Cambridge University Press
978-1-107-03584-3 - GPS, GLONASS, Galileo, and BeiDou for Mobile Devices
Ivan G. Petrovski
Excerpt
More information

Part I

GNSS: orbits, signals, and methods

1 GNSS ground and space segments

Global Navigation Satellite Systems (GNSS) at the time of writing comprise four systems, two of which are fully operational and two of which are on their way (see Table 1.1). A brief history of GNSS is given in Chapter 10 along with a timeline of application development and prospects for this development, especially concerning mobile applications.

Each GNSS comprises a constellation of satellites, called a space segment, and a ground segment (Figure 1.1). The main idea behind GNSS is to measure distances between a satellite and a user located on the surface of the Earth or in a lower atmosphere. Satellite coordinates can be calculated at any moment of time. The information that allows the calculation of satellite position is uploaded to and then broadcast from satellites to the user. The ground segment is responsible for determining satellite orbits, which it then uploads to the satellites, and also for defining the coordinate frame and time frame in which satellite and user positions are estimated.

Having received the information on satellite orbits and measured distances to the satellites, a user can calculate receiver position as an intersection of four spheres in a four-dimensional space-time continuum. If the receiver clock is perfectly synchronized with the satellite time frame, only three satellites would be required to determine receiver position in three-dimensional space (Figure 1.2).

This chapter describes the GNSS space segment and provides the information required to understand the process of calculating satellite coordinates. Algorithms used to calculate satellite positions using various orbit parameters for all existing GNSS are given. The first operational GNSS were the American Global Positioning System (GPS) and the Russian GLObal NAvigation Satellites System (GLONASS), or GLObal'naia NAvigacionnaia Sputnikovaia Systema in Russian. Next came those GNSS that are not yet operational: the European system, Galileo, named after Galileo Galilei (1564–1642), the Italian astronomer and philosopher; and the most recent, the Chinese BeiDou, named for the Chinese pronunciation of the Big Dipper constellation.

1.1 Ground segment and coordinate reference frames

The ground segment comprises tracking stations and facilities, which provide coordinate and clock reference frames, calculate satellite coordinates, and upload this information to satellites.

Table 1.1. GNSS as of 2013

	System			
	GPS	GLONASS	Galileo	BeiDou
Country	USA	Russia	EU	China
Signal from space availability (operational satellites)	1973	1982	2011	2010
Start of operation (positioning service worldwide)	1993	1995	2019 (expected)	2020 (expected)

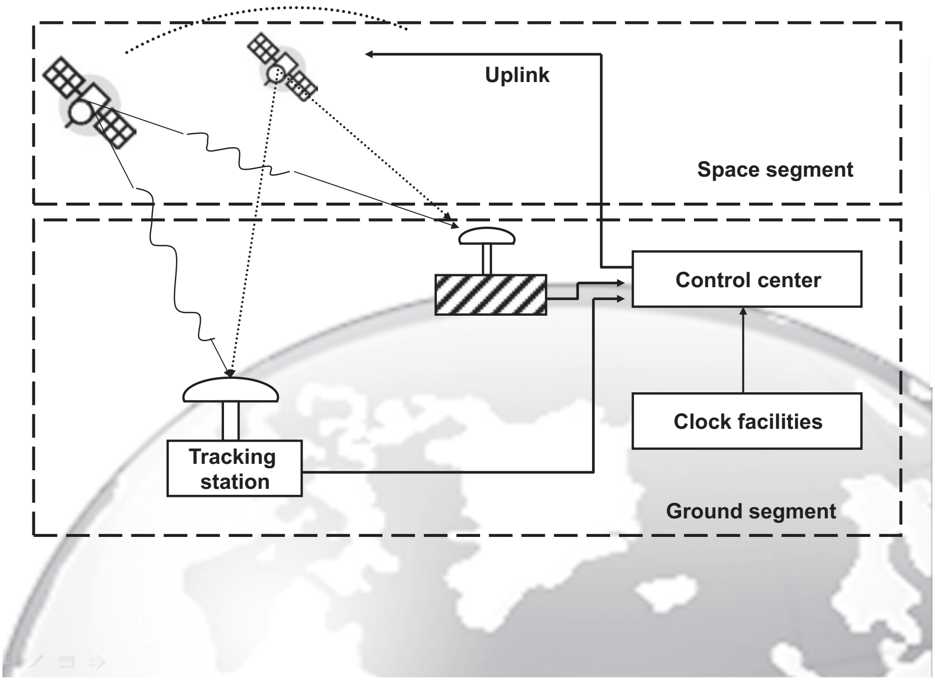

Figure 1.1 GNSS space and ground segments.

The functions of the ground segment can be summarized as follows.

(1) To establish a coordinate frame, in which satellite and user positions are calculated.
(2) To establish a time scale, which, along with the coordinate frame, defines the complete four-dimensional space-time continuum, in which satellite and user positions are calculated.
(3) To collect satellite measurements.
(4) To calculate satellite orbits.
(5) To monitor the satellite signal.
(6) To upload information on satellite orbits and health to the satellites.

A satellite's orbit is defined in a coordinate system, which combines three-dimensional coordinate frame and a time frame. First, we define two essential coordinate frames: Earth-centered inertial (ECI) and Earth-centered, Earth-fixed (ECEF) (see Figure 1.3).

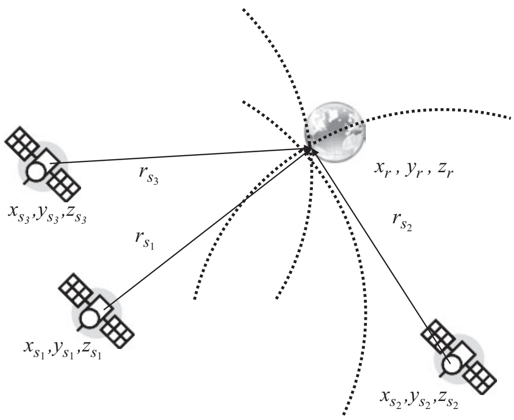

Figure 1.2 Positioning with satellites.

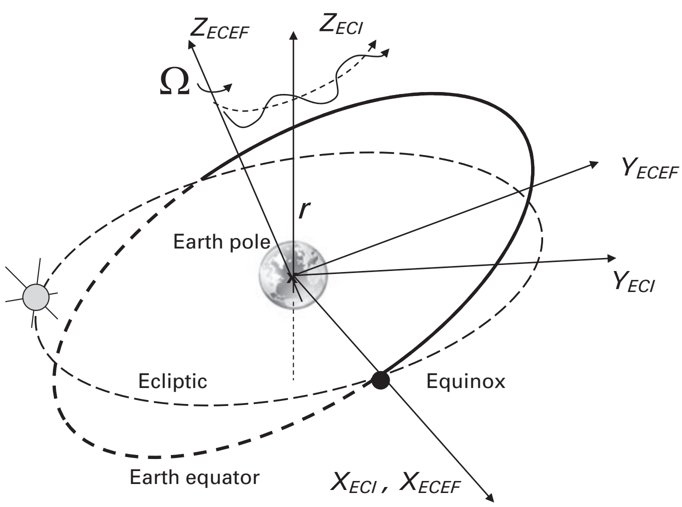

Figure 1.3 ECI and ECEF coordinate frames.

The ECI coordinate frame is a Cartesian coordinate frame with an origin placed at the Earth's center of mass and axes fixed relative to distant stars. The *z*-axis coincides with Earth's spin axis and the *x*-axis is defined by the direction from the Earth to the Sun on the first day of spring, when the Sun crosses the Earth's equatorial plane.

This point of intersection between the Sun's trajectory (ecliptic) and the Earth's equatorial plane is called the vernal equinox, or sometimes the First Point of Aries, as it was named thousands of years ago when the vernal equinox was in the zodiacal constellation of the Ram (Aries). The Aries zodiacal symbol ♈ is still used to mark the vernal equinox. The vernal equinox, defined as the intersection line of the equatorial and the ecliptic planes, is not fixed in space due to precession and nutation. This vernal equinox precession has a period of 26 000 years, which results in an ECI precession rate of 0.014° per year. This drift makes it necessary to reference coordinate frames to a certain date. The vernal equinox coordinates are then adjusted to the epoch of interest through precession and nutation transformations, which are given as sequences of rotations.

In order to make this precise transformation between the ECI and the ECEF, one can apply Earth orientation parameters (EOP), which are freely available from the International GNSS Service (IGS). A navigation message for an L2C GPS signal also contains EOP. The coordinate transformation of a satellite position from ECEF to ECI is described by a series of rotations as follows:

$$\overline{X}_{ECEF} = [R_{PM}][R_{ER}][R_N][R_P]\overline{X}_{ECI}, \tag{1.1}$$

where $[R_{PM}]$, $[R_{ER}]$, $[R_N]$, $[R_P]$ are the rotation matrices for polar motion, Earth rotation, nutation, and precession, respectively. For most mobile applications, all these movements, with the exception of the Earth daily rotation, can be safely neglected. However, when we consider tasks that require higher accuracy, including orbit determination by a ground network, it becomes necessary to account for all movements, including polar motion, nutation, and precession.

The ECEF frame is a Cartesian coordinate frame with its origin placed at the Earth's center of mass. The z-axis coincides with the Earth's spin axis, and the x-axis goes through the Greenwich meridian. The ECEF frame rotates with the Earth.

The satellite orbit parameters are defined in a tracking network, which is effectively an ECEF frame. However, as described in Section 1.2, the mathematical presentation of an orbit in the inertial frame is much simpler (see Figure 1.4). However, in order to be useful for a user located on the Earth's surface, satellite coordinates in the inertial frame must be transformed back to the ECEF frame.

A satellite navigational message embedded in the transmitted signal provides a user with orbital parameters. These parameters allow the user to calculate the satellite's position in the ECEF frame. Keplerian parameters given in a navigation message for almanac and ephemeris are not strictly speaking Keplerian parameters for the satellite orbit. These parameters are defined in the ECEF frame by control segments by fitting these parameters to a set of measurements from a control segment reference station network. The orbital parameters are calculated not in the inertial frame, which is fixed relative to stars, but in a special "non-rotating" with the Earth ECEF frame. This frame is different from the ECI frame due to small rotations caused by polar motion, nutation, and precession.

As a result of using this modified inertial frame, we can transfer satellite coordinates to the ECEF frame simply by multiplying them by a matrix that describes only the Earth's rotation. This provides us with satellite coordinates accurate enough for mobile applications.

For mobile applications, the relative motion of ECEF in relation to ECI can be confined to a rotation around the z-axis with the Earth's angular velocity, i.e.

$$\overline{X}_{ECEF} = \begin{bmatrix} \cos(\omega_E \cdot t) & \sin(\omega_E \cdot t) & 0 \\ -\sin(\omega_E \cdot t) & \cos(\omega_E \cdot t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \overline{X}_{ECI}, \tag{1.2}$$

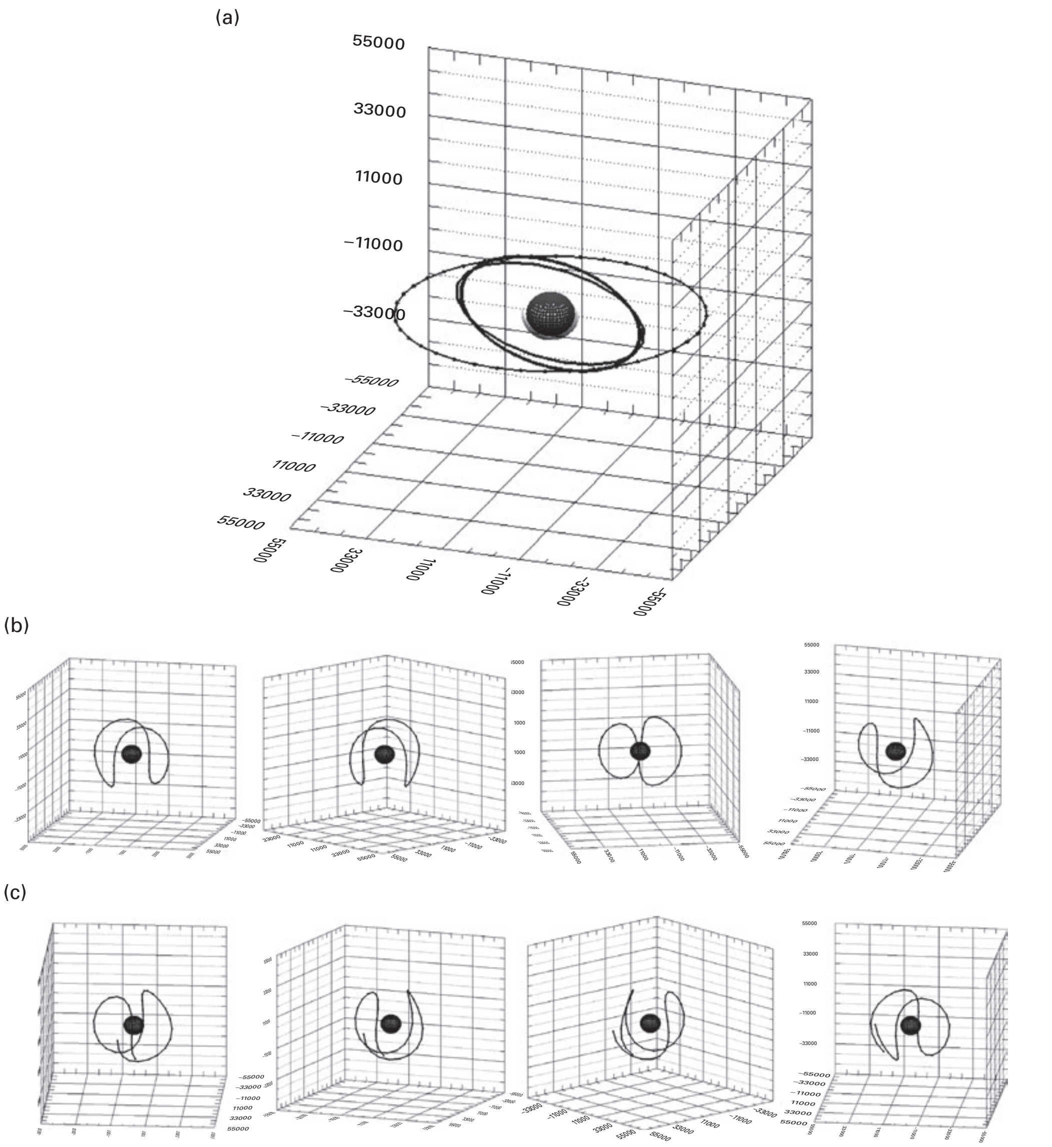

Figure 1.4 (a) GPS, GLONASS, and GEO orbits in ECI; (b) GPS orbit in ECEF; (c) GLONASS orbit in ECEF.

where ω_E is the angular velocity of the Earth's rotation.

In matrix form,

$$\overline{X}_{ECEF} = [R_{\omega_E}] \times \overline{X}_{ECI}, \tag{1.3}$$

where the rotation matrix $[R_{\omega_E}]$ satisfies the following conditions:

$$[R_{\omega_E}(\omega_E)]^{-1} = [R_{\omega_E}\omega_E]^T = [R_{\omega_E}(-\omega_E)]. \tag{1.4}$$

Correspondingly,

$$\overline{X}_{ECI} = \begin{bmatrix} \cos(\omega_E \cdot t) & -\sin(\omega_E \cdot t) & 0 \\ \sin(\omega_E \cdot t) & \cos(\omega_E \cdot t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \overline{X}_{ECEF}. \tag{1.5}$$

In order to specify an inertial frame, we need to specify a reference epoch for both the equator and the equinox. The J2000.0 reference system is one of the most common, used, for example, by the Center for Orbit Determination in Europe (CODE) Analysis Center. More than one realization of the ECI frame exists, for example the International Celestial Reference Frame (ICRF), which is determined using a catalog of extragalactic stars based on Very Long Baseline Interference (VLBI) observations. The International Terrestrial Reference Frame (ITRF) is an ECEF frame, which corresponds to the ICRF. Although there are various realizations of the ITRF, we are particularly interested in the IGS realization, which is based on GNSS observations.

The IGS ITRF consists of coordinates and velocities of a set of globally distributed tracking stations, which are in fact GNSS receivers, for specific epochs. One can say that the ITRF frame is fixed to a network of IGS tracking stations on the Earth's surface.

In a similar way, unless ITRF is used, a ground segment defines a particular realization of the ECEF frame for its system. The problem is that the ground segment should rely on its own tracking stations in order to ensure integrity of the system. Therefore, the ground segment tracking stations ultimately define the underlying coordinate frame for each system, because a control segment reference network, which defines these coordinates, is, in a manner of speaking, the ECEF frame itself.

GPS uses the WGS-84 coordinate frame. The main difference between the WGS-84 and the ITRF is that the WGS-84 may only be realized by users with a resolution of about one meter in geocentric position (because of the quality of the broadcast orbits and satellite clocks). The ITRF may be employed with centimeter accuracy if IGS orbits and ITRF coordinates of the IGS sites are included in the processing. The two systems are therefore consistent at about the one-meter level.

Another very important issue concerning the ground segment is that the accuracy of orbit determination depends on the network distribution. This is in effect similar to the GNSS dilution of precision (DOP) factor [1], which is considered in detail in Chapter 3. Figure 1.5 demonstrates this effect. The achievable accuracy for a regional network is much less than for a global one. The effect can be mathematically described

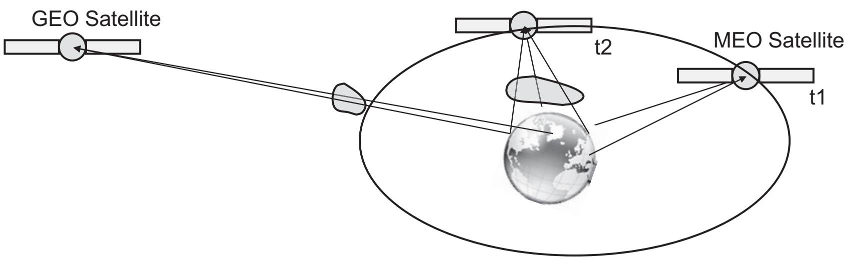

Figure 1.5 Ground segment geometry.

by reversed DOP, which is considered in detail in Chapter 3. At two moments in time, t_1 and t_2, the GPS satellite position can be measured with a much wider angle and by different stations, whereas geostationary satellites can see only the same stations because its position in relation to the Earth is fixed and the angle between the lines of sight to these stations is much narrower. In Chapter 3 we consider the satellite to station DOP, whereas here the effect is described by the station to satellite DOP. The satellite orbit stability and ground network geometry define not only the accuracy of ephemeris determination, but also how often satellite maneuvers are required. For example, for GPS maneuvers are required once a year, whereas geostationary satellites may require them once a week.

The DOP factor, however, is not the only effect that comes from the station network distribution. A globally distributed network allows for continuous satellite tracking over the whole orbit, whereas a regional network can track only part of the orbit. GNSS orbits operate at a distance of about 20 000 km, whereas the Earth radius is about 6000 km. Network density is a factor which affects accuracy significantly. To compensate for the regional character of the network, one can add additional constraints, which may include time synchronization [2] or inter-satellite measurements. Such constraints significantly improve accuracy and the time required for data collection.

The ground networks therefore define slightly different frames, unless they are using ITRF and correspondingly the same satellite tracking stations. Despite this, for mobile applications the difference is negligible. Even for much more demanding geodetic applications, various GNSS coordinate frames can be easily unified. The time frame, however, may differ significantly for various GNSS. We will look at these time scales in Section 1.2.

The GLONASS ground network has a regional character in comparison to GPS, which affects its achievable accuracy. Figure 1.6 shows a schematic distribution of GPS and GLONASS ground segment networks. In 2013, Russia had 19 ground network tracking stations, and in addition there are correction and monitoring stations. Three of these stations are located in Antarctica; others are in Russia, Belarus, and Ukraine. The Russian Space Agency has estimated that at least 40 tracking stations would be required to increase GLONASS's accuracy tenfold. There are negotiations between Russia, the USA, China, and some other countries to extend the monitoring network. However, not all of these stations can be considered as part of the control segment because of the integrity issue.

The first correction and monitoring station outside of the former Soviet Union territory was established in February 2013 in Brazil. The purpose of this station is to provide differential corrections for GLONASS and signal monitoring [3]. This station should significantly improve GLONASS's accuracy, especially in those parts of the world that previously had no tracking stations.

The IGS ground network has a global distribution and very high density. It provides the highest available accuracy of orbit estimation. However, it cannot guarantee real-time integrity. In post-processing mode, if invalid measurements occur or the quality of the measurements is not sufficient, for example due to problems with antennas, the environment, or interference, then such measurements can be caught and removed.

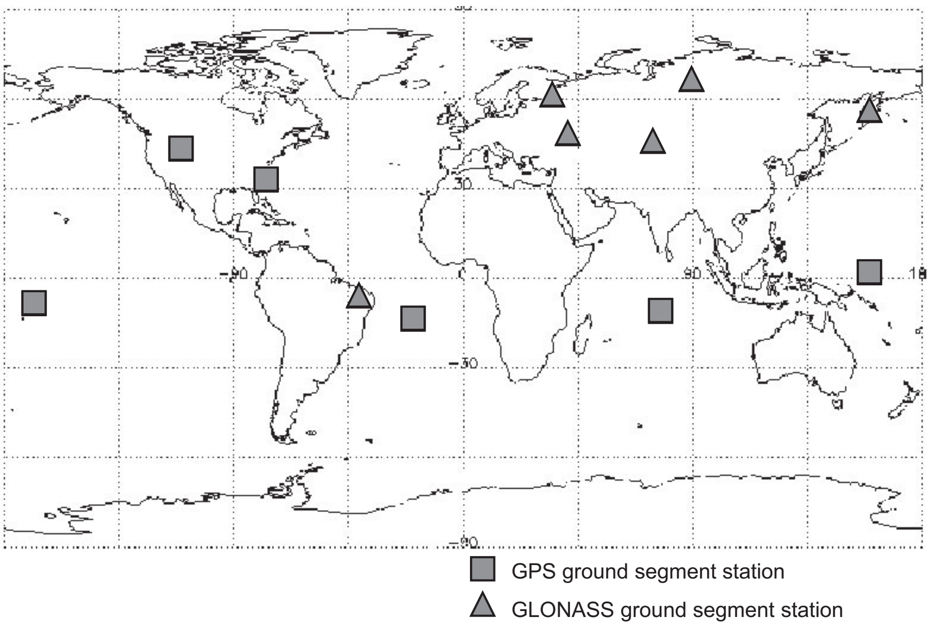

Figure 1.6 Ground segment networks for GPS (after [1]) and GLONASS (after [3]).

1.2 Space segment and time references

1.2.1 GPS time and calendar time

Satellite system time frames are connected to specific world or country time frames. In relation to time scales, we should mention Coordinated Universal Time (UTC). UTC is the main internationally accepted time standard, which has specific realizations for each GNSS within the country that owns that specific GNSS. For example, GPS time is referenced to UTC as it is maintained by the U.S. Naval Observatory (UTC (USNO)). Galileo is referenced to UTC as it is provided by the Galileo Time Service Provider based on information from European UTC laboratories.

All GNSS time scales differ from their corresponding UTC. For example, GPS time is continuous, whereas UTC is periodically adjusted to compensate for leap seconds. Also, the GNSS time scale may drift in relation to UTC. Control segments maintain their scales so the difference to the respective UTC realization is kept within certain limits. For GPS it is within one microsecond.

In terms of mobile applications, we can consider Greenwich Mean Time (GMT) instead of UTC. For practical use we express GMT via the Gregorian calendar. For mobile applications, we may for now reduce all considerations regarding various time frames to the following: connecting all GNSS time frames to GPS time and then converting GPS time to calendar time. It is important to remember that we also need to include leap second correction in the conversion of GPS time to calendar time.