drops off very quickly with altitude. The effect of inclination is also easy to see. If the spacecraft is in a perfect polar orbit (90° inclination), the orbital plane cuts symmetrically through the equatorial plane and therefore the torque generated by the bulge must go to zero (to first order). As you decrease inclination, the influence of the bulge steadily increases, although you reach a singularity at inclination zero since the right ascension of the ascending node becomes undefined.

To get a more quantitative feel for the size of the nodal rotation induced by the bulge, let's look at a couple of standard orbits. For a 500-km altitude near-circular orbit with an inclination of 28.5° (the inclination you get for "free" by launching from Kennedy Space Center (KSC), whose latitude is 28.5°), the nodal rotation will be about −6.7 deg/day, i.e., a westward nodal rotation. By contrast, if you're in an orbit with inclination 97.4° and altitude 500 km, the nodal rotation rate (eastwards) becomes 1 rotation/year (about 1 deg/day), which keeps the orbit synchronized with the annual movement of the Sun, i.e., Sun synchronous.[3] We'll talk more about these very special, and useful, orbits in the last chapter. Also, keep in mind that this is the simple picture you get from just looking at the first order perturbations, also called the J_2 term in reference to the coefficient of the second term (after the point mass term) in a spherical harmonic expansion of the gravitational potential (that mathematical derivation we mentioned but refused to present earlier). When the inclination is near 90°, the effects of Earth oblateness on the orbit typically become smaller than that produced by less symmetric components of the Earth's mass distribution, and the real evolution of the orbit becomes correspondingly more complicated.[4]

But wait, there's more (don't worry; we're not selling Ginsu knives). In addition to causing the line of nodes to rotate, the Earth's oblateness will also cause the line of apsides (the line joining apogee and perigee) to rotate. Again, imagine we've generated a static orbit as a result of the gravitational force exerted by a spherically symmetric Earth, which we can shrink to a single point at the focus of the spacecraft's elliptical orbit. Again, pull some of that mass into a ring modeling the equatorial bulge. For validity of the approximation, assume the bulge is even smaller than the Earth's 20-km one. Let's also (referring to Fig. 2.2) visualize that the spacecraft orbit is lying in the equatorial plane (i.e., has inclination zero) and that the spacecraft's orbit is highly elliptical. At perigee (in particular), the spacecraft will feel a stronger force than that resulting from the entire bulge mass being concentrated at the Earth's center. The effect is dominated by the portion of the bulge closest to the spacecraft.[5] Since the extra tug at perigee is perpendicular to the velocity vector, it won't add or detract from the spacecraft's kinetic energy, so the height at apogee will stay constant, unlike the case where we can raise (or lower) apogee by adding (or removing) kinetic energy at perigee by thrusting parallel (or anti-parallel) to the direction of the velocity vector. Instead, the extra gravitational pull at perigee will cause a tighter deflection

---

[3] Wertz, Spacecraft Attitude Determination and Control, pp. 68–69.

[4] In more formal mathematical terms, the secular J_2 term becomes smaller than the periodic terms associated with the zonal, sectorial, and tesseral coefficients that model the more complex (latitude, longitude) dependent mass distributions, as we'll discuss in the section on the Earth's geopotential in Chap. 6.

[5] Using a simple model in which the bulge is represented by a thin, massive wire girdling the equator, one can show that in the limit in which the spacecraft almost grazes the wire at perigee, the gravitational pull from a small fraction of the wire closest to the spacecraft will be inversely proportional to the distance between the wire and the spacecraft (becoming infinite if the separation is zero) and will completely dominate the pull from the rest of the wire.

of the spacecraft's path than would have been the case for an orbit about a point-mass Earth. This will cause the new apogee point to be shifted in a positive orbit sense relative to the previous apogee without changing the distance between apogee and the Earth's center. As a shift in apogee will occur on each orbit, what we end up with is (to first order) a constant rotation of apogee (and its associated perigee) in the direction of the spacecraft's orbital motion.

By contrast, for a polar orbit with perigee initially located at a pole, the spacecraft at perigee (in particular) will feel a little bit weaker force from the bulge than that resulting from all the bulge mass being concentrated at the Earth's center. Note that the Earth's center is closer to the pole than is any point on the equator, so moving mass from the center to an equatorial ring reduces the pull felt at the pole. Once again, the diminished tug at perigee is perpendicular to the velocity vector, so no change in the spacecraft's kinetic energy results. But the reduced gravitational attraction at perigee will cause the spacecraft's path not to be deflected as tightly as would be true if the spacecraft had been orbiting a point-mass Earth, so the new apogee point will be shifted in a negative orbit sense relative to the previous apogee without changing the distance between apogee and the Earth's center. So for polar orbits we get a rotation of perigee in the direction opposite of the spacecraft's orbital motion.

Logically, if perigee rotations at zero inclination are positive and perigee rotations at polar inclination are negative, there should be some "magic" inclination angle between 0 and 90° at which the perigee rotation is nulled, and in fact there is. At an inclination of 63.435° (ignoring terms of higher order than the secular J_2 term), the perigee rotation rate goes to zero, yielding a "frozen" orbit. This type of orbit has very useful communications applications for countries located in high latitudes, like Russia, as we'll see in the last chapter if you survive the intervening 9 chapters of material. As an example of the size of the effect, for a 500-km near-circular orbit with inclination 28.5°, the perigee rotation rate is 10.9° per day in a positive direction.[6]

Nearby the Earth, oblateness effects are the dominant influence on spacecraft orbits. However, as you get further from the Earth, perturbative effects arising from the Sun's and Moon's gravity become more important. As a rule of thumb, if the spacecraft altitude is below 700 km, lunar and solar gravitational effects can be ignored, while at 8000 km, lunar perturbations are comparable in importance to Earth oblateness.[7] For a great many missions, for example those orbiting Lagrange points, two-body approximations break down completely and the spacecraft's orbital behavior can only be defined by analyzing the combined influences of two celestial bodies. Continuing the running gag for this chapter, a discussion of Lagrange points and their associated missions also will be provided in Chap. 12.

[6] J. Wertz, Spacecraft Attitude Determination and Control, p. 69.

[7] Ibid, p. 63.

2.7 Orbit Perturbations - Aerodynamic Drag

Returning closer to home, aerodynamic drag dominates at altitudes below 100 km, and is an important influence on spacecraft orbits up to about 1,000 km.[8] Of the various perturbations discussed here, drag may be the most complicated, as there are so many factors entering into the picture. In particular, the drag force is proportional to atmospheric density (which is affected by altitude, latitude, constituent composition and chemistry, time of day, season of year, point in solar cycle, butterfly migration patterns, etc.), the square of the spacecraft velocity relative to the atmosphere, and the spacecraft cross-sectional area normal to the wind. Like friction, drag is a retarding force, so it acts to oppose the spacecraft direction of motion. Over the long term, drag eventually causes the orbit to decay, leading to a decrease in the semimajor axis. So unlike Earth oblateness, drag actually removes energy from the orbit. Drag also tends to make the orbit more circular for the following reason. Since drag has its greatest effect at the point in the orbit where the atmosphere is densest (drag decreases exponentially with altitude), the effects of drag will be highest at perigee. Therefore at perigee, more energy will be removed than at apogee. But removing energy at perigee (or apogee) means the swingby at apogee (or perigee) will not reach as high an altitude as on the previous orbit. So the perigee and apogee altitudes will be lower on each succeeding orbit, but the amount of lowering will be greater at apogee than at perigee. This also is due to the inverse distance behavior of the gravitational potential energy, which will cause the same amount of energy reduction to produce a larger altitude decrease at apogee than at perigee. So until the two altitudes roughly become identical, the effect of drag will be to make the orbit steadily more circular, i.e., it will steadily lower eccentricity. Note that these changes from energy loss are only in-plane. Drag has no effect on inclination or right ascension of ascending node, nor for that matter on argument of perigee (at least until it circularizes the orbit) .

Before leaving our discussion of atmospheric drag, it's worthwhile to spend a little time talking about how you model the key input to drag computations, atmospheric density. Of course, the higher you go, the thinner the air, following an exponential drop-off. The atmospheric density will also be influenced by temperature (via the Ideal Gas Law), which decreases as we move away from the equator. But the dominant factor determining the molecule's energy content is solar heating, the effects from which vary over several time scales. The highest frequency influence is the daily variation caused by ultraviolet radiation heating the atmosphere by conduction, causing atmospheric density to increase to its maximum level by 2–3 h after local noon. There also is a 27-day solar activity period, the yearly period (i.e., seasonal patterns), and the 11-year solar cycle, which reached a maximum in 2001. Lumped in with these fairly predictable events, we also have short-term fluctuations in atmospheric density arising from the solar wind and completely unpredictable, violent perturbations from solar flares and coronal mass ejections. At GSFC, two models in particular have found popular use, especially as input to GSFC's orbit

[8] Ibid, p. 63.

determination program, the Goddard Trajectory Determination System (GTDS).[9] The simpler of the two is the Harris-Priester model,[10] which accounts for diurnal (a fancy word for *daily*) influences. The Jacchia-Roberts model[11] includes seasonal variations and constituent composition (at least at the level of molecular weight behavior as a function of altitude) as well as atmospheric density changes caused by solar flares and geomagnetic activity.

2.8 Orbit Perturbations - Solar Radiation Pressure

Once you get up high enough that the atmosphere has virtually disappeared and the Earth looks round (i.e., away from the neighborhood that Low Earth Orbit (LEO) spacecraft call home), solar radiation pressure takes over as the major environmental orbital perturbation. Solar radiation pressure is much easier to model, depending largely on solar luminosity, spacecraft reflectivity, distance from the spacecraft to the Sun, and spacecraft cross-sectional area. The cross-sectional area itself is a function of attitude, which for celestial-pointing missions will change whenever the science target changes. So the actual solar radiation pressure orbit perturbations are closely tied to the science observing schedule. In practice, as when atmospheric drag is modeled, GTDS is not set up to deal with this dynamic attitude behavior, so an average cross-sectional area approximately valid for the time duration of interest is what's input. In any event, solar radiation pressure is proportional to the spacecraft cross-sectional area normal to the sunline (as opposed to the cross-sectional area normal to the spacecraft velocity vector as in the case of atmospheric drag). It also is proportional to the momentum flux from the Sun. The flux, in turn, is proportional to the solar luminosity and is inversely proportional to the square of the distance from the Sun. To see that, no pun intended, just imagine light emitted by the Sun evenly distributed over a sphere of radius equal to the separation distance. The constant of proportionality for the pressure calculation is a function of the spacecraft reflectivity. A reflectivity of 2 is perfectly reflective, while a value of 1 denotes perfectly absorbent, and a value of 0 means perfectly transparent. Most spacecraft are pretty shiny, so aluminum's reflectivity of 1.95 is characteristic of most spacecraft.

---

[9] A. Long, J. Cappellari, C. Velez, and A. Fuchs. Goddard Trajectory Determination System (GTDS) Mathematical Theory, Revision 1. Technical report, NASA/GSFC Flight Dynamics Division Code 550, 1989.

[10] I. Harris and W. Priester, Time-dependent Structure of the Upper Atmosphere, Journal of Atmospheric Science, 19, pp. 286–301, 1962.

[11] C. Roberts, An Analytic Model for Upper Atmosphere Densities Based Upon Jacchia's 1970 Models, Celestial Mechanics, Volume 4, Issue 3, pp. 368–377, Dec 1971.

2.9 Orbit Perturbations - Orbit Maneuvers with Thrusters

The last perturbation we'll talk about, thrust, is completely artificial and, unless something goes badly wrong, planned, predictable, and easily modeled. Depending on the size of the thrusters mounted on the spacecraft, thrust can be the dominant influence on the spacecraft orbit. Thrusters often come in two functional flavors, *orbit thrusters* and *attitude thrusters*. Orbit thrusters, of course, are designed to generate major orbit changes and trims. Although attitude thrusters, as their name implies, are designed to enable attitude control and angular momentum management, misalignments in the thrusters can produce minor, undesirable, but somewhat predictable orbit perturbations. We'll have a lot more to say about thruster hardware and electronics in Chap. 5, but for now we'll just mention briefly some strategies for their use.

Recall in our discussion of drag, frictional removal of energy at perigee will lower apogee while frictional removal of energy at apogee will lower perigee. This same physical effect can be achieved artificially by firing thrusters so as to decelerate the spacecraft at those points (actually on an arc centered on those points). By contrast, firing orbit thrusters so as to accelerate the spacecraft (i.e., artificially adding kinetic energy) at perigee will raise apogee while doing the same at apogee will raise perigee. These kinds of orbit maneuvers are called in-plane maneuvers, and tend to cost much less fuel than out-of-plane maneuvers (nodal rotation or inclination change), which change both the magnitude and direction of the orbital angular momentum vector (in-plane maneuvers can only change the magnitude). If you want to produce pure orbit inclination changes (say, lowering the "automatic" 28.5° you get with a KSC launch to an equatorial orbit), the most cost-effective place to perform your "burns" is at the orbital nodes. Note that because of fuel budget limitations you often accomplish the desired orbital modification simply by initiating a slow, steady drift in the orbital elements in the appropriate direction. At a later time, possibly months later, an inverse orbit maneuver can be performed to cancel the drift, leaving the spacecraft in the desired orbit.

http://www.springer.com/978-1-4471-7324-3

ACS Without an Attitude
Hallock, H.L.; Welter, G.; Simpson, D.G.; Rouff, C.
2017, XVII, 279 p. 54 illus., Hardcover
ISBN: 978-1-4471-7324-3