

Chapter 2

Geodetic Data Collection Techniques

In order to accomplish the disciplinary tasks of geodesy, various geodetic data need to be collected extensively. This chapter briefly introduces the methods and principles of the data collection techniques that are commonly used in geodetic survey, such as terrestrial triangulation, height measurement, space geodetic surveying, physical geodetic surveying, and so on.

2.1 Terrestrial Triangulation

2.1.1 Angle Measurement

In establishing national geodetic control networks, it is often necessary to carry out horizontal and vertical angle measurements. The theodolite is an instrument for measuring angles with specific observational methods.

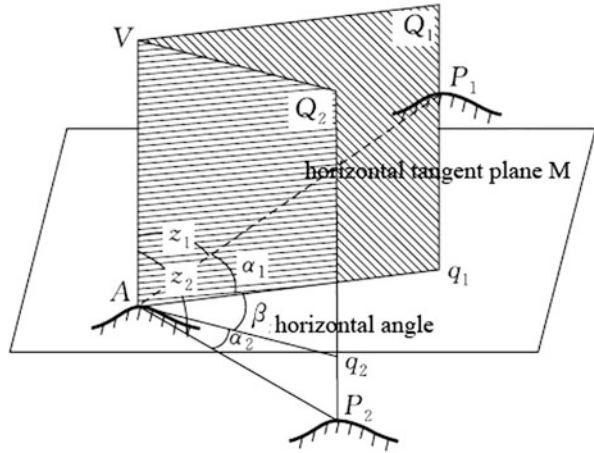
Horizontal and Vertical Angles

Horizontal Angle

In Fig. 2.1, A , P_1 , and P_2 are three geodetic control points on the Earth's surface. Let A be the point of observation and P_1 and P_2 be the target points. Through point A , draw a plumb line AV (the direction of gravity) and a plane M , which is perpendicular to AV . The plane M is called the horizontal plane through point A .

The line of intersection Aq_1 between the horizontal plane M and the vertical plane containing the line of sight (line of collimation) formed by the plumb line AV and the line of sight AP_1 is called the projection of AP_1 on the horizontal plane, which is usually called the horizontal line (horizontal) of AP_1 . Similarly, Aq_2 is the horizontal of AP_2 . The angle β between Aq_1 and Aq_2 is known as the horizontal

Fig. 2.1 Horizontal and vertical angles



angle to P_1 and P_2 from A . The horizontal angle is measured clockwise in the horizontal plane from 0° to 360° .

Vertical Angle

The angle between the line of sight AP_1 and its horizontal Aq_1 is called the vertical angle to the sighted point P_1 from A , denoted by α_1 . Likewise, the angle between the line of sight AP_2 and its horizontal line Aq_2 is referred to as the vertical angle to the sighted point P_2 from A . Therefore, a vertical angle is the angle between the line of sight, which is the collimation axis of the telescope, and its corresponding horizontal, which includes the angle of elevation and angle of depression.

A vertical angle is measured in the vertical plane from 0° to $\pm 90^\circ$, positive above the horizontal (see Fig. 2.1, α_1) and negative below (see Fig. 2.1, α_2).

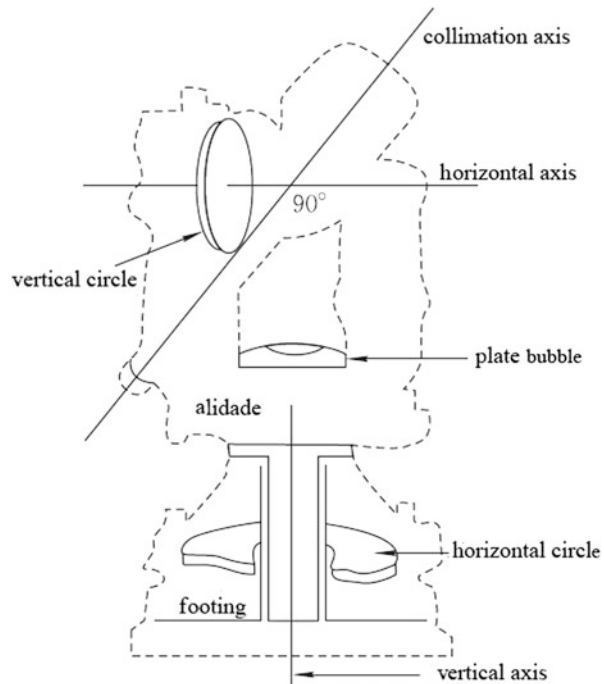
The angles Z_1 and Z_2 between the plumb line AV and the lines of sight AP_1 and AP_2 are called the zenith distances from point A to the sighted points P_1 and P_2 .

As illustrated in Fig. 2.1, the sum of the vertical angle and the zenith distance of a target is 90° , namely

$$\alpha + Z = 90^\circ. \quad (2.1)$$

According to this relation, the vertical angle and the zenith distance can easily be converted one to the other.

Fig. 2.2 Basic structure of a theodolite



The Theodolite

There are many types of theodolite, from the classical optical theodolite to the modern electronic theodolite. Theodolites are classified into J_{07} , J_1 , J_2 , and J_6 in China according to their precision. “J” is the initial letter of Chinese Pinyin (Jing Wei Yi) for theodolites and the subscript indicates their accuracy for angle measurement (mean square error). J_{07} and J_1 are of high accuracy and are used for the first- and second-order national control surveys, while J_2 is of medium-level accuracy for third- and fourth-order surveys (see, e.g., Xu et al. 1991).

A theodolite consists of the principal components shown in Fig. 2.2. These components are related as follows:

1. The vertical axis must be perpendicular to the axis of the plate bubble. When the plate bubble is centered, the vertical axis coincides with the plumb line.
2. The vertical axis must be at right angles to the horizontal circle and pass through its center. When the vertical axis is perpendicular, the horizontal circle is parallel to the horizontal plane through the point of observation and the angle measured in such cases will be the true horizontal angle.
3. The horizontal and vertical axes must be perpendicular and the collimation axis must be at right angles to the horizontal axis. So, when the vertical axis is perpendicular and the telescope is elevated or depressed, the plane formed by the collimation axis will be the vertical plane of sighting.

4. The horizontal axis must be perpendicular to the vertical circle and pass through its center. When the vertical axis is perpendicular and the horizontal axis is horizontal, the vertical circle is parallel to the vertical plane of sighting passing through the point of observation and the angle measured in such cases will be the true vertical angle.
5. Once the vertical circle index bubble is centered, the index for reading the vertical circle must be horizontal or vertical. Thus, the angle between the index of the reading scale and the collimation axis of the telescope will be the vertical angle.

The major parts of a theodolite are structured according to the relationships mentioned above. It is generally required that the relationship between three axes (the vertical and horizontal axes and the collimation axis) and between two circles (horizontal and vertical circles) be free of errors, which is crucially important.

Unlike the optical theodolite, the electronic theodolite provides a visual digital display of the circle readings instead of having to view through a reading eyepiece. It is therefore also called an electronic digital theodolite. The electronic theodolite is composed of optical mechanical devices, electronic sensor, microprocessor, etc. The configuration of its shafting, telescope, and clamp (tangent) screw are identical to that of the optical theodolite. The difference is that the electronic sensor is used to replace the index of the reading scale in the ordinary theodolite. Following the method of analog-to-digital conversion, it first receives electrical signals from the circle and then converts these electrical signals into angles and displays them on the monitor. Generally, there are kinematic and static angle measurements according to the rotation of the circle, and there are circular encoders and circular raster scales for angle measurements according to the different ways of circle graduating. Detailed principles are not discussed further here.

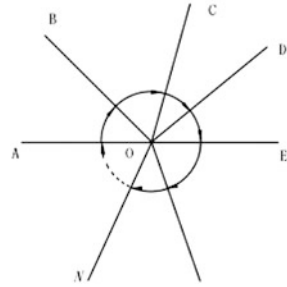
Methods for Observing Horizontal Angles

The Direction Method and the Closing the Horizon Method

For each set of observations, all directions at the station should be observed to get the angles. As shown in Fig. 2.3, the directions that need to be observed at station O are OA , OB , \dots and ON . Let OA be the starting direction (also referred to as zero direction). First, point A is sighted facing left and the reading is recorded. Then the alidade is rotated clockwise and the points observed in order from A to N and the readings noted, as one half of the full set of observations. Then the telescope is turned to face right, the alidade rotated anticlockwise, and the observations repeated facing right. The horizontal angles are then recorded again in reverse order from N to A as the other half of the full set. At this stage, one set of angles has been completed. Such a method is known as the direction method.

While using the direction method, before completing the direct-mode or reverse-mode readings, zero direction A will be measured again (called back-to-zero or being zeroed). Given the fact that each half of the whole set will combine and close up to the starting direction as a full circle, it is referred to as closing the horizon

Fig. 2.3 The direction method



method. The purpose is for an additional check to see if any variations in the tribach of the instrument have appeared during the direct or reverse mode of observations. The direction method and closing the horizon method are basically alike, and can be collectively called the direction method. When the number of the observed directions is equal to or less than three, the time consumed by one set of angle observations is quite short, and the direction method can be adopted (non-back-to-zero or not being zeroed); when the number of directions is greater than three (multiple angles), the closing the horizon method should be employed.

As to closing the horizon, whether the selection of zero direction is appropriate or not will exert some influence on both the accuracy and time of observation at the station. Therefore, the zero direction should be the one with appropriate length, good intervisibility, and clear target image.

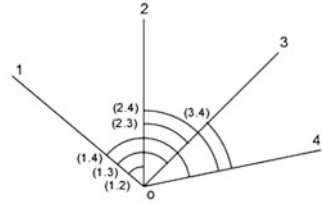
Currently, the direction method is primarily used in angle measurement where lower precision is acceptable. Using the direction method, one obtains the values of all directions at a station. The value of the chosen zero direction is equal to zero. The angle between directions is the difference between the two directions.

Method of Angle Measurement in All Combinations (Schreiber's Method of Observation)

The direction method is simple and requires less effort in observation. However, the sides of the national higher-level control networks are of greater length, and the different target images cannot all maintain good quality at the same time. Moreover, the time consumed in one set of observations is rather long. It is therefore hard to achieve results with significantly greater accuracy. To overcome these deficiencies, the method of angle measurement in all combinations can be used. The major characteristic of this method is that it only measures the angle between two directions each time. In so doing, it is possible to overcome the difficulty in maintaining the clarity and stability of various target images simultaneously. Meanwhile, it also helps shorten the time used in one set of observations and makes it possible to achieve awesome results with higher accuracy, making it the preferred method for accurately measuring horizontal angles.

Each time, two directions are selected out of all the directions to be observed at the station and these are combined to form single angles; this is called the angle in

Fig. 2.4 Method of angle measurement in all combinations



all combinations, for example, if four directions need to be observed at the station, six single angles can be formed: (1.2), (1.3), (1.4), (2.3), (2.4) and (3.4) (cf. Fig. 2.4). If the number of directions at the station is n , then the number of angles in all combinations is given by:

$$K = \frac{1}{2}n(n - 1). \quad (2.2)$$

For each set of observations, only a single angle is observed and the observation set for each combination angle is the same. The characteristic for such observation is that the alidade is rotated in the same direction for both direct-mode and reverse-mode reading. This is intended to better eliminate the errors due to backlash as the alidade rotates. However, for the entire observation and each set of observations for every single angle, the alidade should be rotated clockwise in a half set and counterclockwise in the other half so as to reduce other errors better. During each time period of observation this is achieved by either changing the rotational direction of the alidade when half of the set is completed or changing the rotational direction of the alidade in between the sets of observation alternately.

The above covers the direction method and angle measurement in all combinations. The key advantage of the direction method lies in its simple observation procedures and in the fact that less effort is required for operation. However, when there are multiple angles at the station, it will be hard to obtain clear target images in all directions. On top of that, the time consumed in a full set of observations is longer, which may make it more likely to be affected by external conditions, so difficulties may arise in achieving results with high accuracy. These are the disadvantages of the direction method. The method of angle measurement in all combinations has certain advantages, i.e., each single angle can be observed with a flexibly selected clear target for each set of observations, the time for observation is quite short, and the impact of external conditions on the observations is relatively small. However, the procedures for observation are relatively complicated; the combinations of single angles increase rapidly as the directions at the station increase and thus measurements require much more effort to perform. These are the drawbacks of such a method. Generally, the direction method is more applicable to angle observation with lower accuracy, whereas the method of angle measurement in all combinations is quite suitable for angle observation with higher accuracy.

The measurement of vertical angles requires little for observation and accuracy. What needs to be done is to observe the target using the crosshairs of the telescope during the direct-and-reverse modes of readings, which constitutes a complete set of observations. The vertical angle can be computed accordingly.

2.1.2 Distance Measurement

For hundreds of years, graduated tapes (measuring ropes, tape measures, and steel tapes) have been used to measure distance by means of direct comparison. The major flaw of such a method, however, is that it is easily subjected to the influence of topographic conditions along the survey lines. To obtain distance measurement with higher accuracy, one has to invest large amounts of human and material resources to choose and arrange the survey routes, which can be complicated and costly. Moreover, such distance measurement cannot be carried out if there are rivers, lakes, hillocks, or ravines along the survey lines.

With the progress of science and technology, in the 1940s a new type of distance measuring instrument came into being—the optical-electro distance measuring instrument, which was also the earliest type of electromagnetic distance measuring (EDM) device. Later, microwave, laser, and infrared EDM instruments emerged one after another. Even today, the total station electronic tacheometer that integrates angle and distance measurements is still available. It has created a new era of using EDM to replace the direct comparison method using graduated tapes and the indirect method using the optical tacheometer.

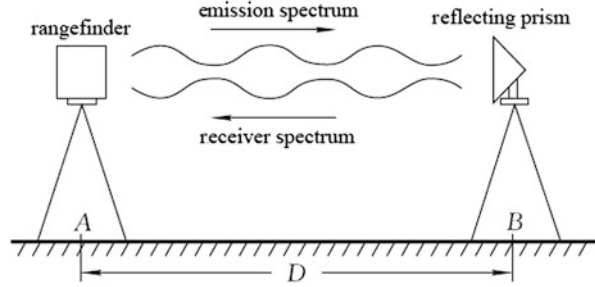
Principles of Electromagnetic Distance Measuring

In Fig. 2.5 an electromagnetic wave transmitted from a rangefinder placed at point *A* travels to the reflector at point *B* and back to point *A*, received by the rangefinder. Thus, the round-trip travel time t_{2D} of the electromagnetic wave between points *A* and *B* can be measured by the rangefinder. The distance (*D*) can be calculated according to the following formula:

$$D = \frac{1}{2} V t_{2D}, \quad (2.3)$$

where *V* is the velocity of propagation of the electromagnetic wave in the atmosphere, *c* is the velocity of the electromagnetic wave in vacuum, and *n* is the index of refraction (refractive index) for the electromagnetic wave:

$$V = \frac{c}{n}. \quad (2.4)$$

Fig. 2.5 Principles of EDM

The value of the index of refraction n is dependent on the wavelength λ of the electromagnetic wave and the meteorological elements of the atmosphere. The relationship between n and the air temperature t , the barometric pressure p , and the humidity e is expressed as (Duan 1996):

$$n = f(\lambda, t, p, e). \quad (2.5)$$

Knowing the wavelength of the electromagnetic wave and the temperature, pressure, and humidity of the atmosphere, the value of the refractive index n can be computed according to (2.5).

To sum up, the principle of EDM is to use instruments to measure directly or indirectly the round-trip travel time t_{2D} of the electromagnetic wave along the distance D and to measure the temperature t , the pressure p , and the humidity e at the same time to compute the distance according to the above formula.

It is clear that we can directly measure the distance between two points at the endpoint using the EDM method. Distances under any topographic conditions can be measured provided that the measurement range is reached and no obstacles interrupt the line of sight between the two points. Distances between mountains, rivers, and even planets for instance, using satellite laser rangefinder can also be measured directly, which can greatly reduce observation time.

Basic Methods of Electromagnetic Distance Measurement

There are three basic methods of EDM:

1. Method of Distance Measurement by Pulse (Pulse Method)

The distance D to the target is attainable if we directly measure the travel time t between the transmitted pulse (dominant wave) and the reflected pulse (echo) from the target. With this, the distance can be obtained with a measurement performed only once, the measurement range varying from several kilometers to hundreds of thousands of kilometers. The precision generally reaches centimeter levels. Such a pulse method is chiefly used for measurements of low accuracy or

long distances, such as the front line tactical reconnaissance and distance measurements from Earth to the Moon and from Earth's surface to artificial satellites.

2. Method of Distance Measurement by Phase (Phase Difference Method)

We can directly measure the phase difference between the transmitted signal and the echo to get the travel time. The measuring accuracy of such a method is better than the millimeter level and its measurement range is within dozens of kilometers. It is commonly used in precision distance measurement on the ground.

3. Method of Distance Measurement by Interference (Interferometric Method)

This method adopts the physical principle of optical interference for precise distance measurement with higher accuracy than that of distance measurement by phase. Its precision generally reaches micrometer levels. It is normally used for calibrating distance measuring instruments and for precise short-distance measurement.

Classification of Electromagnetic Distance Measuring Instruments

EDM instruments can be classified into the following three categories according to the band of their carrier waves:

1. Microwave EDM Instrument

The carrier wavelength ranges from 8 mm to 10 cm in the microwave band.

2. Laser EDM Instrument

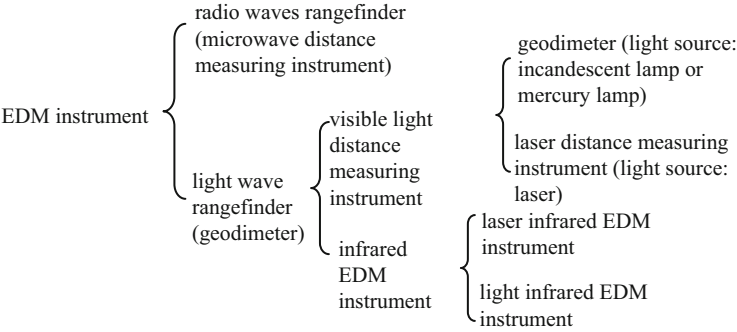
The carrier is usually red visible light of $0.6328\ \mu\text{m}$ wavelength and it is stimulated emission of radiation (i.e., laser emission).

3. Infrared EDM Instrument

The carrier wavelength usually ranges from 0.75 to $0.95\ \mu\text{m}$ and it is stimulated emission of radiation (i.e., laser emission) or spontaneous radiation (fluorescent light).

Generally, electromagnetic distance measurements involve distance measurements with radio waves and with light waves. Measuring distance using radio waves refers to microwave distance measurement. Light wave distance measurement includes two categories, one being visible light distance measurement and the other infrared distance measurement.

Visible light wave distance measuring devices can be categorized into two types. At the early stage of development, the instrument was based on spontaneous emission with an incandescent lamp or mercury lamp as its light source. In later models, a red laser ($\lambda = 0.6328\ \mu\text{m}$) is stimulated by emission of radiation and the light source is generally a He-Ne-gas laser, as summarized below.



2.1.3 Astronomical Measurement

Astronomical observation is a technique utilized to determine the position of a point on the Earth’s surface (astronomical longitude and astronomical latitude) and the astronomical azimuth by observing celestial bodies (especially stars). Astronomical observation is an ancient technology dating back to the era when human culture first took shape. To figure out direction and determine time and season, the sundial and gnomon were invented successively and Polaris was used for determining North.

Definition of Astronomical Coordinate System

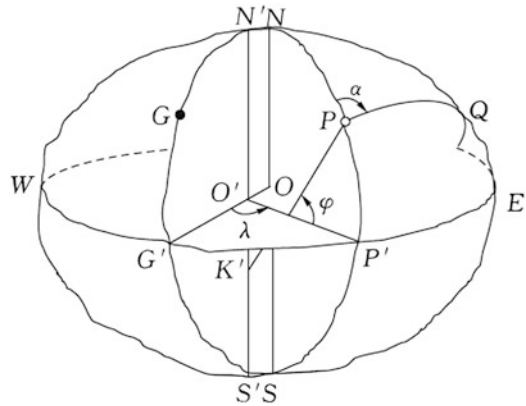
In astronomical observation, the astronomical longitude, latitude, and azimuth of a surface point obtained with reference to its plumb line and the geoid are the representations of the surface point in the astronomical coordinate system.

As shown in Fig. 2.6, NS is the Earth’s axis; the two points N and S , where the Earth’s axis intersects the Earth, are the North Pole and South Pole, respectively; O is the geocenter; the plane $OWG'P'E$ through the geocenter perpendicular to the spin axis is the Earth’s equatorial plane; P is a point on the Earth’s surface; PK' is the plumb line direction of point P ; the plane that contains the plumb line of point P is called the vertical plane of point P , in which the vertical plane $N'PP'S'K'$ parallel to the Earth’s axis is called the astronomical meridian plane; $NGG'S$ is the initial astronomical meridian plane, also referred to as the astronomical meridian plane (see e.g., Xia and Huang 1995).

Astronomical Longitude

The astronomical longitude of a point P on the Earth’s surface is the angle between the initial astronomical meridian plane and the local astronomical meridian plane of

Fig. 2.6 Geoid and astronomical coordinate system



this point, denoted by λ . It is measured eastward or westward from the initial meridian and ranges from 0° to 180° ; proceeding eastward is referred to as east longitude and westward is west longitude. East longitude is positive and west longitude is negative.

Astronomical Latitude

The astronomical latitude of a surface point P is the angle between the Earth's equatorial plane and the plumb line of this point, denoted by φ . It is measured northward or southward from the equator to the poles and ranges from 0° to 90° ; proceeding northward is referred to as north latitude and southward is south latitude. North latitude is positive and south latitude is negative.

Astronomical Azimuth

Let P be the point of observation and Q the target point, then the vertical plane that contains point Q and the plumb line of point P is the plane of sighting of direction PQ . Therefore, the astronomical azimuth of direction PQ is the angle between the astronomical meridian plane of point P and the vertical plane of point Q , denoted by α . Its value is in the horizontal plane of the point of observation, measured from due north of the meridian clockwise from 0° to 360° .

Initial meridian is the intersection of the plane containing the meridian of 0° longitude with the Earth's surface. It is internationally agreed. In 1884, the International Meridian Conference decided to adopt the meridian passing through the Observatory at Greenwich, England (Airy Transit Circle) as the initial meridian, known also as the prime meridian or zero meridian.

Methods for Astronomical Observation

Traditional Methods for Astronomical Observation

Currently, astronomical observation usually adopts the traditional methods, mainly to determine time by receiving time signals emitted from the observatory and to record time by a chronometer. The instruments used in observations are primarily a Wild T4 theodolite and a 60° astrolabe, and the methods extensively applied are as follows:

1. The Wild T4 theodolite allows determination of the first-order astronomical latitude by applying Talcott's method.
2. The Wild T4 theodolite allows determination of the errors of a timepiece (time corrections) by the method of equal altitudes of two stars, one east and the other west of the meridian (Zinger's method), in order to determine the first-order astronomical longitude.
3. The hour angle of Polaris is applied for determining the astronomical azimuth.
4. A 60° astrolabe (composed of T3 and 60° prisms) allows simultaneous determination of the astronomical longitude and latitude of the second, third, and higher orders using the method of equal altitudes of multiple stars.

New Methods for Astronomical Measurement

The new methods primarily employ the GPS OEM (Original Equipment Manufacturer) board with time transfer service to receive satellite signals. Electronic theodolites are adopted for observation instead of optical theodolites; portable computers with advanced programming are used to replace the chronometer and timepiece for time comparison and timekeeping; and the autonomous recording and calculation of the observational data are also enabled. The currently adopted methods are as follows:

1. Use the method of observing multiple stars at approximately equal altitudes to determine the first- and second-order astronomical longitude and latitude simultaneously.
2. Carry out repeated observation using the method of hour angle of Polaris to determine the first- and second-order astronomical azimuth.

2.2 Height Measurement

2.2.1 Leveling

Principle of Leveling

Leveling is a method used for accurate determination of height difference between two points. The basic principle of leveling is that a precisely graduated staff is held vertically over the two points whose height difference is to be determined and then the scale readings are made with the horizontal line of sight. The difference between the two readings will be the height difference between the two points. As shown in Fig. 2.7, A and B are two surface points with unknown height difference. Leveling rods (leveling staffs) R_1 and R_2 are held vertically on each point while a level is placed at point S_1 in between these two points. From the horizontal line of sight, a reading of the rod R_1 is made as “ a ” (known as the backsight reading) and that of the rod R_2 is “ b ” (known as the foresight reading). Then the height difference h_{AB} between A and B is:

$$h_{AB} = a - b. \quad (2.6)$$

The height difference is positive when $a > b$ and negative when $a < b$.

Knowing the height H_A of point A , we can obtain the height of point B following $H_B = H_A + h_{AB}$. To determine the height H_P of an arbitrary point P , one needs to move the level to S_2 and the rod R_1 to point C after the height difference between A and B is determined. Then, the height difference h_{BC} between points B and C can be obtained. Likewise, the difference in height between A and P is $h_{AP} = h_{AB} + h_{BC} + \dots$.

The height of point P is:

$$H_P = H_A + h_{AP}. \quad (2.7)$$

Such a method of transferring heights is referred to as geometric leveling.

Level and Leveling Rod

It can be seen from the principle of leveling that the leveling instrument should be developed to set up a horizontal line of sight. Therefore, the level should have a telescope capable of creating a line of sight (collimation axis) and a component that can direct the line of sight to the horizontal direction (a bubble is one of the simplest kinds). To make the line of sight horizontal and rotate horizontally, foot screws and a vertical axis are also necessary. Integrating these components as shown in Fig. 2.8 will constitute the simplest level. These principal components should satisfy the following conditions:

Fig. 2.7 Principle of leveling

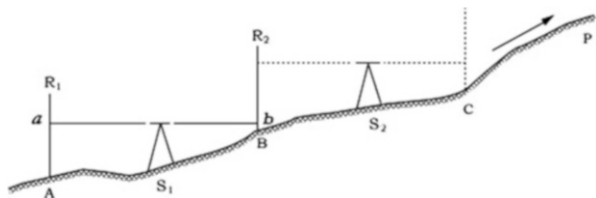
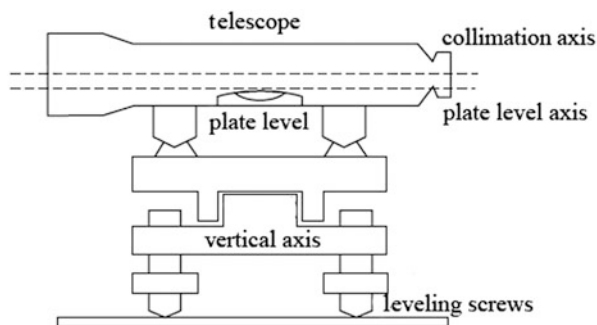


Fig. 2.8 Structure of the level



1. The line of sight is parallel to the axis of the bubble (level) tube.
2. The axis of the level tube is perpendicular to the vertical axis.

In this case, when the instrument is leveled with the bubble tube, the line of sight will be horizontal in all directions.

Precision dictates whether the leveling instrument can be classified as a precise leveling instrument or an ordinary leveling instrument. A precise leveling instrument is chiefly used in high-precision height measurement as in national first- and second-order leveling and precise engineering surveying. An ordinary leveling instrument is used in general engineering construction and topographic surveys. The major difference between these two types is that the precise leveling instrument has a built-in optical micrometer for accurate readings.

The leveling rod is an important leveling instrument that can be used to determine the difference in height between points.

A precise leveling rod has a 26 mm-wide and 1 mm-thick Invar strip placed in the grooves of the wooden part of the rod, with one end fixed to the base plate and the other to the metal frame at the top of the rod by a spring. Graduations of the rod are painted to fill the grooves cut in a scribed rule and the graduation lines are painted on the wooden part of the rod. The rod is approximately 3.1 m in total length.

The rod can be graduated at intervals of 10 mm or 5 mm, according to the measurement range of the level micrometer. Graduations are painted in two columns on the left and right sides of the rod.

The rod holder loop is also configured to both the back sides of the leveling rod to help hold it. The rod stand or stake is installed to keep the rod steady and upright.

Electronic Levels

The first electronic level was invented in March 1990 as a brand-new leveling instrument that integrated electronic technology, encoding, image processing, and computer technologies, and marked the direction for development of leveling instruments. Today, several corporations throughout the world are manufacturing electronic leveling instruments such as the DNA03 and DNA10 of Leica Microsystems, DiNi10 and DiNi20 of Carl Zeiss, and DL-101 and DL-102 of Topcon.

Different to the optical level, the rod face of the electronic level is graduated with a bar code and there also is an inbuilt digital image recognition and processing system. Using digital image processing technology, the image of the bar code can be processed and compared through a telescope, which enables the naked eye of the observer to be replaced by the array detectors (sensors). Observations (including clamping and reading) can therefore be completed automatically. In surveying operations, the leveling instrument needs to be only roughly leveled, the line of sight is automatically made horizontal by the compensator, the leveling rod is sighted, and the focus is adjusted. In such a case, by pressing the “measure” button, the reading of the rod and the distance between the rod and leveling instrument will be displayed on the monitor.

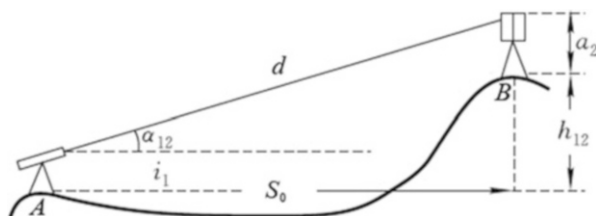
2.2.2 Trigonometric Leveling

Trigonometric leveling is a method for determining the difference in height between two ground control points by using the distance and vertical angle observed between the two points and then transferring the heights of ground control points. Compared with geometric leveling, trigonometric leveling is much simpler and more flexible. It is independent of terrain conditions and enables faster transfer of heights. The flaw in trigonometric leveling, however, is its slightly lower accuracy in determining heights. If controlled by leveling with sufficient density, trigonometric leveling can therefore not only ensure the accuracy of ground control point measurement but also overcome terrain constraints and improve work efficiency (see, e.g., Kong and Mei 2002).

Basic Principle of Trigonometric Leveling

As shown in Fig. 2.9, A and B are two points on the Earth’s surface and their heights are H_1 and H_2 , respectively. The vertical angle from A to B is α_{12} , S_0 is the horizontal distance between the two points, i_1 is the height of the instrument (HI), and a_2 is the height of the target (HT) of point B . The difference in height between A and B will be:

Fig. 2.9 Principle of trigonometric leveling



$$h_{12} = H_2 - H_1 = S_0 \tan \alpha_{12} + i_1 - a_2. \quad (2.8)$$

If the measured slope distance is d , the height difference is:

$$h_{12} = d \sin \alpha_{12} + i_1 - a_2, \quad (2.9)$$

which is the basic formula for computing the height difference using trigonometric leveling. Given the height of point A, that of point B becomes:

$$H_2 = H_1 + h_{12}. \quad (2.10)$$

EDM Height Traversing

Electromagnetic distance measurement (EDM) height traversing is also called precise trigonometric leveling. With the development of the electronic tachymeter, accuracies of angle and distance measurements have been greatly improved. The accuracy of distance measurement reaches over 1/100,000 and that of angle measurement can amount to 0.5", which provides favorable conditions for precise trigonometric leveling. At present, third- and fourth-order leveling can be completely replaced by EDM height traversing and, accordingly, in China specifications have been made by the departments concerned. Replacing leveling with EDM height traversing has proved to be notably economical in mountainous and hilly regions.

The methods of height traversing include reciprocal, leap-frog, and unidirectional. For the reciprocal method, the instrument is set up at each station to conduct reciprocal trigonometric leveling. The leap-frog method involves setting up an instrument midway between two targets. The targets remain at a particular change point. Observations are carried out in a pointwise manner. The targets should be set alternately and an even number of setups is used. This method is similar to leveling, except for using an oblique instead of a horizontal line of sight. The unidirectional method is based on the first and second methods, which is to observe twice with different heights of instrument at one station or to observe twice the two targets at each station. Tailor-made sighting vanes are used as the targets for all three methods described above.

2.3 Space Geodetic Surveying

2.3.1 GPS Surveying

Overview of GPS

Authorized to start in November 1973 by the US Department of Defense, GPS is a second-generation American satellite-based navigation system. It is accessible by the armed forces and cost 1,000 million dollars. It became fully operational in 1994 as the third greatest project after the Apollo lunar spacecraft and space shuttle. In surveying, navigation, guidance, precision positioning, dynamic surveying, time transferring, velocity measurement, and so on, GPS is convenient to use, easily operational in observation, precise in positioning, and beneficial economically. It has displayed its powerful functions and unparalleled superiority (Xu 2001).

The entire GPS consists of the space segment, the ground control segment, and the user segment.

Structure of GPS

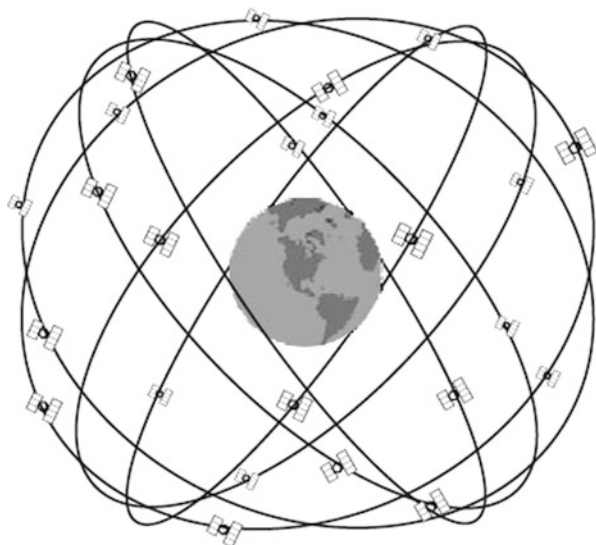
The Space Segment

As shown in Fig. 2.10, the space segment is composed of 24 GPS operational satellites which form the GPS satellite constellation. Of these, 21 are navigation and positioning satellites and 3 are spares. These 24 satellites orbit around the Earth in six orbits at an inclination angle of 55° . Except at the two poles, theoretically more than four satellites can always be in view anywhere on the Earth's surface at any time. Orbiting at an altitude of about 20,000 km, each satellite makes one complete orbit every 12 sidereal hours. Every operational GPS satellite can transmit signals for navigation and positioning, which are then utilized by GPS users.

The Control Segment

The control segment is a monitoring system composed of several tracking stations around the globe. These tracking stations are categorized into master control stations, monitor stations, and up-link stations based on their functions. There is one master control station, which is located in Falcon Air Force Base in Colorado, America. Based on the observational data of GPS from every monitor station, it calculates the correction parameters of ephemeris and clocks of the satellites and uploads these data to the satellites through up-link stations. Meanwhile, it takes control of and gives instructions to the satellites. When one operational satellite goes wrong, it will dispatch a spare to replace the invalidated one. Additionally, the master control station also possesses the functions of monitor stations. There are five monitor stations. Apart from the master control station, the other four are

Fig. 2.10 GPS satellite constellation



located in Hawaii, Ascension Island, Diego Garcia, and Kwajalein. They are designed to receive signals from the satellites and monitor the satellite working status. The three up-link stations are located at Ascension Island, Diego Garcia, and Kwajalein. These stations upload to the satellites the correction parameters of ephemeris and clocks of the satellites computed by the master control station.

The User Segment

The user segment consists of the GPS receivers, the data processing software, and corresponding auxiliary equipment for users, etc. It is intended to receive signals sent by GPS satellites and to use these signals for navigation, positioning, and so on.

Signals From GPS Satellites

GPS satellites transmit carrier signals for civilian use at three frequencies: 1,575.42 MHz (L_1 carrier wave), 1,227.60 MHz (L_2 carrier wave), and 1,176.45 MHz (L_5 carrier wave). Their wavelengths are 19.03, 24.42, and 25.48 cm, respectively. Many signals, chiefly the C/A, P, and D codes, are modulated on carrier waves L_1 , L_2 , and L_5 .

The C/A code, also known as coarse acquisition ranging code, is a pseudo-random noise code (PRN code) with a frequency of 1.023 MHz. The total code period contains 1,023 chips and lasts 1 ms. Different satellites can be distinguished by their PRN names because each satellite differs in its C/A code.

The P code, known as precision ranging code, is a PRN code at a frequency of 10.23 MHz.

The D code, known as the navigation message, at 50 bits per second, carries the position of satellites, status information, etc.

GPS Positioning Services

GPS offers two positioning services: the Precise Positioning Service (PPS) and the Standard Positioning Service (SPS).

PPS. Authorized users of the PPS, including the US military, certain US government agencies, and civil users specifically approved by the US government, need cryptographic equipment and special receivers. The positioning accuracy of PPS is several meters and the time accuracy reaches 40 ns.

SPS. For common civilian users, the US Government provides the SPS to take control of positioning accuracy. Users worldwide use the SPS without charge or restrictions. In the initial stages of GPS system implementation, SPS accuracy was intentionally degraded by the US Department of Defense by the use of so-called selective availability (SA). Under the effect of SA, the positioning accuracy of SPS is degraded to approximately 100 m and the time service accuracy is about 340 ns. In May 2000, the USA announced it was discontinuing the use of SA. At present, SPS provides a positioning accuracy of approximately 10 m and a time service accuracy of about 100 ns.

GPS Coordinate System and Time System

World Geodetic System 1984 (WGS84)

For a worldwide unified geodetic coordinate system, the US Defense Mapping Agency (DMA) has provided WGS60 since the 1960s and later developed the improved WGS66 and WGS72. WGS84, currently used by GPS, is a more accurate global geodetic coordinate system.

The coordinate origin of WGS84 is at the Earth's center of mass. Its Z-axis is the direction of the Conventional Terrestrial Pole (CTP), as defined by BIH1984.0. The X-axis points to the intersection of the zero meridian plane defined by the BIH1984.0 and the plane of the CTP's equator. The Y-axis constitutes a right-handed coordinate system.

The coordinates of GPS single point positioning and the baseline vector in relative positioning solution belong to the WGS84 geodetic coordinate system on which the GPS satellite ephemeris is based. However, practical measurement results often belong to a national or local coordinate system. In real applications, one needs to solve the transformation parameters in order to transform coordinates.

GPS Time System

GPS has a dedicated time system for precision navigation and positioning. The GPS time system, abbreviated as GPST, is provided by atomic clocks in GPS monitoring stations.

GPST belongs to the atomic time system. It has the same interval unit of one SI second as TAI (International Atomic Time), but a different point of origin from TAI, so there is an integer-second offset of 19 s in any instant of time between GPST and TAI, such that $TAI - GPST = 19$ s. GPST was consistent with UTC (Coordinated Universal Time) at 0 h on January 6, 1980. Over time there is an offset of integral multiples of 1 s.

Features and Functions of GPS

With the employment of high-orbiting ranging systems, GPS defines its basic observed quantity as the distance between stations and the satellites. There are two main GPS measurement strategies for obtaining the observed quantity. One is the pseudo-range measurement, which measures the propagation time taken for the pseudo-random code to travel from the satellite to the user's receiver. The other is the carrier phase measurement, which records the phase difference between the carrier signals from the GPS satellites with Doppler frequency shift and the reference carrier signals produced by receivers. Pseudo-range measurement has the highest speed in positioning whereas carrier phase measurement has the highest precision in positioning. The three-dimensional position of a receiver can be deduced through the simultaneous pseudo-range or phase measurements of four or more satellites.

With the appearance of GPS, electronic navigation technology has entered a brilliant period. Compared with other navigation systems, GPS is mainly distinguished by the following (see Xu et al. 1998):

Continuous Global Coverage. Since there are enough GPS satellites in a reasonable distribution, at least four satellites can theoretically be observed continuously and synchronously from any point on the globe, which guarantees all-weather global continuous navigation and positioning in real-time.

Multifunction and High Precision. GPS can provide three-dimensional position, velocity, and time information continuously with high precision for all kinds of users.

High Speed in Real-Time Positioning. One-time positioning and velocity measurement of GPS receivers can be done within 1 s or even less, which is especially important for high dynamic users.

Remarkable Anti-interference Capacity and Adequate Confidentiality. Because of the employment of pseudo-noise spread spectrum technology in GPS, the signals from GPS satellites have remarkable anti-interference capacity and sufficient confidentiality.

GPS technology has developed into a new multidomain (land, marine, aerospace), multipurpose (in-transit navigation, precise positioning, precise timing, satellite orbit determination, disaster monitoring, resource survey, construction, municipal planning, marine development, traffic control, etc.), multimodel (geodesic type, timing-based, hand-held, integrated, vehicle-borne, ship-borne, airborne, satellite-borne, missile-borne, etc.), high-tech international industry. GPS has been widely applied in aerospace, fishing, guided tours, and agricultural production. It is said that “the applications of GPS are only restricted by human imagination.”

GPS Measurement and Positioning Methods

GPS positioning methods are diverse. Users can use different positioning methods that are appropriate for their different purposes. GPS positioning methods can be classified according to different criteria as follows (see e.g., Liu et al. 1996):

According to Observed Values Adopted by Positioning

Pseudo-Range Positioning. The observed values adopted in the pseudo-range positioning are GPS pseudo-ranges, which can be a C/A code pseudo-range or a P code pseudo-range. This positioning method has the advantages of simple data processing, a low demand for positioning conditions, no integer ambiguity, and an easier realization of real-time location. The disadvantage is the low accuracy of observed values. The accuracy of the C/A code pseudo-range observations is generally 3 m and that of the P code pseudo-range observations is about 30 cm, which results in low accuracy of positioning results.

Carrier Phase Positioning. The observations adopted in the carrier phase positioning are GPS carrier phase observations; namely, L_1 carrier, L_2 carrier, or a linear combination of these. The advantage of carrier phase positioning is the high accuracy of observations, with a tolerance of better than 2 mm; however, it is complicated in data processing and has integer ambiguity.

According to Modes of Positioning

Absolute Positioning. Absolute positioning, also known as precise point positioning (PPP), is a positioning model in which one receiver is used. The absolute coordinates of the receiver antenna are determined in this mode. The mode is simple in operation, so it can be used in stand-alone operation. It is generally used for navigation and other applications of low accuracy.

Relative Positioning. Relative positioning, also known as differential positioning, employs more than two receivers to observe simultaneously in order to determine the mutual relationship between positions of the receiver antennae.

According to the Time Used for Obtaining the Results of Positioning

Real-Time Positioning. Real-time positioning, based on the observational data of the receiver, calculates the position of the receiver antenna in real time.

Non-Real-Time Positioning. Non-real-time positioning, also known as post-processed positioning, determines the position of the receiver antenna through the post-processing of the data received by the receiver.

According to the Receiver's State of Motion During Positioning

Kinematic Positioning. So-called kinematic positioning means that the position of the receiver antenna is changing with the time in GPS positioning. That is to say, in data processing, the position of the receiver antenna is seen to be variable over time.

Static Positioning. So-called static positioning is where the position of the receiver antenna in the whole process of observation remains the same. That is, in data processing, the position of the receiver antenna does not vary with the time. In a survey, static positioning is generally used for precise positioning. The specific observation model is to carry out static synchronous observations by multiple receivers at different stations for several minutes, hours, or days.

GPS Receiver

Navigation Receiver

GPS pseudo-range navigation is the most basic GPS service mode. GPS navigation uses the observation of distance (i.e., pseudo-range observations containing errors) to more than four satellites to determine the position of the receiver. Navigation-based GPS receivers (different from the phase measurement-based ones) generally carry out the pseudo-range and Doppler measurements only by C/A code or P code. They can receive navigation messages and calculate the position and velocity of the antenna in real time. Except for American military and authorized users, users can generally only use C/A code. As the most widely used receivers at present, they can be used in military and civil navigation, providing positioning with a medium degree of accuracy, and time transfer of relatively high precision.

Although GPS navigation receivers vary in type, their functions and operations are similar. The work process of a common navigation receiver is as follows:

1. Power on.
2. Wait for satellite searching. The receiver searches automatically for satellites that can be observed in the sky and locks onto the target, which will take a period of time varying from seconds to minutes according to the different types of receivers.

3. Display the positioning results. As soon as the receiver locks onto four (or more than four) satellites, it will begin the positioning and display the results. Position and velocity are generally displayed. The receiver constantly updates positioning and velocity results based on the selected data update rate.

Phase Measurement Receiver

Since the carrier wave is much shorter in wavelength than the pseudo-random code, in the case of the same resolution the observation accuracy of the carrier phase is much higher than that of the code phase. For example, the wavelength of the carrier L_1 is 19 cm, so the error of the corresponding distance observations is about 2 mm, whereas for the carrier L_2 the corresponding error is about 2.3 mm. The carrier phase measurement is the most accurate method nowadays and many companies have produced GPS phase measurement receivers of different types. The MacrometerV-1000 produced by the American company Litton Aero Service is a single frequency (L_1) phase measurement receiver and is the earliest manufactured for commercial use.

2.3.2 *Satellite Laser Ranging*

Satellite Laser Ranging (SLR), rising in the mid-1960s, is a new space geodetic technique that determines the distance between laser station and a satellite using the laser ranger to trace and observe the satellites installed with laser reflectors. To begin with, SLR employed the BE-C satellite. Then, in 1976 NASA launched a laser geodynamic satellite LAGEOS-1; in 1992 America cooperated with Italy and launched the satellite LAGEOS-2 to expand the observation range of SLR on the Earth. Meanwhile, France, the former Soviet Union, Japan, and Germany successively launched their SLR satellites. During its development over more than 40 years, the SLR system has improved from 1 m to the present 1 cm in distance accuracy. It is now one of the main technical means of precise satellite positioning as well as the most precise among the current various space observation technologies in terms of data sampling rate and absolute positioning. It not only plays a decisive role in establishing and maintaining the Global Geocentric Coordinate System (GGCS) but has also led to great achievements in the field surveying of modern plate motion, improvements in the Earth gravity model and geocentric gravitational constant, and the accurate measurement of Earth's rotation parameters (ERP), etc. (see e.g., Ye and Huang 2000).

Principles of Satellite Laser Ranging

SLR is a physical distance-measuring method, using the laser as its light source and the time of flight of the optical pulse for measurements. Its main features are:

1. The output power of the laser can reach orders of magnitude of 10^9 W and its optical energy density per unit area can be greater than that of the surface of the sun. Thus, the effective distance of the laser can reach the artificial Earth satellites tens of thousands of kilometers away, or even the lunar surface.
2. The laser spectrum is very sharp and has a halfwave width of about 5° , which benefits from adopting a narrow-band filter to eliminate sky background noise in a receiving optical system and to improve signal-to-noise ratio in observation.
3. The divergence angle of light beam output by the laser is very small, at about 1 mas. Through the optical system alignment, the divergence angle can be further compressed. Therefore, the light energy can still be concentrated within a very small scope far away.
4. The laser burst of a pulsed laser can reach a very small order of magnitude in width. Because the pulse width is one of the main factors in determining ranging accuracy, the laser ranging can be very accurate.

Due to the aforementioned characteristics of the laser, it is possible to realize long-distance laser ranging. There are three methods of laser ranging: the pulse method, phase method, and interference method. The pulse method is usually applied in SLR. Its basic principle is very simple: laser pulse signals are sent from a laser ranger placed at the observation station to a laser satellite equipped with a back-reflecting prism and go back to the receiving system of the rangefinder after being reflected by the tested satellite. If the time difference Δt between the sending and receiving of these laser pulse signals is measured, we can get the distance ρ between the satellite and the station according to the formula:

$$\rho = \frac{1}{2} c \Delta t, \quad (2.11)$$

where c is the velocity of light. Suppose the equation of motion of satellites in the Earth-Centered Inertial (ECI) Coordinate System is:

$$\dot{\mathbf{X}} = F(\mathbf{X}, P_d, t), \quad \mathbf{X}(t_0) = \mathbf{X}_0, \quad (2.12)$$

where \mathbf{X} is the satellite's state vector at an instant of time t ; $\mathbf{X} = (r, r_0)^T$ or $\mathbf{X} = \sigma$, σ being the six Keplerian elements; \mathbf{X}_0 is the satellite's state vector at initial time t_0 ; and P_d is the physical parameter to be estimated. Then the solution of (2.12) can be expressed as:

$$\mathbf{X} = Q(\mathbf{X}_0, P_d, {}^d t). \quad (2.13)$$

Suppose Θ_O is the satellite's observed quantity (i.e., the distance between the satellite and the Earth); its corresponding theoretical value Θ_C can then be defined as:

$$\Theta_C = \Theta(\mathbf{X}, \mathbf{R}), \quad \mathbf{R} = \mathbf{PNSR}_0. \quad (2.14)$$

where \mathbf{R} and \mathbf{R}_0 represent the position vectors of the station expressed in the inertial system and in the Earth-fixed coordinate system, respectively, and \mathbf{P} , \mathbf{N} , and \mathbf{S} stand for precession of the equinoxes matrix, the nutation matrix, and the rotation matrix of the Earth (see Vermeille 2002).

The above is the general principle of satellite dynamic geodesy. In practice, we should, based on different situations, purposes, and demands, select appropriate parameters as the adjusted quantities and keep the theoretical values of other parameters unchanged. Whichever dynamic geodesy is chosen, the satellite orbit is generally needed as the adjusted quantity, i.e., they all have a process of orbit determination and measurement.

SLR System

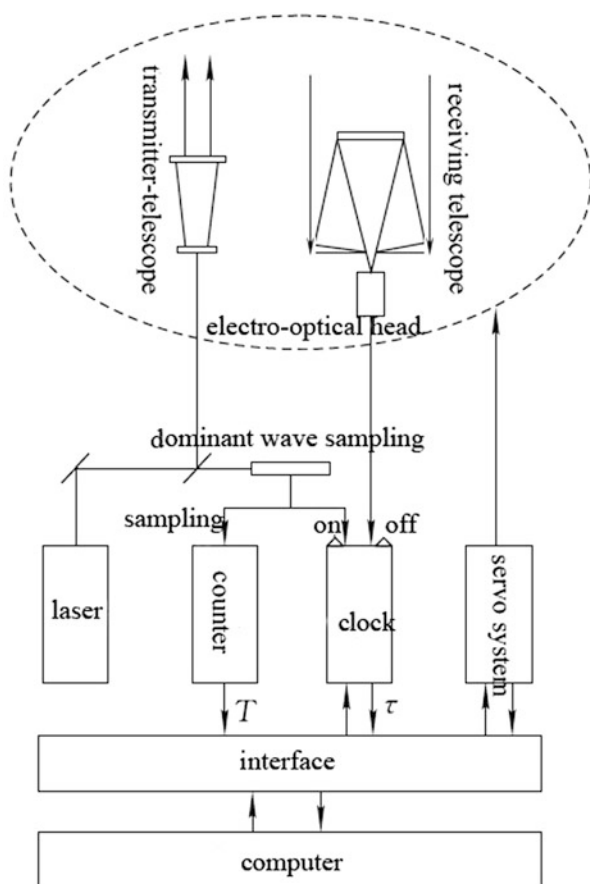
The SLR system consists of two main segments, a laser ranger on the ground and a laser satellite in space. The hardware devices of the ranger include seven parts: laser, telescope, electro-optical head, pulse position measurement system, time and frequency system, servo system, and computer; see Fig. 2.11 for an outline of the structure.

The working principle of the satellite laser ranging system is as follows:

The light pulse generated by the laser is led through the guiding optical path into the transmitter-telescope, which then emits the light beam to the target laser satellite after collimation. A small segment taken from the emitted light beam forms two electric pulses through the dominant wave sampling circuit. One is called the dominant wave pulse, which, as the enabling signal, is used to initiate the laser flight time interval counter. The other pulse is the electrical pulse, used to sample the clock and record the time of the laser emission. The laser pulse reflected by the satellite to the ground is received by the receiving telescope. Low-light detection equipment, which detects the reflected light, is set on the focus of the receiving telescope. The light is converted into electrical signals that are then amplified and reshaped into an echo pulse, which, as the stopping signal, stops the counter. In this way, the counter records the time interval $\Delta t = \tau - T$ between the dominant wave and the echo pulse (shown in Fig. 2.11), where Δt is the round-trip flight time of the laser between the station and the satellite.

The telescope of the ranger has three functions; namely, emitting, receiving the laser, and targeting satellites. We can design it as three independent telescopes or one telescope with three functions. Its time and frequency system has two functions: one is to provide a stable frequency source for counters, lasers, computers, and other devices (the stability of counters' frequency should be better than 10^{-10}); the other is to record the time for laser emission. The accuracy of time recording for

Fig. 2.11 Structure of the SLR system



the rangefinder with centimeter-level accuracy is $1\ \mu\text{s}$, and good quartz or rubidium clocks can meet the above requirements.

The laser ranger can only make an observation of satellites with dedicated reflectors. The laser reaching the satellite should return in the same direction as that of the laser emission. Such a reflector is also called a retroreflector, which is mainly composed of glass prisms. To achieve the accuracy required, the reflector must be carefully designed for the geometric shape and orbital altitude of dedicated satellites. To adjust the energy balance between the emitting laser and the receiving photons, the reflector should be big enough to reflect enough energy. In most cases, several single reflectors with 2–4 cm diameter set in a certain array can acquire the necessary energy. We should pay close attention to the alignment and adjustment of a single reflector lest the signal overlap causes a pulse distortion. The reflector is a passive device, easy to install on a satellite as an attachment. Therefore, many satellites are equipped with laser reflector arrays. Satellites LAGEOS-1 (in 1976) and LAGEOS-2 (in 1992) launched by NASA, and satellites Etalon-1 (in January

1989) and Etalon-2 (in May 1989) launched by Russia (the former Soviet Union), and others are dedicated laser ranging satellites for the applied research of geodynamics and geodesy. Satellites are especially suitable for geodetic research and observation because of their high and stable orbits, small area-to-mass ratio, symmetrical spherical shape, long accumulation time for observational data, etc. Figure 2.12 shows the shape of such satellites. They are spherical with a diameter of 60 cm, installed with 426 laser reflectors on their surfaces. They are still used for the common geodetic observation of SLR nowadays and have, among others, provided much data for establishment of the terrestrial reference frame and determination of the Earth rotation parameters. SLR has become one of the main techniques for satellite orbit determination because of its high precision distance measurement ability. Many reconnaissance satellites, meteorological satellites, Earth resources satellites, and oceanic satellites have all been equipped with laser reflectors so as to carry out more precise measurement and control of satellites by means of the SLR technique.

2.3.3 *Very Long Baseline Interferometry*

Very long baseline interferometry (VLBI) came into development in the late 1960s and is a radio interferometric observation technique that can combine two radio telescopes thousands of kilometers apart into a radio interferometry system with very high resolution. Since the line between the two telescopes is known as the baseline, VLBI is called very long baseline interferometry. Its resolution has now been improved to (the magnitude of) 0.1 mas with extension of the baseline. Because of the super-high resolution of VLBI, it has been widely applied in many fields like astronomy, geophysics, geodesy, and space technology for applications such as radio astronomy, accurate determination of the Earth's rotation parameters, crustal deformation detection, exploration of deep space and the ionosphere, etc.

Principles of Geodetic VLBI

Celestial bodies observed by VLBI are extragalactic radio sources, which are usually in deep space 100 million light years from Earth. When the electromagnetic wave radiated from the celestial bodies reaches the Earth's surface, its propagation distance is much further than that of the baseline in VLBI, so at this moment the movement of the wave front can be assumed to be parallel propagation and the wave is called a plane wave. On account of the different distances between the two antennae and a certain radio source, we get a distance difference L . Therefore, the time span of the radio signal from the same wave front to either antenna will be different, resulting as a time delay τ_g . According to the geometric relationship shown in Fig. 2.13 we get:

Fig. 2.12 Satellite
LAGEOS. Source: NASA

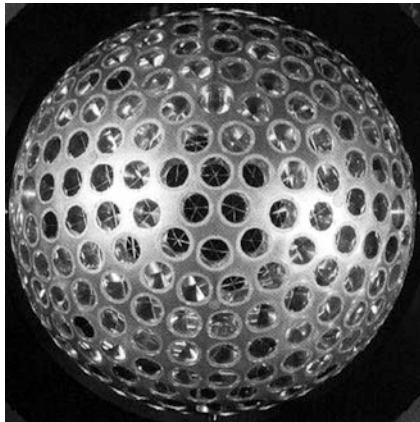
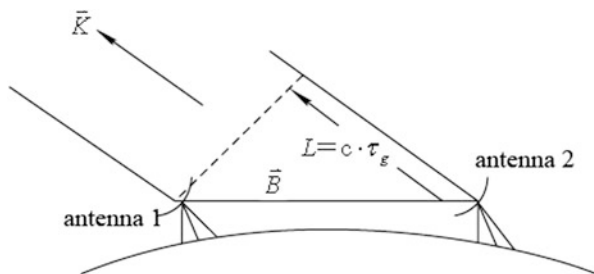


Fig. 2.13 Geometric
principle of VLBI



$$L = c \cdot \tau_g, \quad (2.15)$$

where c is the velocity of light in vacuum. Suppose \vec{B} is the baseline vector from antenna 1 to antenna 2 and \vec{K} is the direction of the electrical source observed; then we get:

$$\tau_g = -\frac{1}{c} (\vec{B} \cdot \vec{K}). \quad (2.16)$$

Due to the movement of the Earth, the direction of vector \vec{K} relative to the baseline vector \vec{B} will change. Suppose τ_g is a time function, then we can denote its time derivative as the delay rate $\dot{\tau}_g$, namely:

$$\dot{\tau}_g = -\frac{1}{c} \frac{\partial}{\partial t} (\vec{B} \cdot \vec{K}). \quad (2.17)$$

The observed values of VLBI in geodesy are mainly the delay and the delay rate.

\vec{B} and \vec{K} in (2.16) and (2.17) must be expressed in the same coordinate system. However, the direction of the radio source is usually represented by the right ascension and the declination (α, δ) in the celestial coordinate system, and the baseline vector is defined as a vector $\vec{b} = (\Delta X, \Delta Y, \Delta Z)$ in the terrestrial coordinate system. As a result, in practical calculation, must be converted to a vector in the celestial coordinate system (see Dennis and Petit 2004):

$$\vec{B} = \mathbf{PNSW}\vec{b}, \quad (2.18)$$

where \mathbf{P} , \mathbf{N} , \mathbf{S} , \mathbf{W} stand for the rotation matrixes of precession, nutation, Earth's diurnal rotation, and Earth's polar motion, respectively.

For simplification, the influence of precession, nutation, and polar motion is not taken into account when discussing the principles of VLBI. Thus, (2.18) can be expressed as:

$$\vec{B} = \mathbf{R}_z(-\theta_g)\vec{b} = \begin{pmatrix} \Delta X \cos \theta_g - \Delta Y \sin \theta_g \\ \Delta X \sin \theta_g + \Delta Y \cos \theta_g \\ \Delta Z \end{pmatrix}. \quad (2.19)$$

Substitute (2.19) into (2.16) and (2.17) to obtain:

$$\tau = -\frac{1}{c} [\Delta X \cos \delta \cos (\theta_g - \alpha) - \Delta Y \cos \delta \sin (\theta_g - \alpha) + \Delta Z \sin \delta], \quad (2.20)$$

$$\dot{\tau} = -\frac{1}{c} [\Delta X \omega_g \cos \delta \sin (\theta_g - \alpha) + \Delta Y \cos \delta \cos (\theta_g - \alpha)], \quad (2.21)$$

where θ_g is Greenwich local sidereal time and ω_g represents the Earth's rotation speed.

The two equations above are the principle formulae for solving geodetic parameters by the use of the observed quantities of VLBI delay and delay rate. Through analysis of the formulae we know that the solution of VLBI parameters has the following characteristics:

1. VLBI delay and delay rate are pure geometric observed values that do not contain information about the Earth's gravitational field. Therefore, the acquisition of these observed values is not affected by the Earth's gravitational field.
2. VLBI, as a relative measurement technology, can only determine the relative position between two antennae, that is, the baseline vector; it cannot get the geocentric coordinates of each antenna. Thus, for determining the geocentric coordinates of the VLBI station we usually make observations through both VLBI and SLR in one station at the same time. Using the geocentric coordinates obtained by SLR as the datum, we can calculate the geocentric coordinates of other VLBI observations.

3. Because of its direct relation with the alteration of the Earth's rotation θg , the right ascension α of the radio source cannot be calculated independently from the observed quantities of the delay and the delay rate. As a result, VLBI alone cannot determine the origin of the right ascension of the radio source reference system and other technologies are required to do so.
4. The observed quantity of the delay rate does not contain the effect of the baseline component ΔZ , which cannot be calculated just by the observation of the delay rate. Besides, adding the data of delay rate to that of delay will not reduce the level of radio sources, which are to be observed to solve all the unknown parameters. In data processing and parameter solution, the delay rate is adopted only as a supplementary observation while the observed quantity of the delay is decisive.

The VLBI System

The VLBI system, as shown in Fig. 2.14, contains antennae, receivers, local oscillators, samplers, recording devices, related processors, and other units. The following is a brief introduction of the process of data collection of VLBI.

1. First, two antennae in the system receive the radio signal emitted by the observed radio source on the focus points of the antennae's paraboloid. Then, the feed source transfers the collected electromagnetic wave into high-frequency current and sends it to the receiver. The accuracy of VLBI observation (time delay, time delay rate) with regard to celestial and geodetic observation is directly proportional to the signal-to-noise ratio (SNR) of the system, and the SNR is directly proportional to the antenna aperture. Due to the weak signals from extragalactic radio sources, the VLBI antenna aperture is often over 20 m to obtain enough SNR.
2. The receiver's main function is to amplify this signal into a radio-frequency signal by using a high frequency amplifier and then convert it to an intermediate-frequency signal with a certain bandwidth through a mixer. In mixing, the mixer requires a local oscillating (LO) signal, which is provided by the station's local oscillator.
3. The intermediate-frequency (IF) signal from the receiver reaches the data recording terminal device, which nowadays employs an MK3 system or upgraded MK4 and MK5 systems. The MK3 recording system mainly contains two IF distributors, 14 video converters, a format cell data collection system, a magnetic tape recorder, and a computer that takes control of the data collection system and the recorder. The IF signal from the receiver is sent to the IF distributors and then to these 14 video converters, which convert the signals in different IF frequency ranges into video signals (also named Base Band) at 0–2 MHz that can be recorded by the magnetic tape recorder. The video signal output by the video converters is sent to the format cell. The main function of the format cell is to digitalize the 0–2 MHz video signals through one-bit sampling; then a format encoder supplies the precise receiving time for each datum and

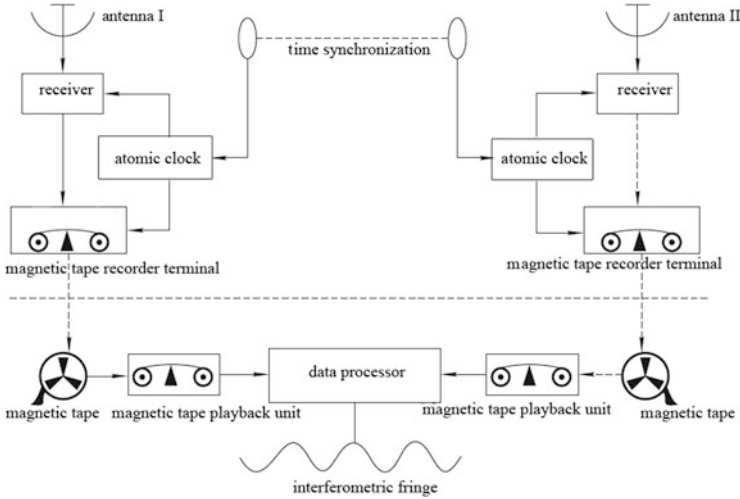


Fig. 2.14 VLBI system

encodes these data by rewriting signals and necessary information in a special format. Next, the formatted data are recorded on dedicated tapes in specialized format by the magnetic tape recorder. It should be noted that these 14 frequency converters have 14 independent LO signals, which will result in phase drift. So, phase calibration is required. A phase calibration system consists of a pulse generator, which transmit an impulse and inject it into the signals every micro-second. This pulse injection point is defined as the reference point of the delay.

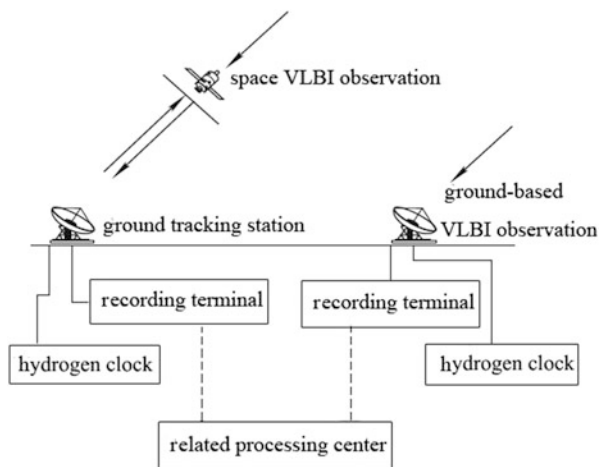
4. Finally, the observed data recorded by the magnetic tape recorder are sent to the related processors. The processors playback the data and input them to the correlators of corresponding channels to carry out cross-correlation computations and acquire the related function value, i.e., the interferometric fringe. After that the computer uses its software system to obtain the required observed value of the time delay and the time delay rate by fringe fitting computation.

The Technique of Space VLBI

To improve the resolution of VLBI, the concept of space VLBI was proposed in 1970 and the establishment of the space VLBI system was also considered. By 1980, space VLBI became more mature in theory and technical realization. In 1997, the first space VLBI satellite (VSOP) in human history was successfully launched in Japan. Although space VLBI was proposed for astrophysics research, conceptually it has more advantages over ground-based VLBI for application in fields like geodesy. Therefore, it will become a more effective geodetic observation technology.

In light of the VLBI principle, there are no differences between space and ground-based VLBI. A space VLBI station can be seen as a component of a

Fig. 2.15 Space VLBI system



ground-based VLBI net extending into space. It has the same function as a ground antenna, i.e., to receive signals from a radio source. Then, the required observational data for scientific research can be acquired through correlation processing of the signals received by both the space VLBI station and the ground antenna. However, space VLBI is different to ground-based VLBI in technical realization since the former places the antenna in space. The system of space VLBI is shown in Fig. 2.15.

Space VLBI is mainly distinguished by:

1. The phase of the space station's local oscillator is locked onto the hydrogen frequency standard of the ground tracking station. This frequency standard is sent to the space station from the tracking station through a radio channel (upwards).
2. The radio signal and other data received by the space station are sent back to the ground tracking station through a radio channel (downwards).
3. The space station must be equipped with highly precise systems for the attitude adjustment of antennae and for orbit control and detection.
4. The space station produces its own source of energy from received solar energy.
5. A ground support system of global coverage that can maintain uninterrupted communication with the space station is required.

The most significant technical advantage of the application of space VLBI to geodesy is to turn the geometric measurement of ground-based VLBI into dynamic measurement. It has been mentioned that the measurement completed by forming a baseline between two ground VLBI observations is a geometric one from the perspective of geodesy. Such measurement alone can only determine the relative position of the two stations but not their geocentric coordinates. Since the orbit of the space VLBI is described in a geocentric coordinate system and its movement is affected by various geodynamic factors, when adopting space VLBI a dynamic

measurement system can be formed by making a baseline between the space and ground stations so as to determine directly the geocentric coordinates of the ground-based station. Because all the VLBI antennae around the world take part in the space VLBI observation, a complete terrestrial reference system can be established independently using space VLBI technology itself. A VLBI station can not only be an observed object for various artificial satellite tracking stations on the ground, it can also be a space station in the orbit of artificial satellites to observe extragalactic radio sources directly. Thus, the direct connection and unification of the artificial satellite dynamic reference system and the radio source reference system can be realized. In addition, by means of space VLBI, an agreement can be made in VLBI about conversion between the Conventional Terrestrial Reference System (CTRS) and the Conventional Celestial Reference System (CCRS) to obtain a unified celestial and terrestrial reference system (i.e., a unified rotation and scale system with a commonly defined origin). Such unification of coordinate systems is of great significance for research in geodesy and other related fields.

2.3.4 Satellite Altimetry

In the 1980s, satellite altimetry (SA) appeared along with the application and development of computer technology, space technology, satellite telemetry, and remote sensing technology. SA employs microwave radar altimeters installed in satellites, radiometers, synthetic aperture radar, and other equipment to measure in real time the distance from a satellite to the ocean's surface, the effective wave height, and the backscattering coefficients, and to carry out research in geodesy, geophysics, and oceanography through data processing and analysis.

SA data can determine the marine geoid and solve the gravity anomaly of the ocean to compensate for the data gap in gravity measurement of marine areas. Therefore, SA plays an important role in establishing an Earth gravity field model with high accuracy and high resolution. The US Federal Geodetic Control Subcommittee (FGCS) noted that what the ocean altimetry satellite Seasat does within 3 months would take 200 years and cost 2 billion US dollars if done by marine gravimetry. Besides, SA data can also be used in oceanographic studies such as the measurement of the width, boundary, and velocity of ocean currents; tidal fluctuations; sea surface topography; and mean sea level changes.

The Basics

In SA, a microwave radar altimeter mounted on a satellite (the carrier) transmits microwave signals to the ocean's surface. This radar impulse reaches the ocean's surface and then returns to the radar altimeter by reflection. According to echo theory, we can obtain three observed quantities after the return of the radar pulse, including:

1. The round-trip time of the radar pulse going from the satellite to the ocean's surface and back to the satellite: the measured value of satellite altitude.
2. The waveform of an echo signal, which contains an ascending front, a flat top, and a decay area.
3. The amplitude of an echo signal, that is, the automatic gain control value of the signal. Based on the analysis of the waveform, structure, and round-trip time of the echo signal, we can obtain information like sea level altitude, sea level tilt, ocean currents, effective wave height, sea surface backscattering coefficients, and wind field (see Leuliette et al. 2004).

In satellite altimetry, the satellite is seen as a mobile platform on which a radar altimeter transmits a microwave pulse to the Earth's surface and receives the signal reflected back. Suppose that the altitude of the satellite is and the propagation velocity of the signal is c ; we can then use Δt , the round-trip time of the radar signal observed, to calculate:

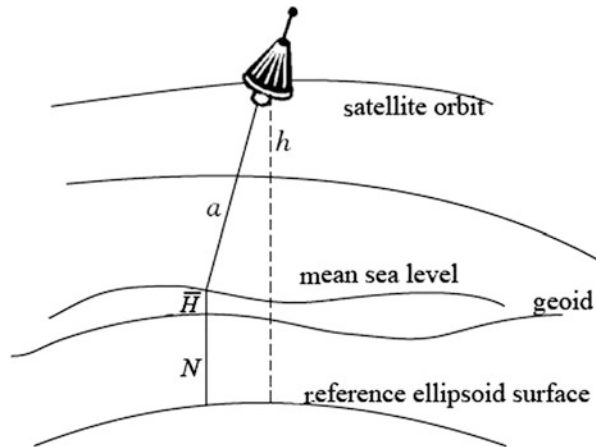
$$a = c \frac{\Delta t}{2}. \quad (2.22)$$

Because of water's good reflective properties, this method is particularly suitable for marine areas. The radiation of radar signals can instantaneously cover an annular region with a radius of thousands of kilometers on the sea surface (often referred to as the signal footprint). The size of the annular region is related to the spatial resolution of the incident microwave beam. So, the observed value is an elevation to the average instantaneous sea level. Its difference from the geoid height is \bar{H} . Assume that the satellite altitude with respect to the reference ellipsoid is h , which can be derived from the satellite orbit in the geocentric reference system through calculation. If neglecting additional corrections, we can simplify the basic satellite altimetry equation to:

$$h = N + \bar{H} + a. \quad (2.23)$$

Figure 2.16 shows that the radar altimeter can be adopted to scan the sea surface directly so as to scan the marine geoid approximately. Therefore, satellite altimetry is an effective method for directly drawing a geoid map. It is important that it can detect a very large marine area in a rather short time and make out a detailed sea level expression with a very high spatial-temporal resolution. \bar{H} means a disturbance (noise) in establishing the geoid, but is an observed signal for the research of ocean dynamics. From the extensive analysis of \bar{H} , we can obtain an important understanding of the structure of the ocean floor and seabed features.

Fig. 2.16 Rationale of satellite altimetry

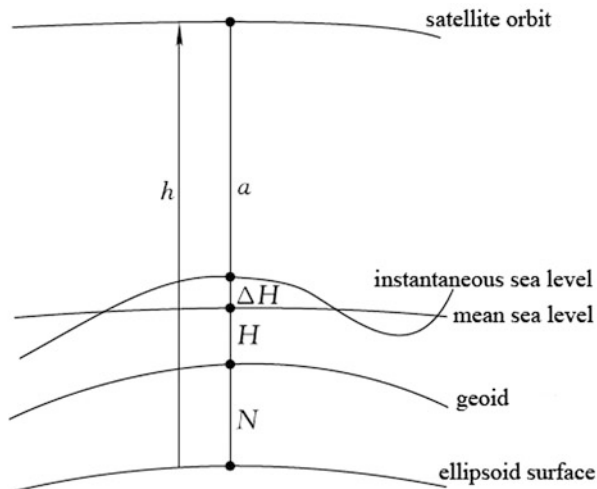


Satellite Altimeter and Its Operational Principle

A satellite altimeter is a satellite-borne microwave radar that usually consists of a transmitter, a receiver, a timing system, and a data collection system. It is generally 13.9 GHz in the emission frequency with 2 kW in transmitting power and works at an altitude of 800 km. Its radar antenna is parabolic with a diameter of 0.6–1 m. To guarantee at the same time the accuracy of measurement, the resolution, altitude, and other indices, the radar pulses transmitted must have a comparatively large time–frequency bandwidth. Thus, the satellite altimeter employs the pulse compression technique to transmit and receive pulses. The compressed pulses can be of nanoseconds (10^{-9} s) in width, which means that the pulse compression technique has solved a big problem in radio theory in that the width of pulses in the time domain and frequency domain cannot be enlarged simultaneously. The product of the width of pulses in the time and frequency domains is referred to as the compression ratio.

The operational principle of a satellite altimeter is as follows: The transmitter sends modulated compressed pulses to the Earth's surface with a certain pulse repetition frequency (PRF) through its antenna; after the reflection by the ocean's surface, the pulses return and are received by the receiver, which measures the time difference between the transmitted pulse and the received pulse. According to this time difference and the reflected waveform, we can determine the distance from the satellite to the ocean's surface. The width of the radar beam transmitted by the satellite is approximately 1° ; as a result, when the radar beam reaches the ocean's surface, the radius of the signal track is about 3–5 km. Therefore, the distance measured by the altimeter is equal to the average distance from the satellite to the circular area with the radius of 3–5 km. On this basis, instrument adjustment, sea state correction, tropospheric refraction correction, ionospheric effect correction, and correction for periodic effects of the sea surface, etc. have to be taken into account.

Fig. 2.17 Geometric relationship in satellite altimetry



The Observed Quantity in Altimetry and Error Analysis

Equation (2.23) is a simplified observation equation of the altimeter, which needs to be refined in practical application. The geometric relationship in SA shown in Fig. 2.17 yields:

$$h = N + H + \Delta H + a + d, \quad (2.24)$$

where h is the ellipsoidal height of the altimeter satellite based on orbit computations, N is the geoid height, H is the sea surface topography, ΔH is caused by the instantaneous tidal effect, and a is the altimeter measurement.

$H + \Delta H = \bar{H}$ in (2.23). The quantity a is observed by the altimeter and requires atmospheric correction, which should be referred to the satellite's center of mass. The difference between the geoid and the mean sea level is called the sea surface topography, which can reach 1–2 m. The mean sea level is defined as a stationary sea surface not changing with time. The difference between the sea surface and the geoid is caused by differences in seawater salinity and temperature, a wide range of pressure differences, strong tidal currents, etc. For a resolution of better than 2 m, it is no longer valid to use mean sea level as a close approximation to the geoid. Besides, there are difficulties in connecting up the height systems obtained by a variety of tidal observational stations.

Errors and corrections of altimetric observations are categorized into three types; namely, the difference between the actual and the calculated orbit (the orbital error), the effects on signal propagation path, and the difference between the instantaneous sea surface and the geoid.

The orbital error mainly results from the accuracy of the Earth gravity model used in orbit calculation, errors in the tracking stations' coordinates, errors or limitations in the tracking systems, and errors in the model used for orbit

calculation. The most important impact generally comes from the Earth's gravity field. Since each satellite is only particularly sensitive to one certain subset of spherical harmonic coefficients, it is an effective method to develop a particular gravity model for the observed quantity of a particular satellite. For instance, use of the gravity model GEM10 in the altimetry satellite GEOS-3 improves the accuracy of the satellite's orbit from 10 m to 1–2 m. The tracking system is the second important factor to affect orbit accuracy. To acquire a high-precision orbit, orbit determination using on-board GPS is required, and the accuracy of the coordinates has reached several centimeters. Even so, the remaining orbit error is still much greater than the altimeter's accuracy. Consequently, we have to improve the orbit determination models, adopt some non-dynamic methods, etc.

The effects on the signal path can be categorized into instrumental errors and propagation errors. The major instrument effects include the distance between the phase center of the radar antenna and the satellite's center of mass, the propagation delay in the electronic circuit of the altimeter, and the timing error in the measurement system. In the manufacture of altimeters, these effects can be reduced to a minimum and can be estimated. All the effects of instrumental errors should be determined and kept under control when calibrating an altimeter in the test area with measured accuracy. The signal propagation error caused by ionospheric refraction is about 5–20 cm and depends on the ionization intensity, whose effects can be corrected by dual frequency. The effect of tropospheric refraction is about 2.3 m. Since only the observed quantity in the vertical direction is adopted, the effect can be well corrected by a proper atmospheric refraction model to reach an accuracy of several centimeters. Propagation errors also include the impact of actual sea conditions on the reflected signals.

The discrepancy between the instantaneous sea surface and the geoid can be divided into a time-invariant part H and a time-dependent part ΔH . Before determining the mean sea level using the altimetry observed quantity, the time-dependent component should be corrected. The wave-induced sea level change, which has been smoothed out in the altimeter's observation, can be negligible. Therefore, the correction contributor to be considered is mainly the sea level change induced by tides and changing atmospheric pressure fields.

The resolution and accuracy of the sea surface height measured directly from satellite altimetry can reach 5 km and 5 cm, respectively. However, under the effect of the sea surface topography, the tides, and errors in the environmental correction model, the accuracy of the sea geoid can rarely be better than ± 10 cm.

2.4 Gravimetry

2.4.1 Absolute Gravimetry

Absolute gravimetry (absolute gravity measurement) is a technique utilized to determine the gravity value (actually, gravitational acceleration) at a defined geometric point. There are two methods for absolute gravity measurement, one using a reversible pendulum, and the other by means of the free-fall motion of bodies. The second method of measurement has been the dominant method since the 1960s and will be discussed here.

Free-fall motion refers to the accelerated linear motion of a body along the plumb line under the action of gravity only. According to mechanics, if the gravity acceleration g in the course of motion along the plumb line is assumed constant (no gravity changes with height), then the equation of motion is:

$$l = l_0 + V_0 t + \frac{1}{2} g t^2, \quad (2.25)$$

where V_0 and l_0 denote the initial velocity of the falling body and the distance from the origin O , respectively, at the computational time $t = 0$, and l is the distance of the falling body from the origin O after a period of time t , cf. Fig. 2.18.

Two methods can be used to determine the gravity value by means of the free-fall motion of bodies: the free-fall method and the symmetrical rise-and-fall method (abbreviated as the rise-and-fall method). Their principles are discussed below.

Free-Fall Method

In (2.25), to avoid determining V_0 and l_0 , it is necessary to measure from at least three positions. Assume at time t_1 , t_2 , and t_3 that the distances of the falling body from the point O are l_1 , l_2 , and l_3 , respectively, as illustrated in Fig. 2.19, where the transverse axis indicates time and the ordinate axis indicates distance. With reference to (2.25), for each time period there will be a corresponding equation of motion which gives:

Fig. 2.18 Free-fall motion of an object

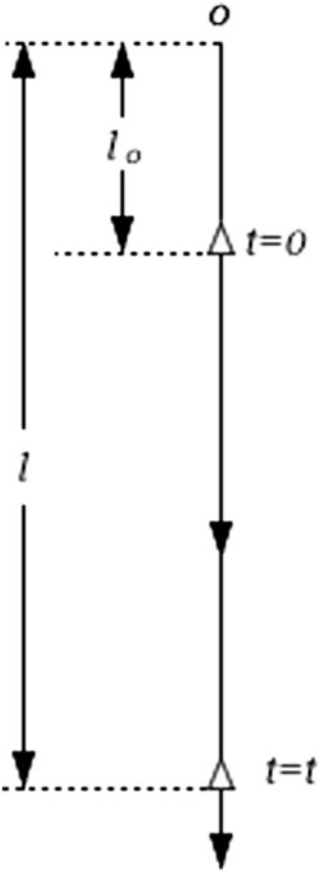
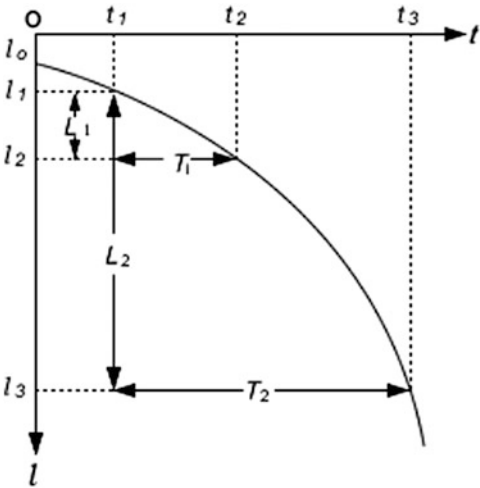


Fig. 2.19 Gravity determination by the free-fall method



$$\begin{aligned}
l_1 &= l_0 + V_0 t_1 + \frac{1}{2} g t_1^2, \\
l_2 &= l_0 + V_0 t_2 + \frac{1}{2} g t_2^2, \text{ and} \\
l_3 &= l_0 + V_0 t_3 + \frac{1}{2} g t_3^2.
\end{aligned}$$

Subtracting the first equation from the second and third equations, respectively, the results are as follows:

$$\left. \begin{aligned}
L_1 &= V_0 T_1 + \frac{1}{2} g T_1 (t_1 + t_2) \\
L_2 &= V_0 T_2 + \frac{1}{2} g T_2 (t_1 + t_3)
\end{aligned} \right\}, \quad (2.26)$$

where $L_1 = l_2 - l_1$ and $L_2 = l_3 - l_1$ are the distances from the first position to the second and third positions, respectively. $T_1 = t_2 - t_1$ and $T_2 = t_3 - t_1$ are the times taken by the falling body in its motion from the first position to the second and third positions. To eliminate V_0 , the two equations in (2.26) are divided by T_1 and T_2 , respectively, and the two results thus achieved subtracted from each other, which reads as:

$$\frac{L_1}{T_1} - \frac{L_2}{T_2} = \frac{1}{2} g (t_2 - t_3).$$

Since $t_2 - t_3 = T_1 - T_2$, the formula of g can finally be written as:

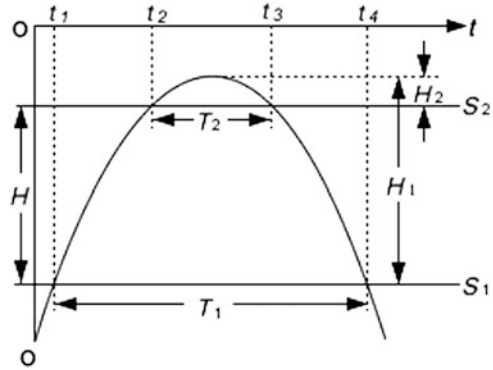
$$g = \frac{2}{T_2 - T_1} \left(\frac{L_2}{T_2} - \frac{L_1}{T_1} \right). \quad (2.27)$$

Thus it can be seen that to determine gravity using the free-fall method requires knowledge of the distances L_1 and L_2 traveled by the falling body within the time periods T_1 and T_2 .

Rise-and-Fall Method

In this method, an object is thrown vertically upward and then allowed to fall freely. To obtain the gravitational acceleration g , it is necessary to label two positions S_1 and S_2 in its course of motion. The time intervals T_1 and T_2 of the falling body past each position are determined, cf. Fig. 2.20. The transverse axis indicates time and the ordinate axis indicates the vertical position of the falling body. Let H_1 and H_2 be

Fig. 2.20 Gravity determination by the rise-and-fall method



the distances from the two measuring positions to the peak of its motion. According to (2.25), where $l_0 = 0$, $V_0 = 0$, we will obtain:

$$H_1 = \frac{1}{2}g\left(\frac{T_1}{2}\right)^2,$$

and

$$H_2 = \frac{1}{2}g\left(\frac{T_2}{2}\right)^2.$$

H is taken to denote the distance between the two positions, and hence yields:

$$H = H_1 - H_2 = \frac{1}{2}g\left[\left(\frac{T_1}{2}\right)^2 - \left(\frac{T_2}{2}\right)^2\right].$$

After rearrangement, the equation for g becomes:

$$g = \frac{8H}{T_1^2 - T_2^2}. \quad (2.28)$$

Thus, it can be seen that to determine gravity using the rise-and-fall method requires the determination of time intervals T_1 and T_2 of the object passing two positions with a distance of H during its rise and fall.

2.4.2 *Relative Gravimetry*

Relative gravimetry is a technique used to determine the gravity difference between two points, and then to obtain the gravity value of each point in a pointwise manner through at least one point of known gravity value.

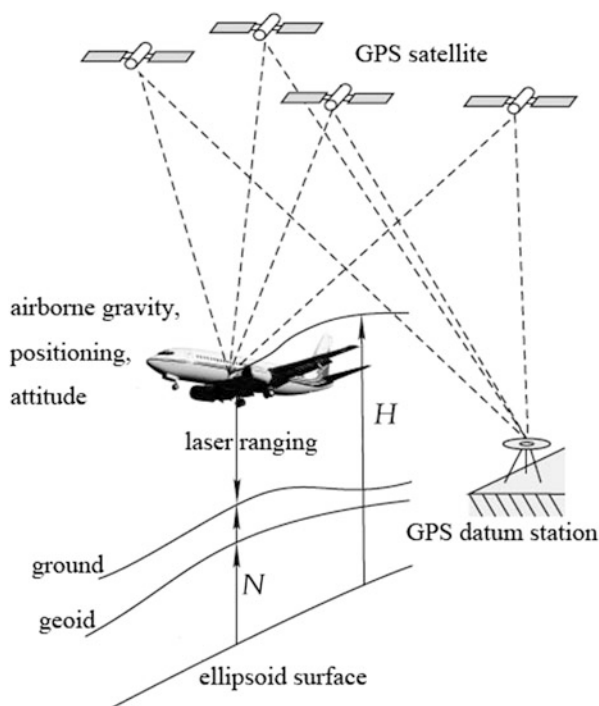
The static method of relative gravimetry is to use a kind of force (such as the spring force) to work against the force of gravity that is acting on the object and balance the gravitational pull. By changing gravity, the location of the equilibrium position (location of the spring) is also changed. As long as the change of the equilibrium position (the amplitude of the spring) is determined, the variation in gravity can be calculated (according to Hooke's Law). The gravity difference between the two locations is thus obtained (Lu 1996).

Currently, the most frequently used gravimeter is called the spring gravimeter, in which the spring force is used to balance gravity. Examples are the quartz spring gravimeter ZSM series manufactured by Beijing Geological Instrument Factory and the (LCR) metal spring gravimeter by LaCoste and Romberg in the USA. Both of these spring gravimeters incorporate a spring mass system, optical system, mechanical device for measurement, instrument panel, and insulated case. The range of gravity difference measured by ZSM is $80 \times 10^{-5} \text{ m/s}^2$ to $120 \times 10^{-5} \text{ m/s}^2$ and the precision of measurement is between $0.1 \times 10^{-5} \text{ m/s}^2$ and $0.3 \times 10^{-5} \text{ m/s}^2$. The LCR gravimeter can also be classified as Model G and Model D. The range of direct measurement of the Model G is up to $7,000 \times 10^{-5} \text{ m/s}^2$. It can be utilized for relative gravity measurement on a worldwide scale and its measurement precision amounts to $\pm 20 \times 10^{-8} \text{ m/s}^2$. The range of direct measurement of Model D is only $200 \times 10^{-5} \text{ m/s}^2$. It is widely used in regional gravity surveys and its precision of measurement is slightly higher than that of Model G.

2.4.3 *Airborne Gravimetry*

Airborne gravimetry is a method employed to determine the near-Earth gravitational acceleration using an integrated airborne gravity remote sensing system, which consists of an aircraft as carrier, airborne gravimeter, GPS, altimeter, and attitude determination devices, etc. (Fig. 2.21). It can operate in areas where terrestrial gravity measurement is hard to conduct such as deserts, ice sheets, marshlands, and primeval forests. It can acquire information on the gravity field at a fast pace, with high precision, on a large scale, and with even distribution. Compared to the classical technique of terrestrial gravity measurement, it is entirely different in terms of measuring instrument, motion carrier, measuring technique, methods of data collection, as well as theory of data reduction, etc. Airborne gravity measurement has fully demonstrated the integrated application of modern technologies in the field of geodetic survey. It is of vital significance to geodesy, geophysics, oceanography, resources exploration, and space science.

Fig. 2.21 Airborne gravity measurement



An airborne gravity measurement test was first conducted in 1958. The precision of navigation was rather low and the 10 mGal accuracy in the vertical disturbing accelerations of the aircraft was difficult to maintain, so until the late 1970s the technology of airborne gravimetry had virtually been in a state of stagnation. The advent of GPS, particularly implementation of the centimeter-level kinematic differential GPS, enabled the separation of gravitational effects with a precision of a few milligals. There are two main categories in airborne gravimetry; namely, scalar gravimetry and vector gravimetry. Scalar gravimetry can only determine the acceleration due to gravity, whereas vector gravimetry can measure both gravity anomalies and deflection of the vertical. Currently, the technology of airborne vector gravimetry is still undergoing research and development and is being used in some routine operations.

Fundamentals of Airborne Gravimetry

The basic principle of airborne gravimetry is to use the airborne gravimeter on the aircraft to determine the gravitational variations of the flight profile relative to the surface reference gravity point and compute the non-gravitational accelerations and corrections to disturbance. Through filtering and data processing, the results can be obtained and then, in the downward continuation approach, the gravity value at a

surface point can be obtained. Airborne gravimetry is relative gravity measurement, i.e., prior to taking off, the aircraft is connected to a surface point of known gravity. Its basic data model is:

$$\Delta g_h = g_b + \delta g - A_v - A_E - A_h + 0.3086H - \gamma_0, \quad (2.29)$$

where Δg_h is the gravity anomaly at a point in space at a height H , g_b is the gravity value at the ground gravity reference station, δg is the gravitational variation relative to g_b observed by the airborne gravimeter, A_v is the vertical acceleration correction of the aircraft, A_E is the Eötvös correction, A_h is the inclination correction to the horizontal acceleration, γ_0 denotes the normal gravity value (referred to in Sect. 4.1) evaluated on the geometric surface of the reference ellipsoid, and $0.3086H$ is the spatial correction of normal gravity.

The vertical disturbing acceleration for aircraft A_v is mainly induced by the vertical motion of the aircraft and the self-excited vibration in the body of the aircraft. This self-excited vibration is chiefly in the high-frequency bandwidth and can be removed by means of a low-pass filtering technique and high-damping of the gravimeter's sensing element. Vertical motion of the aircraft can be corrected by determining the flight altitude in progression with an appropriate computation method. It is fairly easy to measure changes in flight altitude relative to sea level, i.e., to measure directly the changes in distance from the aircraft to the sea surface using an altimeter. However, over land surfaces, what the altimeter measures are the changes in altitude from the aircraft to the ground; therefore, in order to obtain changes in the flight altitude, measurements of changes in the topographic surface of the predetermined flight course are also needed at the same time

To our knowledge, gravity is the resultant of the universal gravitation of the Earth's masses and the centrifugal force due to the Earth's rotation. When measuring gravity on a moving platform, the centrifugal force will change due to the resultant force of the carrier's velocity and the rotation velocity of the Earth, and this change is known as the Eötvös correction (A_E). The computational formula is written as:

$$A_E = \left(1 + \frac{H}{R}\right) \left(2\omega V \sin A \cos \varphi + \frac{V^2}{R}\right), \quad (2.30)$$

where H denotes the flight altitude, R is the average radius of the Earth, V is the velocity of the carrier, A indicates the azimuth of the motion, ω is the angular velocity of the Earth's rotation, and φ is the geocentric latitude at a measuring point.

When determining gravity, the gravimeter and the level surface should be strictly parallel to each other. For airborne gravity measurement, if the platform of the gravimeter is not strictly parallel to the level surface, it will not only affect the gravitational acceleration but also exert influence on the vertical component of the horizontal acceleration. This effect is called inclination correction to the horizontal acceleration. Assuming that g is the actual gravity value, g_r is the value measured by the gravimeter, θ is the inclination between the platform surface and the level

surface, and A_e denotes horizontal acceleration, the inclination correction to the horizontal acceleration (A_h) can be expressed as:

$$A_h = g(\cos \theta - 1) + A_e \sin \theta. \quad (2.31)$$

In the above equation, when $A_e = 500$ mGal, $\theta \leq 3.4'$, A_h is less than 1 mGal. Since the horizontal precision of the gyro platform is better than $0.2'$, this correction is generally neglected and the corresponding error is less than 0.05 mGal.

System of Airborne Gravimetry

The airborne gravimetry system is the product of a combination of modern technologies like gravity sensing, satellite positioning, inertia and precision altimetry, etc. It mainly consists of five systems:

1. Gravity Sensor System. This mainly comprises the airborne gravimeter and the platform. The airborne gravimeter should have enough dynamic range and be able to provide information, like the large, short-time accelerations while the aircraft is taking off and landing, so as to facilitate the computation of corrections to the gravity disturbance.
2. Dynamic Positioning System. The major role of this system is to guarantee the optimal real-time navigation by using GPS, to provide data of initial orbit and precise location, and to compute accelerations in relation to the motion of the carriers. It is feasible to use merely pseudo-range measurement in real-time navigation. However, in order to obtain the precise flight path, it is necessary to integrate the utilization of pseudo-range measurement, phase measurement, and Doppler measurement to observe data.
3. Attitude Sensor System. The flight attitude of the aircraft is usually referred to as “pitch, roll, and yaw” and is determined by inertia measuring instruments. Because of the disadvantages that inertia measuring instruments are highly expensive, suffer from high drift, and are hard to maintain, in recent years GPS attitude measuring instruments with high precision, zero drift, and low price have come into use.
4. Altitude Sensor System. The major function of this system is to provide data on height for computing the Eötvös correction. The correction is applied to reduce the airborne gravity anomaly to the Earth’s surface using microwave altimeter, radar altimeter, pressure altimeter, or GPS survey.
5. Data Collection and Processing System. This includes airborne data collection devices and ground data processing devices. The airborne devices are used to record the input data from the subsystems of the gravity sensor, navigation positioning, attitude sensor, and altitude sensor in synchronization. Each set of recorded data should have a unified accurate time scale to facilitate computation and processing for the ground devices.

2.4.4 *Satellite Gravimetry*

The features of satellite gravimetry chiefly include the ground tracking satellite, satellite-to-satellite tracking (SST), satellite gravity gradiometry (SGG), and satellite altimetry (SA) (see, e.g., Torge and Müller 2012). SA has been described in Sect. 2.3.

Determining the Earth's Gravity Field by Means of a Ground Tracking Satellite

The Earth's gravity field can be determined by means of a ground tracking satellite, using techniques such as satellite laser ranging (SLR), Doppler orbitography and radiopositioning integrated by satellite (DORIS), and precise range and range-rate equipment (PRARE).

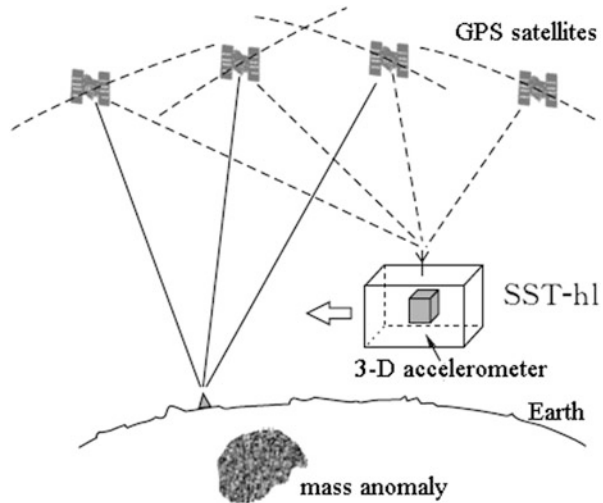
The observed quantities involved with use of a ground tracking satellite include primarily the direction, range, range rate, and phase from the ground tracking station to the satellite being tracked. The geometric and physical functional relationships between the satellite orbit and the ground tracking station can be established on the basis of these observational data. Since the satellite orbit is the implicit function of the perturbation factors of the Earth's gravity field, the gravity field of the Earth can be computed.

Determining the Earth's Gravity Field by Means of Satellite-to-Satellite Tracking

The technologies of SST can be sorted into two modes: high–low satellite-to-satellite tracking (SST-hl) and low–low satellite-to-satellite tracking (SST-ll). SST-hl utilizes the space-borne GPS receiver and the GPS satellite constellation (altitude about 21,000 km) on the low Earth orbit satellite (LEO, altitude 400 km), forms the high–low satellite space tracking network, and estimates the three-dimensional location, velocity, and acceleration of the low Earth orbit satellite, namely the first derivative of gravitational potential (GPDF). SST-ll employs two identical satellites in the same orbit with an inter-satellite distance of 200–400 km, measures precisely the relative motion of the two satellites or the changes of inter-satellite distance by using a microwave interferometer, and determines the coefficients of the Earth's gravity field based on the rate of change of the inter-satellite distance (see Nin et al. 2006).

Germany's CHAMP (Challenging Mini-Satellite Payload for Geophysical Research and Application) employs the SST-hl tracking mode, as illustrated in Fig. 2.22. CHAMP was successfully launched into an orbit of 418–470 km altitude in July, 2000 on a 5-year mission. One of the scientific missions of CHAMP was to

Fig. 2.22 Schematic diagram of SST-hl tracking mode



determine the medium- and long-wavelength static part and the temporal variations of the gravity field.

GRACE (Gravity Recovery and Climate Experiment) is a joint project of the USA and Germany. It employs the combination of two tracking modes: SST-hl and SST-II, as shown in Fig. 2.23. The GRACE satellite was successfully launched into orbit in March, 2002 on a 5-year mission. One of the scientific missions of GRACE was to determine precisely the medium- and long-wavelength static part of the gravity field and to analyze and determine the variations in the Earth's gravity field every 2–4 weeks (see Klees et al. 2008).

Determining the Gravitational Acceleration Differences in the Earth's Gravity Field by Satellite Gravity Gradiometry

Satellite gravity gradiometry (SGG) allows determination of the differences in gravitational acceleration in three mutually orthogonal directions by using the differential accelerometer on one or more fixed baselines (about 70 cm) inside the satellite. The signals observed indicate gradients of the gravitational acceleration component, i.e., the second derivative of the gravitational potential. Non-gravitational accelerations (e.g., air resistance), in the same way, exert some effects on all the accelerometers inside the satellite, and the difference can be removed perfectly when differentiating. One of the missions of SGG is to detect the Earth's gravity field and its variations with higher temporal and spatial resolution.

The European Space Agency (ESA) launched a third gravity satellite in March, 2009, the GOCE (Gravity and Ocean Circular Exploration) with SGG mode (Fig. 2.24) (see Bouman et al. 2004).

Fig. 2.23 Schematic diagram of SST-II tracking mode

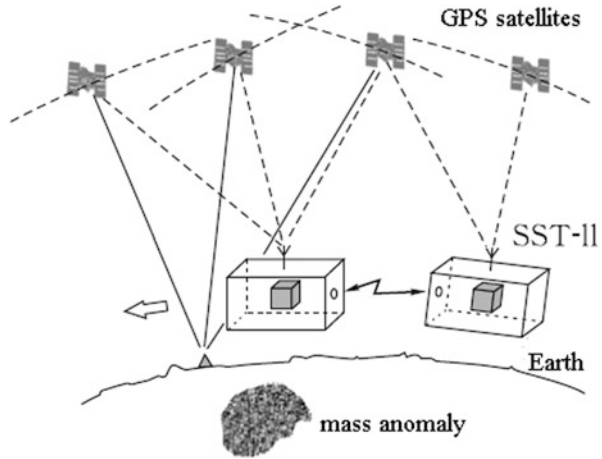
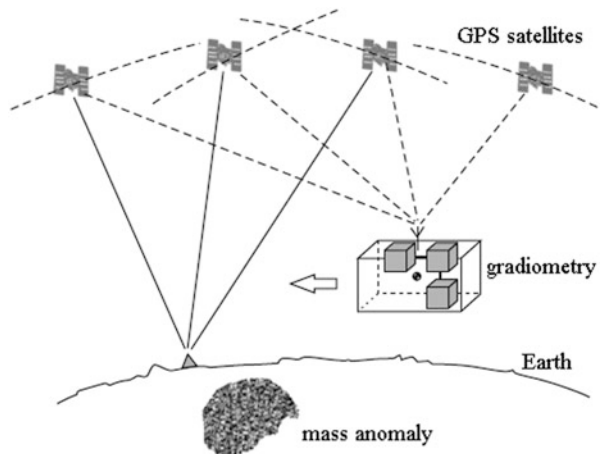


Fig. 2.24 Schematic diagram of SGG mode



Review and Study Questions

1. Briefly discuss the instruments employed and the observational quantities for different geodetic data collection techniques.
2. Briefly explain the concepts of horizontal and vertical angles and introduce the methods used for measuring horizontal angles.
3. Illustrate the concepts of astronomical longitude, astronomical latitude, and astronomical azimuth.
4. Briefly describe the basic principles of and methods for electromagnetic distance measurement (EDM).
5. Briefly discuss the methods for GPS measurements.

6. What are the basic principles for determining marine geoid height using satellite altimetry?
7. Explain the measuring principles of the VLBI technique.
8. How many categories can satellite gravimetry be classified into and what are they?
9. Explain the advantages of space geodetic techniques with respect to the classical geodetic survey.

Geodesy

Introduction to Geodetic Datum and Geodetic Systems

Lu, Z.; Qu, Y.; Qiao, S.

2014, XXI, 401 p. 147 illus., 15 illus. in color., Hardcover

ISBN: 978-3-642-41244-8