8. Geodesy

Matthias Becker

Geodesy is the basis for all Geographic Information System (GIS) applications as it provides all information that is required for describing the location of a point at or close to the Earth. In this chapter the basic definitions, quantities and mathematical relations used in Geodesy are described. It should provide the important understanding of reference frames, coordinates, height systems, their variation in time and their relation to plane coordinates. It also includes a review of the importance of the gravity field and basic methods to determine coordinates.

8.1 Basics

Geodesy is closely related to natural sciences as well as to technical sciences. It works strictly with the *Système International d'Unités* (SI) system of physical units [m–kg–s] and can to a large extent be represented by Euclidian geometry and Newtonian mechanics. Geodesy deals with the determination of the size and shape of the Earth, its gravity field, and the geometric coordinates of surface or other points related to the Earth. Since the availability of artificial Earth satellites and the advances of space geodesy, this task has been tremendously facilitated. For the first time the precise figure of the Earth, its gravity field, and well-defined global reference frames to monitor coordinates and their changes in time could be derived. Tasks that geodesists worked on for centuries are now completed within weeks. Presently the Earth is being covered with a network of thousands of continuous global positioning system (GPS) stations [8.1] which continuously record variations in coordinates at the millimeter level. GPS, as the precursor of similar global navigation satellite systems (GNSS) to come, is the main tool for positioning, in particular for Geographic Information System (GIS) applications. From the level of tens of meters in navigation down to the submillimeter level in deformation analysis, coordinates can be measured in a global, common, and homogeneous reference system.

The geometrical information obtained from GPS has to be supplemented by gravity field information [8.2] if the vertical or height component is of interest. Today's gravity field information comes from the combination of satellite mission data with high-resolution terrestrial gravity observations and has a global accuracy of a few centimeters to decimeters if used for determination of the physical reference surface for height, the geoid (commonly referred to as *mean sea level*). The combination of geoidal height and GPS height leads to the vertical coordinate that is relevant in GIS and many other surveying and mapping applications, commonly referred to as *the height*. At the present level of accuracy, time-dependent geoid changes can be monitored at the millimeter level averaged over regions of about 500 km^2. These changes may be caused by gravity changes due to mass redistribution, such as melting icecaps or glaciers or other global and local effects.

Geometrical monitoring of regional surface deformations is enabled by using remote sensing satellites equipped with Synthetic Aperture RADAR (SAR) and interferometric differential SAR. These techniques allow the derivation of digital terrain models at the submeter level and the detection of variations in terrain at the millimeter level.

The basic task of geodesy is the definition and realization of coordinate systems and their interrelations that allow the description of the continuously changing Earth. By the collection and administration of all data describing the geometrical and physical structure of the Earth surface, users can apply this information to produce GIS systems to archive, display, and utilize these data for all types of applications. For the proper use of spatial data it is essential to understand these basic principles and relationships in order to correctly assess the data quality and its uncertainty, and to allow for correct handling of geodetic data.

8.2 Concepts

Geodesy distinguishes a number of surfaces that have to be clearly distinguished (Fig. 8.1). These are the solid Earth surface (i. e., the topography), the reference ellipsoid, and equipotential surfaces such as the geoid or a local level surface. The position of a point in space can be described purely geometrically by three-dimensional Cartesian coordinates referenced, e.g., to the center of mass of the Earth.

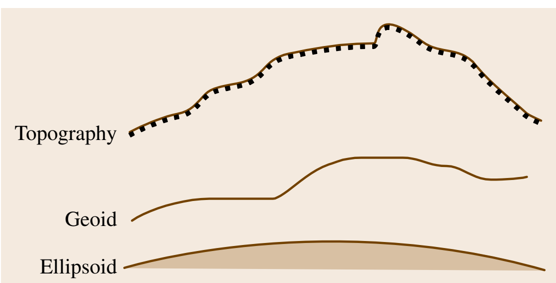

Fig. 8.1 Main geodetic surfaces

However, thinking of the spatial relation of the position of this point to features on the Earth's surface, description in terms of latitude, longitude, and ellipsoidal height is more appropriate and informative. Being still purely geometric, the introduction of a sphere, or better a reference or mean Earth ellipsoid, allows the separation of horizontal position and vertical position above the ellipsoid. A mean sphere of radius $R = 6371.0$ km may be sufficient for some applications, but the use of an ellipsoid is more appropriate and still simple enough to handle in computations. The flattening, i. e., the oblateness, of an ellipsoid fitted optimally to the Earth results in a difference of 23 km between the equatorial and the polar axis.

For an even better approximation of the Earth's figure, the concept of level surfaces plays a major role. Level surfaces are defined as being everywhere normal to the direction of the plumb line. On a level surface, therefore, water cannot flow, and any liquid will be at rest if it is part of a level surface. In partic-

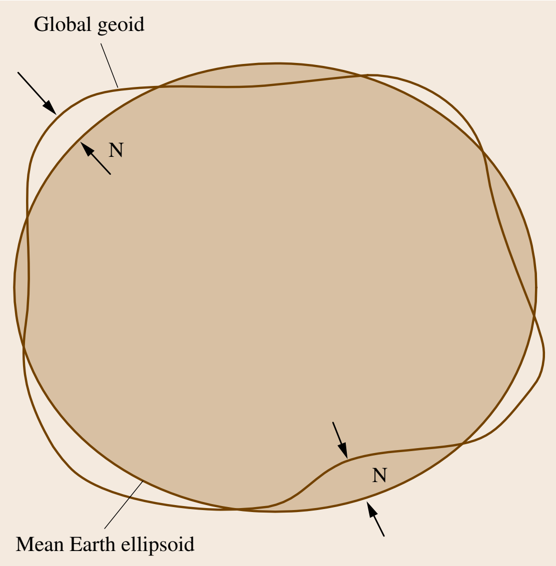

Fig. 8.2 Mean Earth ellipsoid and global geoid

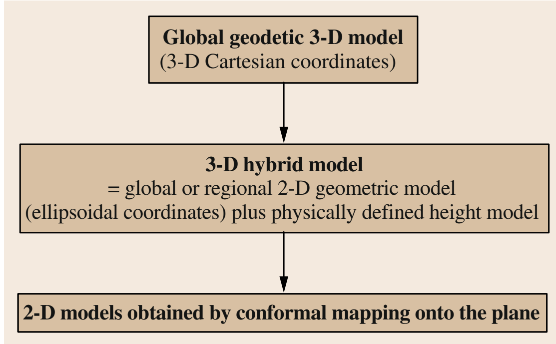

Fig. 8.3 The hierarchy of the three layers of coordinate systems

ular the undisturbed surface of the oceans, the mean sea surface, is very close to a level surface. This level surface best fitting to the mean sea surface is called geoid, which extends inside the Earth's crust and serves as a reference for height determinations. The geoid is a complicated surface that cannot be described analytically and therefore cannot be used for computations. It is described by the distance to the best-fitting ellipsoid, which deviates by up to about 100 m from the geoid (Fig. 8.2). This vertical separation of ellipsoid and geoid determining the shape of the geoid is computed globally from models of the geopotential, i.e., Earth gravity models (EGM), or locally from more precise regional and local geoidal models.

For the use of coordinates in GIS we can distinguish conceptionally between the three layers illustrated in Fig. 8.3. The first of these is global geodetic three-dimensional (3-D) Cartesian coordinates for a unique description in geometry space. The next level is the use of a global or regional two-dimensional (2-D) surface model, such as the ellipsoid, that allows the separation of the vertical coordinate and the introduction of a level surface such as the geoid to work with physically defined height models. The third layer contains 2-D models that are obtained from the conformal projection of the curvilinear coordinates onto the plane, e.g., the most commonly used universal transverse Mercator (UTM) coordinates.

8.3 Reference Systems and Reference Frames

This section describes the basic reference systems and their realization in the form of reference frames. Reference systems are maintained by the International Earth Rotation and Reference System Service [8.3] based on the actual standards as defined by the International Association of Geodesy (IAG) and the International Astronomical Union (IAU). Fortunately, due to standardization and the predominant use of GNSS or other satellite techniques, virtually all actual positionings are based on these systems, and coordinates are given in the International Terrestrial Reference Frame (ITRF) [8.4]. Current regional or national systems are mostly based on the ITRF or are in the process of being updated to it, which highlights the utmost importance of the ITRF and the need for a proper definition and understanding. The celestial system [8.5] is as important, although not directly visible to the user of positioning services. In addition we will introduce the World Geodetic System 84 (WGS 84) due to its wide use and the Geodetic Reference System 80 (GRS80) due to its importance in mapping and national surveying.

As the Earth is rotating and as satellites are revolving in an inertial space, a supreme system, the celestial reference system, must be used. It constitutes an inertial frame of reference in which any body initially at rest will remain at rest indefinitely, or in which a moving body moves in a straight line with constant speed indefinitely; in other words, it is free from any inertial forces. It can be defined as a frame of reference in which Newton's laws of motion apply exactly. Terrestrial, Earth-fixed systems are not inertial because they are revolving around the sun and rotating with the Earth, so virtual forces such as the Coriolis force and the centrifugal force have to be taken into account. Reference systems are constructed from observations by geodetic space techniques, and linking them allows the unique realization of both the celestial inertial system and the terrestrial system. In Fig. 8.4, the three pillars of geodesy are shown together with the observation techniques that are involved. Geometrical observations, observations on Earth rotation, and gravity field observations are combined to maintain the reference frame.

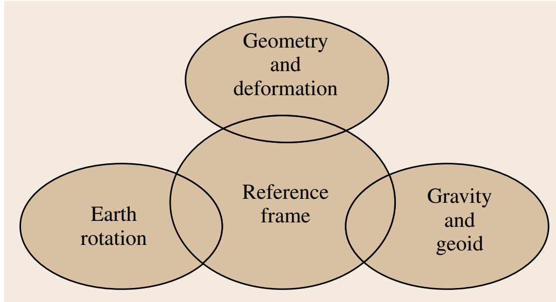

Fig. 8.4 The three pillars of geodesy linked together by the reference system

8.4 Coordinate Reference System

Coordinate reference systems (CRS) are a combination of at least one coordinate system together with its spatial datum. In a CRS, positions or locations of geographic information are described by coordinates. In GIS [8.6] the schema for the definition of a CRS contains two different elements: the datum and the coordinate system (Fig. 8.5). The datum defines how the CRS is related to the Earth: the position of the origin, the scale, and the orientation of coordinate axes, e.g., ED50 (European Datum 1950) and ETRS89 (European Terrestrial Reference System 1989). A geodetic datum in addition includes the parameters of a reference ellipsoid. A vertical datum defines the reference potential of physical heights (Sect. 8.5). The datum may also be a local engineering datum. The coordinate system describes how the coordinates are expressed in the specified datum, e.g., as Cartesian coordinates, ellipsoidal coordinates or coordinates of a map projection such as UTM. The coordinate system, as the mathematical part of the coordinate reference system (CRS), is a set of rules, e.g., projection equations, for specifying how coordinates are to be assigned to points. The list of coordinates in a specified CRS constitutes the coordinate reference frame (CRF).

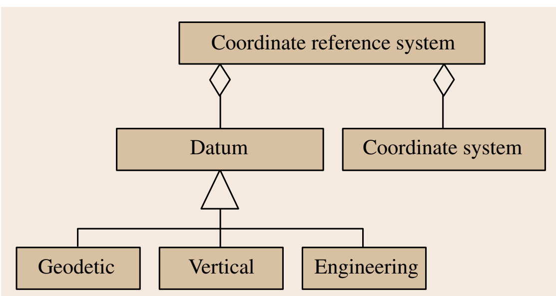

Fig. 8.5 Schema of CRS definition (after: ISO 19111:2007 *Spatial referencing by coordinates*)

8.4.1 Coordinate Systems and Coordinate Types

In order to specify a location, three coordinates and (as material points may be subject to motion) a time stamp are required. The methods and concepts that are used to fix a point in space are called the coordinate system. They are defined by conventions. The most common conventional coordinate system is the orthogonal system of Cartesian coordinates (Fig. 8.6).

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3 = \sum_{i=1}^{3} x_i\boldsymbol{e}_i \,,$$

$$\text{where} \quad e_i \cdot e_j = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} . \tag{8.1}$$

In three-dimensional geodetic applications the components for the three axes are labeled x, y, z, thus

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \tag{8.2}$$

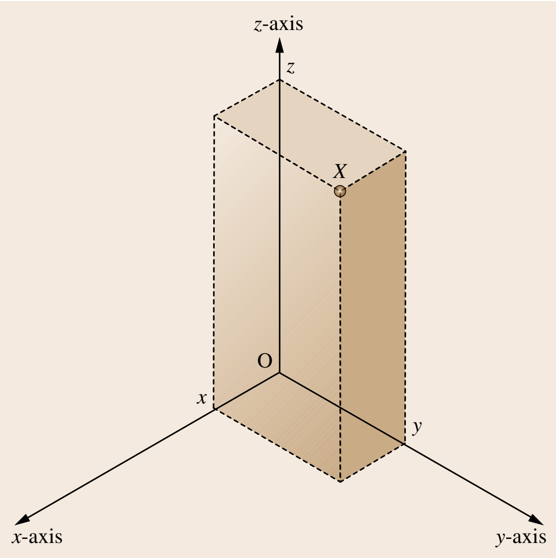

Fig. 8.6 Three-dimensional Cartesian coordinates

The other main system of orthogonal coordinates used in geodesy is that of the geodetic surface coordinates of an ellipsoid of rotation. These are the ellipsoidal coordinates geodetic latitude φ, geodetic longitude λ, and ellipsoidal height h (Fig. 8.7); φ is the angle between the equatorial (x, y)-plane and the normal to the ellipsoid, λ is the angle between the (x, z)-plane located at the Greenwich mean meridian and the normal to the ellipsoid, and h is the normal distance to the ellipsoid. The size and figure of the ellipsoid are defined by the major and the minor axes a and b. The flattening f and the first and second eccentricity e and e' as well as the polar radius of curvature c are derived

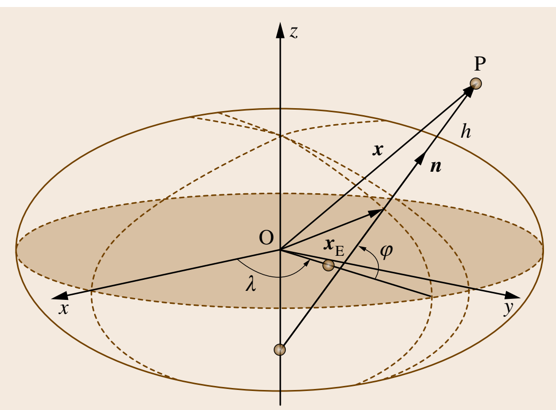

Fig. 8.7 Ellipsoidal coordinates and position vectors

constants that are used in calculations with ellipsoidal coordinates, e.g., in conversion from and to Cartesian coordinates (Sect. 8.8.1, Fig. 8.13). In general, the term *geodetic coordinates* is associated specifically with the ellipsoid, used for large- and medium-scale mapping and in geodesy. The term *geographic coordinates* is more general and is used for spherical coordinates (Fig. 8.8) that are used for small-scale mapping or approximations to the ellipsoid.

An important distinction has to be made for natural or astronomical coordinates. These describe the direction of the plumb line. At each point they give the normal direction to the associated equipotential surface, i. e., the local zenith direction. The astronomical coordinates are dependent on the irregularities of the local gravity field, but may differ from the ellipsoidal latitude and longitude by up to 30″ or more in mountainous regions.

8.4.2 International Celestial Reference System and Frame

The celestial reference system is a conventional system that constitutes an inertial space-fixed system [8.7]. Its origin is located at the barycenter (center of mass) of the solar system. For the definition of the coordinate axis some further explanation on the Earth and its rotational motion is needed. The rotation axis of the Earth, or more precisely the angular momentum axis, is used as the x_3 axis of a Cartesian system. The position of this reference axis is called the celestial ephemeris pole (CEP). As the rotation vector of the Earth oscillates for a number of reasons, a specific date (J2000.0, Sect. 8.3) has to be defined at which the direction of the rotation axis is used. The x_1 axis of the celestial system points towards the vernal equinox; it is specified by the direction of the intersection of the equatorial plane of the Earth and the ecliptic. The x_2 axis completes the orthogonal system of the International Celestial Reference System (ICRS).

The ICRS is realized by very-long-baseline interferometry (VLBI) estimates of equatorial coordinates of a set of extragalactic compact radio sources, the International Celestial Reference Frame (ICRF). By appropriate modeling of VLBI observations in the framework of general relativity, the directions of the CEP and the vernal equinox are maintained fixed relative to this selected set of precise coordinates of quasars. The catalogue of stars as used for optical astronomical observations (the *Fundamentalkatalog* FK5 or FK6) is aligned to ICRF and provides the primary realization of ICRS at optical wavelengths.

The relation between the celestial and the terrestrial frames is of importance if astronomical methods of positioning or GNSS orbit determination in the framework of navigation or positioning are required.

8.4.3 International Terrestrial Reference System and Frame

For all tasks that are related to the Earth surface or close to it an Earth-fixed system that rotates with the Earth is appropriate. This again is a conventional system that is geocentric, the center of mass being defined for the whole Earth, including oceans and atmosphere. The x_3 Cartesian axis is attached to the mean Earth rotation axis [8.8]. The position of this mean axis on the Earth surface is called the Conventional International Origin and was initially given by the Bureau International de l'Heure (BIH) orientation at 1984.0. The direction of the x_1 axis is the BIH 1984.0 mean meridian of Greenwich, and x_2 completes the orthogonal system. The latter two axes define the terrestrial plane of the equator orthogonal to the mean rotation axis.

The realization of the terrestrial conventional system is called International Terrestrial Reference System (ITRS), its realization is the International Earth Rotation and Reference Frame Service (IERS) ITRS Product Center. It is a set of points with their three-dimensional Cartesian coordinates, respective point velocities, and

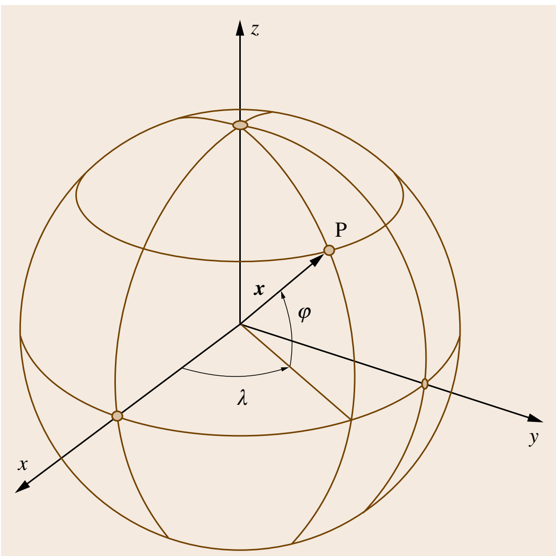

Fig. 8.8 Earth-fixed terrestrial system with geographic coordinates

a reference epoch. It is important to keep in mind that, at the present level of observational precision, the Earth and in particular the Earth's crust is constantly in motion and deforming. A purely static description may be sufficient for some mapping or GIS applications. However, in view of the reference frames, for positioning and GNSS applications, kinematic modeling of points is mandatory. Changes in coordinates are at the 1–3 cm level per year on average. Some regions, such as the Pacific and South East Asia, however, exhibit locations with velocities up to 24 cm per year. These motions can, for the majority of the solid Earth surface, be described quite well by plate motion models [8.9] as long as one stays away from the plate boundaries. Within the deformation zones at the plate boundaries, irregular and large motion rates have to be expected.

The current procedure to compute the realization is to combine the observations of space geodesy techniques: VLBI, lunar and satellite laser ranging (LLR, SLR), GPS, and Doppler orbitography and radiopositioning integrated by satellite (DORIS) in a least-squares adjustment. The combination method makes use of local ties in collocation sites where two or more geodetic systems are being operated. ITRF solutions are published in intervals that depend on the number of new observations and the expected changes in the realization expected. The numbers (yy) following the designation ITRF specify the last year whose data were used in the formation of the frame. Hence ITRF98 designates the frame of station positions and velocities constructed using all of the IERS data available until 1998. The latest issue is ITRF2005 [8.8] with more than 300 points. The coordinates, velocities, and full covariance matrix of the least-squares adjustment used for their estimation with the related uncertainties can be obtained from the ITRF Product Center [8.4]. Actual positions of points at the time of observation or any other specific epoch in time t can be computed by (8.3)

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \boldsymbol{v}(t - t_0) + \Sigma[\Delta\boldsymbol{x}_i(t)] , \tag{8.3}$$

where $\boldsymbol{x}(t)$ is the vector of coordinates at epoch t, $\boldsymbol{x}_0$ the vector of coordinates and velocities at the reference epoch t_0, $\boldsymbol{v}$ the vector of velocities, $\Delta\boldsymbol{x}_i(t)$ the site-specific time-dependent corrections (if available, e.g., in case of earthquakes or other known variations of station position in time).

The reference epoch of an ITRF solution is specified to be the central epoch of the data span that was used in the computation. Presently it is January 1, 2000. The site-specific corrections are needed if the station position is to be determined with highest accuracy, at the centimeter or subcentimeter level. In general the station positions and velocities in the latest ITRF realization are accurate to the level of a few centimeters. For national reference systems and networks or regional densifications the coordinate solutions of the respective points are transformed into the ITRF by a suitable three-dimensional transformation (Sect. 8.8.3).

It is important to understand the concept of time-varying coordinates and the consequences for reference frames and coordinate determination. All computations, e.g., in GPS positioning, should be performed with reference coordinates at the measurement epoch. Only the instantaneous coordinates correspond to the measured coordinate differences that come from actual geodetic observations. After that, the resulting coordinate sets may be transformed to any desired reference epoch as required, e.g., by some national reference system definitions. An example is the European Reference System [8.10] that is defined as a system with a static datum and a reference epoch at January 1, 1989. Therefore, after computation of new coordinates in the actual ITRF at the observation epoch, the coordinates have to be rotated back into the reference epoch by using the corresponding velocities. For Central Europe, for example, this corresponds to a coordinate change of almost 50 cm if observed in 2008. For stations that do not have their own known velocity vector from repeated observations, the revised no net rotation Northwestern University velocity model (NNR-NUVEL-1A) for plate motion [8.9] can be applied in stable regions. Other countries, e.g., New Zealand, that lie on an active fault, use a dynamic datum that is adapted to the rapid and irregular changes of coordinates [8.11]. The typical procedure for computing the coordinates of a new site in ITRF by differential GPS is

1. compute the baseline coordinate components from GPS observations at epoch t;
2. transform the known ITRF coordinates of the reference point to the actual values at the observation epoch t by using its known ITRF velocities;
3. compute the coordinates of your new station;
4. in case one needs the coordinates in a national reference frame at the respective reference epoch: Transform the resulting coordinates by using, e.g., NUVEL plate motion velocities to the reference epoch;
5. alternatively to point 4: For some national networks transformation parameters are given which enable the transformation of points in a certain region and pertaining to a particular network by using empirically determined values that keep the consistency of a particular reference frame realization. Then the back-rotation of step 4 using model velocities may be replaced by a three-dimensional seven-parameter coordinate transformation.

8.4.4 World Geodetic System (WGS 84)

The WGS 84 is a particular realization of the terrestrial reference system that is implicitly connected to the US GPS system. It is realized by the coordinates of the tracking station antennae of the GPS ground segment monitoring stations that are used to compute the broadcast ephemeris of the GPS satellites. In its predecessors and earlier versions before 1996 (WGS 72, for example) the WGS was determined exclusively from these observations. Since 1996, these tracking station coordinates have been computed from the ITRF coordinates and velocities, so WGS 84 is now identical to the ITRF [8.12]. Slight differences at the centimeter level may occur as the WGS may use an older version of the ITRF and the realization itself is less accurate than the primary International GNSS Service (IGS) stations used in ITRF. The difference does not matter in most navigation and positioning applications. So, by using the GPS broadcast ephemeris in GPS point positioning, the resulting coordinates will be in the ITRF at the observation epoch. For geodetic work of utmost precision the methods described in the previous section have to be applied.

The WGS 84 definitions in addition include an associated WGS 84 ellipsoid to convert the Cartesian x, y, z coordinates to ellipsoidal latitude, longitude, and height and a gravity field model to relate ellipsoidal height to the geoid. The ellipsoidal flattening of the WGS 84 is slightly different from that of the GRS80 (Table 8.1) ellipsoid that should be used in geodesy, however the difference for applications at or close to the Earth surface is negligible. The gravity field model that is used to compute the geoid as a global vertical reference surface is the enhanced Earth Gravitational Model 1996 (EGM96), a spherical harmonic expansion complete to degree and order 360. By use of this model, height above mean sea level can be computed as $H = h - N$ to better than 1 m anywhere on the Earth.

8.4.5 Geodetic Reference System 1980

The Geodetic Reference System 1980 (GRS80) is the official geodetic reference system recommended by the International Association of Geodesy. It should be used

Table 8.1 Parameters of the GRS80 and the WGS 84

Parameter	GRS80	WGS 84
Equatorial radius of the Earth	6 378 137 m	6 378 137.0 m
Geocentric gravitational constant of the Earth, GM (including the atmosphere)	$3\,986\,005 \times 10^8\ \mathrm{m^3/s^2}$	$(3\,986\,004.418 \pm 0.008) \times 10^8\ \mathrm{m^3/s^2}$
Reciprocal flattening of the reference ellipsoid	298.257222101	298.257223563
Angular velocity of the Earth, ω	$7\,292\,115 \times 10^{-11}$ rad/s	$7\,292\,115 \times 10^{-11}$ rad/s

in all geodetic work and in computations of the gravity field both on the Earth's surface and in outer space. The definition is given in [8.13]. Four defining parameters are used to uniquely define the best-fitting ellipsoid and the reference gravity field (normal gravity and potential) of the Earth based on the theory of the geocentric equipotential ellipsoid. The defining conventional constants for GRS80 are given in Table 8.1.

Its origin and orientation are such that the minor axis of the reference ellipsoid, defined above, is parallel to the direction defined by the Conventional International Origin, and such that the primary meridian is parallel to the zero meridian of the BIH-adopted longitudes, which is coincident with the x, y, z Cartesian coordinate system constituted by the ITRF. The GRS80 should eventually replace all nongeocentric regional reference systems and geodetic datums. Together with the ITRF, it provides the basis for the computation of globally homogeneous coordinates, mappings, and reference gravity field parameters.

8.5 Height Systems and Vertical Datum

8.5.1 Definition of Heights in Geodesy

The term *height* has an intrinsic problem in geodesy as its definition can be purely geometric or based on physics, i. e., the potential of the Earth. There is no "best" height, as the use of a particular height definition depends on the application. Ellipsoidal heights are defined in a purely geometric way and are not suited for technical purposes. They are defined as the distance from a point to the chosen reference ellipsoid along the ellipsoidal normal.

Heights based on the potential primarily are not measured in meters but in potential differences. The difference in potential $W_\mathrm{B} - W_\mathrm{A}$ of two points is the work done in transporting a unit mass of 1 kg from A to B. This work is independent of the path taken from A to B. Potential differences are measured by a combination of geometric leveling and gravity measurements (=geopotential leveling). The physical dimension of a potential difference is $\mathrm{m}^2\,\mathrm{s}^{-2}$. As users in general prefer a geometric value in meters, a suitable definition of a metric height system has to be agreed on. It should fulfill three requirements [8.14]

- the height of a point should be unambiguous and independent of the way it was measured;
- the height should ideally be free of any hypothesis;
- corrections to measured height differences from geometric leveling to obtain the system adopted should be small so that their possible neglect does not cause too large effects.

As seen in Fig. 8.9, equipotential surfaces are not parallel, i. e., points with the same value of the potential do not have the same ellipsoidal height and their ellipsoidal height varies with position and the height itself. In order to eliminate these variations that cause path dependency of geometric leveling, gravity values along

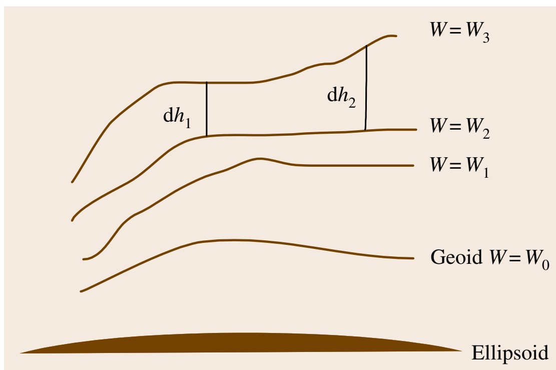

Fig. 8.9 Ellipsoid, geoid, and nonparallel equipotential surfaces W_i

the leveling lines have to be known. For a potential increment of $\mathrm{d}W$, gravity of g, height increment of $\mathrm{d}h$, and leveled height difference of δh, the basic relation to compute potential difference is: $\mathrm{d}W = -g\,\mathrm{d}h$. Integration along the leveling line gives the potential difference $W_\mathrm{A} - W_\mathrm{B}$ between a starting point A and an end point B as

$$W_\mathrm{A} - W_\mathrm{B} = \int_\mathrm{A}^\mathrm{B} g\,\mathrm{d}h \approx \sum_\mathrm{A}^\mathrm{B} g\delta h\;. \tag{8.4}$$

For the definition of a height system, a zero level has to be defined. This is accomplished by the geoid, the potential of which is labeled as W_0 in the corresponding relations. The potential difference to this value W_0 is called the geopotential number (C_A):

$$C_\mathrm{A} = W_0 - W_\mathrm{A} = \int_{\mathrm{P}_0}^\mathrm{A} g\,\mathrm{d}h \approx \sum_{\mathrm{P}_0}^\mathrm{A} g\delta h\;. \tag{8.5}$$

The geopotential number can formally be converted to a metric quantity by dividing it by a gravity value. The value used in Europe is the gravity value of the reference ellipsoid at 45° latitude, called normal gravity γ_{45}. The result is called dynamic height

$$H_\mathrm{A} = \frac{C_\mathrm{A}}{\gamma_{45}}\;. \tag{8.6}$$

Dynamic heights are strict, as no water can flow between two points with identical heights and there are no hypotheses involved. Their main drawback is the fact that they require large corrections to the leveled height differences and that they cannot be combined with GPS heights as they do not have a defined zero level and no geometrical interpretation.

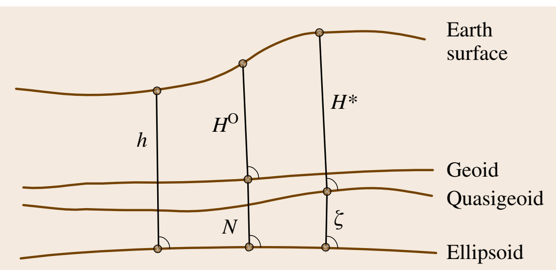

Fig. 8.10 Height types: ellipsoidal height h, orthometric height H^O, normal height H^*, and the related reference surfaces with their respective distance to the ellipsoid, geoid height N, and height anomaly ζ

8.5.2 Orthometric Height H^0

The orthometric height, H^O, is the length of the slightly curved plumb line from the geoid to the Earth surface (Fig. 8.10). The orthometric height usually reflects local variations in gravity as well as changes in topography. A fictitious geopotential leveling along the plumb line from the geoid to A will give the geopotential number C_A. Because C_A is independent of the leveling path, the same value results from a geopotential leveling along the Earth's surface. The following relation between H_A^O and C_A holds:

$$C_\mathrm{A} = W_0 - W_\mathrm{A} = \int_{\mathrm{A}_0}^\mathrm{A} g\,\mathrm{d}h = h_\mathrm{A}\frac{1}{h_\mathrm{A}}\int_{\mathrm{A}_0}^\mathrm{A} g\,\mathrm{d}h$$
$$= \bar{g}_\mathrm{A} H_\mathrm{A}^\mathrm{O}\;, \tag{8.7}$$
$$H_\mathrm{A}^\mathrm{O} = \frac{C_\mathrm{A}}{\bar{g}_\mathrm{A}^*}\;, \tag{8.8}$$

where $\bar{g}_\mathrm{A}^*$ is the integral mean of the gravity along the plumb line, which has to be computed from gravity values measured at the Earth's surface. This is where hypothetical assumptions on the density enter. As a consequence it is hardly possible to get orthometric heights with millimeter accuracy, and in mountainous areas even the centimeters may be uncertain. The requirements for a good height system given above are fulfilled for the first item, to a large extent for the third one, but not for the second one. One drawback is that water can flow between two points with equal H^O. However, the most important advantage of H^O values is that they can be combined with ellipsoidal heights: The difference of ellipsoidal heights equals the difference of orthometric heights plus the difference of geoidal undulations:

$$(h_\mathrm{B} - h_\mathrm{A}) = (H_\mathrm{B} - H_\mathrm{A}) + (N_\mathrm{B} - N_\mathrm{A})\;. \tag{8.9}$$

8.5.3 Normal Heights (H^*)

Normal heights are the most advanced concept. They are related to the geodetic theory of the Russian geodesist Molodenskij. According to his idea, the surface of the Earth is mapped point by point onto another surface. Each point A on the surface of the Earth receives a partner point Q on the same ellipsoid normal, above or below A. The ellipsoidal height of Q depends on values that can be calculated without hypotheses, which is what makes Molodenskij's theory so attractive.

The normal height H^* of point A is defined as the ellipsoidal height of the partner point Q. The calculation of the height of Q depends on potential differences. The basics are formulated according to

$$U_0 - U_Q = W_0 - W_A = C_A \,, \tag{8.10}$$

where U_0, the mean Earth ellipsoid, generates an associated theoretical gravity field. Its potential, labeled U_0, is by definition equal to the potential W_0 of the geoid. C_A, the geopotential number, can be determined from (8.5).

Thus, U_Q is computable without hypotheses as

$$U_Q = U_0 - C_A = W_A - W_0 + U_0 \,.$$

The entirety of all points Q creates a surface that is near but not identical to the Earth's surface. This surface is called the *telluroid*. The distance between the Earth's surface and the telluroid ζ is termed the *height anomaly*.

Similar to the orthometric heights (8.7), the following relation holds for normal heights:

$$C_A = U_0 - U_Q = \int_0^{H^*} \gamma \, \mathrm{d}H_A^* = \bar{\gamma}_Q H_A^* \,.$$

This leads to the normal height of point A, H_A:

$$H_A^* = \frac{C_A}{\bar{\gamma}_Q} \,, \tag{8.11}$$

where $\bar{\gamma}_Q$ is the integral mean of the theoretical gravity from the ellipsoid to point Q. It can be computed as shown from the normal gravity formula once the iterative determination of point Q is accomplished.

The quasigeoid is another surface commonly used in geodetic science. The distance between a point A on the Earth's surface and the quasigeoid is exactly the normal height of the point H_A. The pair of geoid height N and orthometric height is equivalent to the pair of quasigeoid height and normal height.

The geometrical interpretation can be derived from Fig. 8.10 as

$$\begin{aligned}&\left(H_B^O - H_A^O\right) + (N_B - N_A)\\ &= \left(H_B^* - H_A^*\right) + (\zeta_B - \zeta_A) \,.\end{aligned} \tag{8.12}$$

The quasigeoid is not an equipotential surface. It coincides by definition with the geoid on the oceans. On the continents it runs slightly above it. The difference between N and ζ depends on the geology and the topography itself; in mountainous regions their difference may amount to 40–50 cm.

The importance of orthometric and normal heights lies in the fact that, as soon as the detailed geoid (or quasigeoid) is known, costly leveling operations can be replaced by GPS observations. The computation of the necessary height anomalies as well as of geoid undulations is the task of physical geodesy and is a very ambitious problem. Regional models allow the computation of the geoid or quasigeoid height to centimeter precision in a limited area. GPS leveling, the determination of physical heights by combining ellipsoidal heights from GPS and geoid or quasigeoid heights, is therefore about to replace traditional leveling on the national and regional scale. Global geopotential models, however, are not yet sufficiently accurate to compute the absolute N or ζ to better than decimeters (Sect. 8.6).

The comparison of the values of the four types of heights in use is given in Table 8.2, to illustrate the numerical values. They are taken from points of the Austrian first-order leveling net ranging from the lowlands to the Alps.

The choice of a particular height system depends on each country, however, for Europe the use of normal heights is recommended by the EUREF Subcommission of the IAG for Europe [8.16]. Based on an adjustment of geopotential numbers, normal heights are recommended as a standard. Future height systems for precise applications in monitoring and geodynamics will have to take height variations with time into consideration.

Vertical Datum

The zero surface to which elevations or heights are referred to is called a vertical datum. Traditionally it is associated with mean sea level. The mean sea level (MSL) reference is realized by continuous measure-

Table 8.2 Comparison of different height types (after [8.15])

Geopotential number C (kGal m)	Dynamic height (m)	Orthometric height (m)	Normal height (m)	$(\zeta - N)$ (m)	Geoidal undulation N (m)	Ellipsoidal height (m)
140.0704	142.839	142.801	142.800	+0.001	+0.60	143.40
694.0876	707.805	707.810	707.721	+0.089	+1.18	708.99
1090.1256	1111.670	1111.797	1111.645	+0.152	+2.35	1114.15
1420.8730	1448.954	1449.340	1449.037	+0.303	+2.77	1452.11

ments at tide gage stations. The average reading over a sufficiently long interval defines MSL. MSL is then used as zero elevation for a local or regional area. MSL is an approximation to the geoid. Differences arise locally because the sea level includes position- and time-dependent components due to currents, winds, tides, and salinity, among others, that cause a deviation from an equipotential surface of the Earth gravity field. This sea surface topography part is typically in the decimeter range; globally it is between -1.5 and $+1.5$ m. Therefore national height systems referred to different tide gages may have offsets at the meter level.

For georeferencing and global monitoring of the Earth the implementation of a unified global vertical datum is needed. This will lead to greater accuracy in the connection of national and continental datums. It will also improve geoid computations, as it will remove systematic regional biases in gravity anomaly databases to refer gravity anomalies to one unique geopotential surface. However, due to a whole range of fundamental geodetic questions such as the choice of the reference potential of the mean Earth ellipsoid, the value of the potential W_0 for the geoid and its time dependency, plus many more, there is no clear definition at the moment. Presently the EGM gravity models are the de facto standard if used with ITRF-derived ellipsoidal heights from GPS to compute H^{O} or H^*. Both height systems, i.e., derived from tide gages and leveling and from GPS leveling and geopotential models, are in use.

Height datums and height values, like all quantities that stem from observations, can be inconsistent for several reasons. In principle the geometric heights h may change by δh due to changes in the reference ellipsoid. The physical height H may change by δH due to changes in gravity, leveling or reference potential at the tide gage or selected zero point. Geoid heights N may change by δN due to changes in the reference potential W_0 or geoid redefinitions. These effects generally appear as a near-constant bias in a given area, which may be expressed by

$$h+\delta h=(H+\delta H)+(N+\delta N)\,. \tag{8.13}$$

The necessity to combine geometric coordinates and a potential-based height led to the introduction of compound coordinate reference system descriptions. The compound coordinate reference system describes the position by two independent coordinate reference systems, e.g., an ITRF-based geometry component and a height system with a particular vertical datum for the vertical.

8.6 Geopotential Models and Geoid

This section gives a brief description of geopotential and geoid models in the framework of georeferencing and GIS applications. They are used to compute physical heights above the geoid in combination with GPS leveling, the realization and unification of the vertical reference system, and the transformation of local geodetic observations to a global reference frame. They have to be consistent with the geometric terrestrial reference system, ITRF. In practice, global representation of the geopotential of the Earth by an expansion in spherical harmonics [8.14] is used. Until 2008, the IAG-recommended global gravity field was the Earth Gravity Model 96 (EGM96), computed by the US National Geospatial-Intelligence Agency NGA (formerly known as the Defense Mapping Agency (DMA) or National Imagery and Mapping Agency (NIMA)) in cooperation with the Ohio State University [8.17]. This data is a set of fully normalized, Earth gravity (geopotential) coefficients, complete to degree (n) and order (m) 360, corresponding to a resolution of the gravity field's features at 100 km scale. These spherical harmonic coefficients are used in Clenshaw summation numerical algorithms or fast Fourier transformation algorithms to compute all quantities of interest to geodesy. These are point gravity anomalies, point geoid heights, point N–S or E–W components of the deflection of the vertical, point total deflection of the vertical, point radial component of the gravity disturbance vector, and point N–S or E–W component of the gravity disturbance vector. Detailed description of background and application formulas are published in [8.12]. In 2009, EGM96 was replaced by the EGM2008 [8.18]. This set of coefficients is based on a combination of terrestrial gravity observations and data from satellite gravity observations. These are satellite laser-ranging missions, satellite altimetry missions over the oceans, and recently the Challenging Minisatellite Payload (CHAMP), the Gravity Recovery and Climate Experiment (GRACE), and the Gravity Field and Steady-State Ocean Circulation (GOCE) dedicated satellite gravity missions [8.19]. The latter missions have led to significant improvements in the modeling of long-wavelength gravity signals.

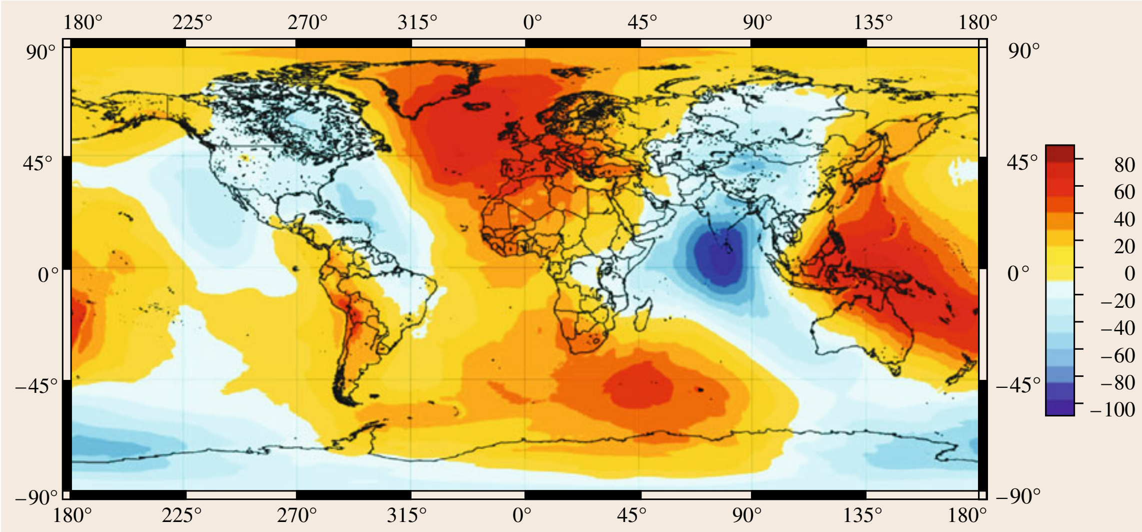

Fig. 8.11 EGM2008 geoid heights (m)

In combination with terrestrial gravity data of good quality (± 1 mGal, $1\,\text{mGal} = 10^{-5}\,\text{m/s}^2$) and coverage, significantly improved continental-scale geoid and quasigeoid models are provided. Accuracy and resolution in EGM2008 are significantly enhanced (Fig. 8.11), the spatial resolution now being about 15 km, or degree and order 2159 in the frequency domain. The accuracy of geoid or quasigeoid height computation is improved from about 0.3 m with EGM96 to 0.13 m, as shown by comparison with selected GPS leveling results. The accuracy, however, may vary in different regions of the world depending on the terrestrial input data available for the EGM2008 computation and reach 0.5–1 m in extreme cases. The standard deviation of the global geoid undulations with respect to the WGS 84 ellipsoid is about 30.5 m, with minima and maxima of −107 m and 85 m, respectively. These values illustrate the deviation of the geoid from the ellipsoid and the error that may arise if ellipsoidal heights and physical heights above mean sea level are not distinguished.

For GIS applications, the gridded data set of EGM2008 and a particular version for the use with WGS 84 are available from the NGA website, as for previous models [8.20]. They have a resolution of 1, 2.5, and 5 arcmin. Software for synthesis of harmonic coefficients or interpolation of the grids to particular positions on the Earth is provided by NGA as well.

8.7 Time Systems

Time systems play a fundamental role when dealing with space geodesy and advanced GNSS data analysis. The typical example is the transformation between the space-fixed ICRS and the Earth-fixed ITRF implicitly included in GNSS positioning. Sensor systems and sensor fusion depends on precise timing and time synchronization. Time is the fundamental quantity that, moreover, is the basis of almost all modern geodetic observation techniques used in geodesy and GIS [8.21]. It is measurable at the level of one part in 10^{15}. Recent developments in optical clocks report two orders of magnitude improvement in accuracy that will have a huge impact on GNSS and positioning in the next 5–10 years [8.22]. This section gives a brief introduction to the time scales used in GIS and geodesy. Basically, we have to differentiate between the unit of time, e.g., the interval of 1 s, and the epoch, i. e., a particular instance of an event in time. Each time system may have its own unit and its own zero epoch.

There are four basic time systems that are in use.

1. Solar time is based on the daily path of a fictitious sun that moves with constant velocity along the equator. It is the basis of universal time (UT).

One second is 1/86 400 of a solar day. Because of the changes in rotational speed of the Earth and the slowing down of Earth rotation, it is not constant.

2. Sidereal time is based on the rotational speed of the Earth. The unit is the period of the Earth's rotation with respect to a point nearly fixed with respect to the stars. One sidereal day is 4 min shorter than the mean solar day, due to the revolution of the Earth around the sun.
3. Atomic time (TAI) has the fundamental interval of one Système International (SI) second. It is defined as the duration of 9 192 631 770 cycles of radiation corresponding to the transition between two hyperfine levels of the ground state of cesium 133 (^{133}Cs). The SI day is defined as 86 400 s, and the Julian century as 36 525 days. TAI is the International Atomic Time scale, a statistical timescale based on a large number of atomic clocks. The origin was established on January 1, 1958. At midnight on January 1, 1958, universal time and sidereal times effectively ceased to function as time systems.
4. Dynamic time is based on the equations of motions of the solar system celestial bodies. The theory of general relativity implies that we have to consider the choice of an adequate inertial reference frame. For events at or close to the Earth it is suitable to use terrestrial dynamical time [*temps dynamique terrestre* (TDT)], which represents a uniform time scale for motion in the Earth's gravity field. By definition it has the same rate as an atomic clock on Earth. TAI is related to the definition of TDT by the definition

$$\mathrm{TDT} = \mathrm{TAI} + 32\,184\,\mathrm{s}\,. \tag{8.14}$$

8.7.1 Time Scales and GNSS Times

TAI is a continuous time scale, and so does not remain synchronized with the mean solar day (UT1), since the Earth's rotation rate is slowing by an average of about 1 s per year. This problem is taken care of by defining universal time coordinated (UTC), which runs at the same rate as TAI but is incremented by leap seconds periodically. Leap seconds are introduced by the IERS so that UTC does not vary from UT1 by more than 0.9 s. Presently

$$\mathrm{UTC} - \mathrm{TAI} = -34\,. \tag{8.15}$$

GPS time is derived from TAI. The time signals broadcast by the GPS satellites are synchronized with the atomic clock at the GPS Master Control Station in Colorado, USA. Global positioning system time GPST zero was set to 0 h UTC on January 6, 1980. It is not incremented by UTC leap seconds. Therefore, there is an integer-second offset of 19 s between GPST and TAI such that

$$\mathrm{GPST} + 19\,\mathrm{s} = \mathrm{TAI}\,. \tag{8.16}$$

As of 2009 there has been a total of 15 leap seconds since January 6, 1980 so that currently

$$\mathrm{GPST} + \mathrm{UTC} = 15\,\mathrm{s}\,. \tag{8.17}$$

For precise applications this offset between UTC and GPST has to be adequately considered by specifying the time system used. GPS time is primarily counted in GPS week numbers and seconds of week. Since January 6, 1980 each week has been designated its own number. For example February 4, 2009 is the day of year 35 in GPS week 1517. To identify a given epoch within the week, the concept of seconds of week is used. This number counts from midnight between Saturday and Sunday, the beginning of the GPS week. Furthermore, for convenience the individual days of the week are numbered: Sunday 1, Monday 2, Tuesday 3, Wednesday 4, Thursday 5, Friday 6, and Saturday 7. Professional GPS software uses the day of week and seconds of day for numerical reasons.

Other GNSS, such as the Russian Global Navigation Satellite System (GLONASS), or the future European Galileo system, will maintain their own time system. However, like GPS, they will be realizations of UTC and steered to be within 1 μs of UTC, modulo whole seconds. GNSS times are not adjusted for leap seconds. Their offsets will be broadcast to users to allow interoperability and seamless use of all GNSS.

A continuous time count often used in astronomy, geodesy, and GIS is the Julian date (JD). It describes a number of days and the fraction of a day after a zero epoch sufficiently in the past to precede the historical record, chosen to be at 12 h UT on January 1, 4713 BC. The JD of the standard epoch of UT is called J2000.0, where

$$\begin{aligned} \mathrm{J2000.0} &= \mathrm{JD}2\,451\,545:0 = 2000 \text{ January } 1.5^{\mathrm{d}}\mathrm{UT} \\ &= \text{January 1st, } 12\,\mathrm{h}\ \mathrm{UT}\,. \end{aligned} \tag{8.18}$$

JD is a large number, so often it is replaced by the modified Julian date (MJD)

$$\mathrm{MJD} = \mathrm{JD} - 2\,400\,000.5\,. \tag{8.19}$$

Hence J2000.0 = MJD51 544.5. MJD, in contrary to JD, starts at midnight.

8.8 Conversions, Transformations, and Projections

There are two basic kinds of coordinate operations: coordinate transformation and coordinate conversion.

Transformations are basic operations in geodesy. They cover coordinate transformations between different types of coordinate systems (Fig. 8.12), or linear, affine, and projective transformations. Typically they comprise translation, rotation, and a change in scale. Formulas are available for Cartesian or for ellipsoidal coordinates; the relevant set of formulas are known as the Helmert transformation and Molodenskij formulas, respectively. The latter also include terms to consider the change in ellipsoid parameters, i. e., the dimension of the ellipsoid. They were used in geodetic datum transformations, e.g., to relate nongeocentric geodetic systems with ellipsoidal coordinates to the ITRF and the GRS80. Today's GNSS-based coordinates are Cartesian and are given in the ITRF or WGS 84; therefore, the Helmert transformation is appropriate in most cases. The transformation parameters in general are derived empirically in a least-squares estimation by a set of identical points known in both systems. Choice, allocation, number, and the quality of coordinates of these points extensively affect the results and the accuracy.

For three-dimensional CRS in general the seven-parameter Helmert transformation is used for coordinate transformations (8.36–8.38). For two-dimensional and for geotopographical data, also a grid-based transformation is usable. Values of ellipsoidal coordinates of the identical points are computed first and stored in a regular grid. The shifts for the transformation of other or new coordinates are then computed by bilinear interpolation inside the grid meshes.

The change from one coordinate system to another based on the same datum is accomplished by a coordinate conversion. In this case, mathematical rules are specified. Generally these conversions are unambiguous and can be realized with high accuracy. Examples are map projections and conversions between Cartesian and ellipsoidal coordinates.

The change of coordinates from one CRS to another may result from a series of operations consisting of one or several transformations and conversions by concatenated operations.

8.8.1 Conversion Between Ellipsoidal and Cartesian Coordinates

Conversion between Cartesian coordinates, e.g., obtained from the ITRF, and ellipsoidal coordinates used to describe the location of a point on or above the ellipsoid is often required in positioning and GIS. An ellipse, and an ellipsoid, is defined by two parameters: the semimajor axis a and the semiminor axis b. The equation of an ellipsoid of revolution in a Cartesian system with origin in the center and z-axis in the minor axis is given by

$$\frac{x^2+y^2}{a^2}+\frac{z^2}{b^2}=1\,. \tag{8.20}$$

A number of frequently used quantities that describe the geometry of an ellipsoid of rotation can be derived from the examination of a meridian curve of the ellipsoid (Fig. 8.13): the flattening (f), the first eccentricity (e), the second eccentricity (e'), and the radius of polar curvature (c), defined as

$$f=\frac{a-b}{a}\,;\quad e^2=\frac{a^2-b^2}{a^2}\,;$$
$$e'^2=\frac{a^2-b^2}{b^2}\,;\quad c=\frac{a^2}{b}\,. \tag{8.21}$$

With the aid of auxiliary quantities

$$V=\sqrt{1+e'^2\cos^2\varphi}\,,$$
$$N=\frac{c}{V}\,,$$
$$M=\frac{c}{V^3}\,. \tag{8.22}$$

where M and N are the radii of curvature of the ellipsoid in the direction of the meridian and orthogonal to it, the transformation of the ellipsoidal coordinates to

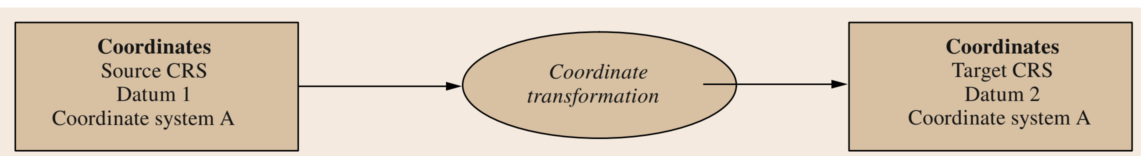

Fig. 8.12 Schema of coordinate transformation (after [8.6])

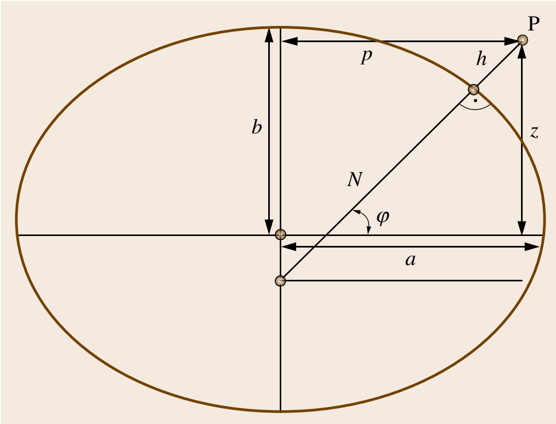

Fig. 8.13 Geometry of the ellipsoid

the Cartesian system is given by

$$x = \left(\frac{c}{V} + H\right) \cos\varphi \cos\lambda$$

$$y = \left(\frac{c}{V} + H\right) \cos\varphi \sin\lambda$$

$$z = \left[\frac{c}{V}(1-e^2) + h\right] \sin\varphi$$

or

$$x = (N+h)\cos\varphi\cos\lambda$$

$$y = (N+h)\cos\varphi\sin\lambda$$

$$z = [N(1-e^2)+h]\sin\varphi\,. \tag{8.23}$$

The inverse transformation is not as straightforward; λ follows from

$$\lambda = \arctan\left(\frac{y}{x}\right)\,. \tag{8.24}$$

Solving (8.23) for φ and h is theoretically possible but very complicated. In practice this is done by an iterative procedure starting with $h = 0$, or by an approximate solution which nevertheless is rather accurate. The auxiliary terms θ and p are defined by

$$\theta = \arctan\left(\frac{az}{bp}\right)\,, \qquad p = \sqrt{x^2+y^2}\,. \tag{8.25}$$

Latitude φ and height h result from

$$\varphi = \arctan\left(\frac{z + e'^2 b \sin^3\theta}{p - e^2 a \cos^3\theta}\right)\,, \tag{8.26}$$

$$h = \frac{p}{\cos\varphi} - \frac{c}{V}\,. \tag{8.27}$$

Care has to be taken to keep millimeter accuracy; latitude and longitude have to be given to $0.0001''$ or to 3×10^{-8} accuracy. Two normals of the ellipsoid subtending the small angle of $1''$ intersect the ellipsoid at two points 30 m apart.

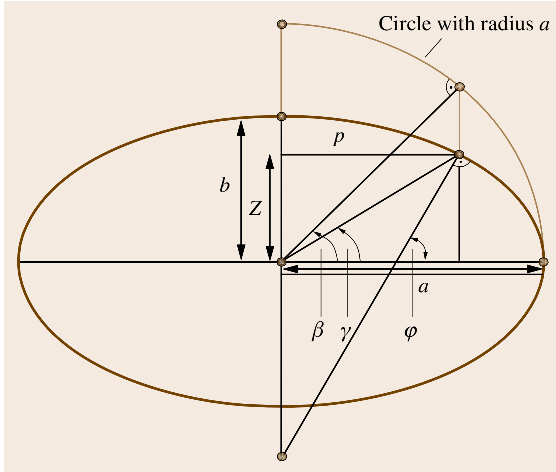

Fig. 8.14 Geodetic latitude φ, reduced latitude β, and geocentric latitude γ

For many computations in a limited area the ellipsoid can be substituted by an osculating sphere which is situated tangent to the ellipsoid in a central point of the region. The radius R of this sphere is equal to the geometric mean of the principal radii M and N of the ellipsoid M and N

$$R = \frac{c}{V^2}\,; \tag{8.28}$$

V has to be calculated for the latitude of the tangent point.

Two different types of latitude at the ellipsoid have to be distinguished: the reduced latitude β and the geocentric latitude γ. They are used in connection with the gravity models of the geopotential (Sect. 8.6). As seen in Fig. 8.14, the reduced latitude β is computed by the circle with radius a and the vertical through the point at the ellipsoid from the components p and a; the geocentric latitude γ is computed from z and p:

$$\beta = \arccos\left(\frac{p}{a}\right)\,,$$

$$\gamma = \arctan\left(\frac{z}{p}\right)\,. \tag{8.29}$$

8.8.2 Local Geodetic Systems

For handling terrestrial data, the local observation systems as shown in Fig. 8.15 have to be related to the global coordinates. The origin of a local system is topocentric, i. e., at some point P on the surface of the

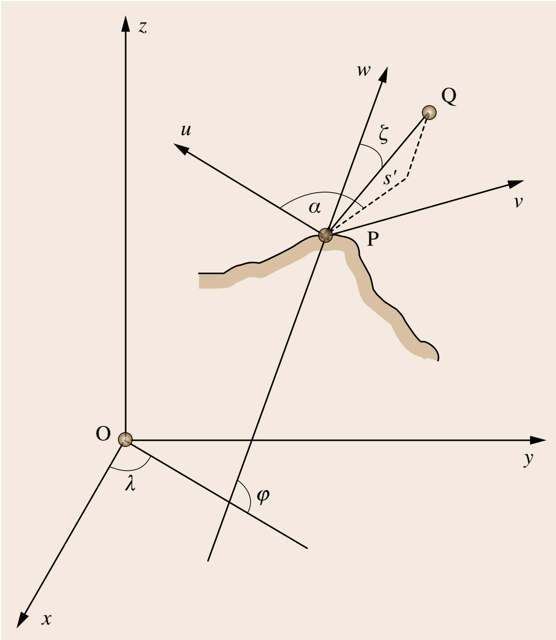

Fig. 8.15 Local observations in the topocentric system and their relation to the global CRS

Earth with the orthogonal axes u in direction of geodetic north, v to the east, and the w-axis parallel to the local normal on the ellipsoid. This system is also termed the *horizon system*. A local observation to a fixed point Q can be done by observing the distance s, the azimuth α, and the zenith distance ζ. Note that actual observations have to be corrected for the deflection of the vertical, i. e., the difference between natural coordinates and ellipsoidal coordinates. The vector $\boldsymbol{s}$ can be computed as

$$\boldsymbol{s} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \cos\alpha \sin\zeta \\ \sin\alpha \sin\zeta \\ \cos\zeta \end{pmatrix} . \quad (8.30)$$

Now the vector $\boldsymbol{s}$ has to be transformed into the difference vector $\boldsymbol{S}$ of the position vectors of points P and Q in the global CRS and its associated ellipsoid. As shown in Fig. 8.15, the two Cartesian systems have different orientations, so the conversion will include two rotations for latitude and longitude and one mirroring to convert the left-handed local system into the right-handed geocentric system. The transformation then reads

$$\begin{aligned} \boldsymbol{S} &= (\Delta x, \Delta y, \Delta z)^{\mathrm{T}} \\ &= \boldsymbol{R}_{\mathrm{w}}(180° - \lambda)\boldsymbol{R}_{\mathrm{v}}(90° - \varphi)\boldsymbol{R}_{\mathrm{v}} \cdot \boldsymbol{s} = \boldsymbol{R} \cdot \boldsymbol{s} , \end{aligned} \quad (8.31)$$

with

$$\boldsymbol{R} = \begin{pmatrix} -\sin\varphi\cos\lambda & -\sin\lambda & \cos\varphi\cos\lambda \\ -\sin\varphi\sin\lambda & \cos\lambda & \cos\varphi\sin\lambda \\ \cos\varphi & 0 & \sin\varphi \end{pmatrix} . \quad (8.32)$$

For the inverse transformation we get, with the difference vector $\boldsymbol{S} = (\Delta x, \Delta y, \Delta z)^{\mathrm{T}}$,

$$s = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} , \quad (8.33)$$

$$\begin{aligned} \alpha = \arctan[&(-\sin\lambda \cdot \Delta x + \cos\lambda \cdot \Delta y) \\ &/(-\sin\varphi\cos\lambda \cdot \Delta x - \sin\varphi\sin\lambda \cdot \Delta y \\ &+ \cos\varphi \cdot \Delta z)] , \end{aligned} \quad (8.34)$$

$$\begin{aligned} \zeta = \arccos\Bigg[&\frac{1}{s}(\cos\varphi\cos\lambda \cdot \Delta x + \cos\varphi\sin\lambda \cdot \Delta y \\ &+ \sin\varphi \cdot \Delta z)\Bigg] . \end{aligned} \quad (8.35)$$

8.8.3 Coordinate Transformation and Transformation of Terrestrial Frames

Today's modern reference frames based on satellite geodesy are orthogonal and homogeneous. The standard relation for transformation between two reference systems is a Euclidian similarity transformation with seven parameters: three translation components, one scale factor, and three rotation angles. At the level of accuracy of ITRF2005, the transformation parameters may partly be time dependent, due to different definitions used for the initial adjustment of a particular frame. Therefore also the transformation of velocities has to be considered by using the time derivative of the seven parameters. At the ITRF website the detailed transformation formulas and the most recent estimates for the transformation parameters are available [8.4].

The transformation of coordinate vector $\boldsymbol{x}$ expressed in a reference system S into a coordinate vector $\boldsymbol{x}'$ expressed in a reference system S' is computed by (8.36), which consists of a translation for the shift in origin, rotation, and scale change in all three axes (Fig. 8.16)

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \partial x \\ \partial y \\ \partial z \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} . \quad (8.36)$$

The rotation matrix also includes a scale change. For the transformation of coordinates, this affine map-

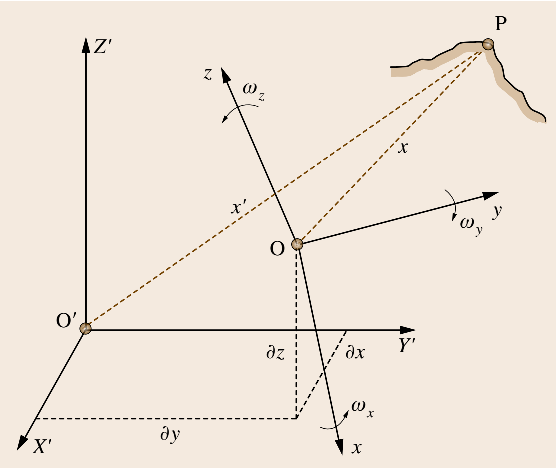

Fig. 8.16 Three-dimensional spatial coordinate transformation

ping is constrained by implying conformity in using only one scale factor m. This leads to the standard seven-parameter similarity transformation widely used in geodesy:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \partial x \\ \partial y \\ \partial z \end{pmatrix} + (1+m) \cdot \boldsymbol{R} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \tag{8.37}$$

$$\boldsymbol{R} = \begin{pmatrix} 1 & \omega_z & -\omega_y \\ -\omega_z & 1 & \omega_x \\ \omega_y & -\omega_x & 1 \end{pmatrix} . \tag{8.38}$$

This transformation is known as the spatial or 3-D seven-parameter Helmert transformation. In general, the rotation angles ω_i around the three coordinate axes are quite small, and the simple form given in (8.38) can be used. The transformation parameters are to be determined by the use of at least three identical points and by solving the equation for the seven parameters. In general this is accomplished by least-squares adjustment.

For sets of ellipsoidal coordinates and the transformation to a new reference ellipsoid an alternative to the computation of new latitude, longitude, and ellipsoid height by concatenation of three operations (geographical to geocentric, geocentric to geocentric, geocentric to geographic) is possible using formulas derived by *Molodenskij* [8.23]. These directly relate the changes in geographical coordinate offsets by use of the transformation parameters for the origin and the change in ellipsoid parameters. Their short form is

$$\begin{aligned}
\varphi' &= \varphi + \mathrm{d}\varphi , \\
\lambda' &= \lambda + \mathrm{d}\lambda , \\
h' &= h + \mathrm{d}h , \\
\mathrm{d}\varphi'' &= [-\mathrm{d}X \sin\varphi \cos\lambda - \mathrm{d}Y \sin\varphi \sin\lambda \\
&\quad + \mathrm{d}Z \cos\varphi + (a\,\mathrm{d}f + f\,\mathrm{d}a) \sin 2\varphi] \\
&\quad /(M \sin 1'') , \\
\mathrm{d}\lambda'' &= \frac{-\mathrm{d}X \sin\varphi + \mathrm{d}Y \cos\lambda}{N \cos\varphi \sin 1''} , \\
\mathrm{d}h &= \mathrm{d}X \cos\varphi \cos\lambda + \mathrm{d}Y \cos\varphi \sin\lambda + \mathrm{d}Z \sin\varphi \\
&\quad + (a\,\mathrm{d}f + f\,\mathrm{d}a) \sin^2\varphi - \mathrm{d}a .
\end{aligned} \tag{8.39}$$

Here, $\mathrm{d}X$, $\mathrm{d}Y$, and $\mathrm{d}Z$ are the geocentric translation parameters, M and N are the meridian and normal radii of curvature at the given latitude φ on the first ellipsoid, $\mathrm{d}a$ is the difference in the semimajor axes of the target and source ellipsoids, and $\mathrm{d}f$ is the difference in the flattening of the two ellipsoids.

8.8.4 Projections and Plane Coordinates

GIS, cartography, and surveying applications need plane coordinates. There are numerous ways to project the ellipsoid onto the plane, but it is not possible to avoid distortions in distances in this process [8.23]. Either areas or angles can be selected as the target quantity not to be distorted so as to obtain an equal area or conformal (equal angular) mapping in the plane. In geodesy and GIS, conformal maps are preferred as the distortions of distances are not dependent on the directions. They are computed from the solution of the Cauchy–Riemann differential equation (8.42). There are many possible choices of conformal maps. The modern form of the transversal Mercator projection, the one of most importance in GIS and geodesy, will be derived below. A comprehensive collection of formulas and parameters can be found in [8.24,25]. The general feature of projections is the fact that they are basically two-dimensional mappings between two surfaces so that heights above these surfaces remain unaffected.

Conformal mapping is facilitated by the one-to-one transformation of an isothermal net of parameter lines from one surface to the other. "Isothermal" means that both sets of parameter lines are orthogonal and of the same scale (isometric). The plane Cartesian coordinate set is isothermal; the net of meridians and parallels, i. e., the geographic longitude λ and the latitude φ, are not. However, an isothermal net on the sphere or the ellipsoid can be generated by use of the Mercator function to convert the latitude φ to the isothermal latitude q:

$$q = \ln\left[\tan\left(\frac{\pi}{4}+\frac{\varphi}{2}\right)\left(\frac{1-e\sin\varphi}{1+e\sin\varphi}\right)^{e/2}\right] . \quad (8.40)$$

The inversion is done iteratively by using $\varphi = 0$ as a starting value

$$\varphi_{i+1} = 2\arctan\left[\left(\frac{1-e\sin\varphi_i}{1+e\sin\varphi_i}\right)^{e/2}\cdot\exp(q)\right] . \quad (8.41)$$

Conformal mappings are now computed from the holomorphic function

$$x+\mathrm{i}\cdot y = f(q+\mathrm{i}\cdot\lambda) , \quad (8.42)$$

where the plane coordinate system components x and y are the real and the imaginary part of the complex function (8.42).

8.8.5 Meridian Strip Projection (Transverse Mercator Projection)

This projection goes back to Gerhard Mercator in 1569 and was later derived by Gauss for ellipsoidal coordinates. The formulas for practical calculations were developed by various geodesists in different countries and this projection is known under several names; the internationally accepted name is *transverse Mercator projection* (TMP). Maps projected by the TMP have straight lines for all meridians and parallels. A particular meridian of the ellipsoid is adopted as the central meridian (CM) of the projection. Usually the central meridian in its true length is chosen with a longitude divisible by 3, i.e., 6°, 9°, 12°. A small region, $\pm\Delta\lambda$ east and west of the CM, is then mapped by (8.42), so that the projection of the CM becomes the x- or North-axis. The projection of the equator is the y- or East-axis (Fig. 8.17). The origin of the plane coordinate system is thus the intersection of the CM with the equator of the reference ellipsoid at the chosen reference meridian. The distortion increases with distance from the CM. Therefore $\Delta\lambda$ is usually limited to 1.5° or 3°, leading to stripes of 3° or 6° width.

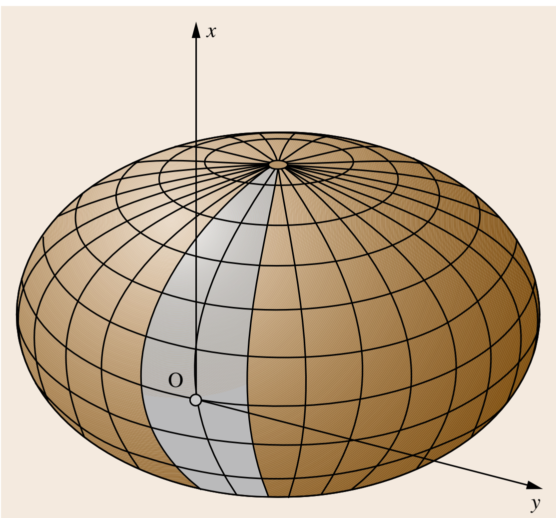

Fig. 8.17 Meridian strip projection scheme

For the conversion, practical formulas are published based on series expansions [8.26]. An analytical solution based on the complex variables of the holomorphic function (8.42) that is easily implemented was published in [8.27] and elaborated upon in [8.28]. It is facilitated in a two-step procedure for forward and backward conversion from ellipsoidal to plane coordinates.

Conversion Between $(\varphi,\lambda) \leftrightarrow (q,l)$

$$\begin{aligned} q &= \arctan h(\sin\varphi) - e\arctan h(e\sin\varphi) , \\ l &= \lambda , \quad \text{and} \\ \varphi_{i+1} &= \arcsin\tan h(q + e\arctan h(e\sin\varphi_i) , \\ \varphi_1 &= 0 \\ \lambda &= l . \end{aligned} \quad (8.43)$$

Conversion Between $(q,l) \rightarrow (x,y)$

Here the transformation between the two sets of isothermal coordinates is solved by the complex function $z = f(w)$, where the analytical function f is determined by the presupposition of an undistorted CM

$$\begin{aligned} w &= q + \mathrm{i}\cdot l \\ z &= x + \mathrm{i}\cdot y \\ b_{i+1} &= \arcsin\tan h[w + e\arctan h(e\sin b_i)] , \\ b_\mathrm{i} &= (\bar{\varphi} + \mathrm{i}\cdot\bar{\lambda}) \\ z &= M_0(1+E)\cdot b - M_0\tfrac{1}{2}\sin(2b)E_b \quad \text{with} \\ E(e'^2) &= \sum_{n=1}^{\infty} d''_n \\ E_\varphi(e'^2,\varphi) &= \sum_{n=1}^{\infty} d''_n \sum_{k=0}^{n-1} k''_n \\ d_{n+1} &= d_n\frac{(2n+1)(2n+3)}{(2n+2)(2n+2)}e^2 ; \quad d_0 = 1 , \\ k_{i+1} &= d_n\frac{(2i+2)}{(2i+3)}\sin^2\varphi ; \quad k_0 = 1 . \end{aligned} \quad (8.44)$$

The plane coordinates (x, y) are the real and imaginary part of the complex variable z. Note that in many countries a constant [false easting (FE) and false northing (FN)] is added to the y-coordinate in order to avoid a negative sign and to indicate the zone or chosen CM. Frequently, these coordinates are then termed *right* or *East* for y, and *high* or *North* for x. Of course, when doing the inverse transformation $x, y \to \varphi, \lambda$ this value has to be eliminated first. In addition, the CM may not be projected isometrically but must be multiplied by a scale factor as in the case of UTM coordinates.

For the inverse transformation the inversion of the above formulas is used, and q and l, respectively, are the real and imaginary part of w.

Conversion Between $(q,l) \to (x,y)$

$$
\begin{aligned}
b_{i+1} &= \frac{z}{M_0(1+E)} + \sin(2b_i)\frac{Eb_{(i-1)}}{2(1+E)}\ ; \quad b_0 = 0\ ,\\
w &= \operatorname{arctan} h(\sin b) - e \operatorname{arctan} h(e \sin b) \quad \text{with}\\
M_0 &= a(1-e^2)\ . \qquad (8.45)
\end{aligned}
$$

The formulas exhibit quick convergence and iterations can be limited to $n = 3$ or $n = 4$ in most cases to arrive at centimeter accuracy.

Distortions of Distance and Areas. The conformal mapping leaves angles unchanged, but distances and areas distorted the farther from the CM they are located. Distances and areas are enlarged by the projection in principle. Let s be the length of a geodetic line on the ellipsoid, s' the corresponding length in the plane, and the areas be A on the ellipsoid and A' in the plane. For a sufficiently good approximation of the distortions it suffices to use an osculating sphere of radius R_m, where R_m is the mean radius at the mid-latitude of the region in question. Then, with y_1 and y_2 as the y-coordinates of the endpoints of a line or of the nearest and farthest point of an area with respect to the CM,

$$
\begin{aligned}
s' &= s + \frac{s}{6R_\mathrm{m}^2}\left(y_1^2 + y_1 y_2 + y_2^2\right)\ ,\\
A' &= A + \frac{A}{3R_\mathrm{m}^2}\left(y_1^2 + y_1 y_2 + y_2^2\right)\ ,\\
R_\mathrm{m}^2 &= \frac{c^2}{V_\mathrm{m}^4}\ . \qquad (8.46)
\end{aligned}
$$

8.8.6 Universal Transverse Mercator System

To limit distortions, a grid system with several grid zones and common defining parameters is used. Coordinates throughout the system are repeated in each zone. To make coordinates unambiguous the easting is prefixed by the relevant zone number. This procedure was adopted, e.g., by German mapping through the Gauss–Kruger systems and later by US military mapping through the universal transverse Mercator (UTM) grid system (Fig. 8.18). The CM is a derived parameter to be computed from two other defining parameters, the initial longitude (the western limit of zone 1) (λ_I) and the zone width (W). Each of the remaining four transverse Mercator defining parameters – CM or latitude of natural origin, scale factor at natural origin, false easting, and false northing – have the same parameter values in every zone.

The standard transverse Mercator formulas above are modified as zone number $Z = \mathrm{INT}[(\lambda + \lambda_\mathrm{I} + W)/W]$ with λ, λ_I, and W in degrees; λ_I is the initial longitude of the zoned grid system, and W is the width of each zone of the zoned grid system. If $\lambda < 0$, $\lambda = (\lambda_\mathrm{I} + 360)°$. Then

$$
\lambda_0 = (ZW) - \left[\lambda_\mathrm{I} + \left(\frac{W}{2}\right)\right]\ . \qquad (8.47)
$$

For the forward calculation, the easting and northing are

$$
\begin{aligned}
\mathrm{E} &= Z \times 106 + \mathrm{FE} + k_0 y\ , \qquad (8.48)\\
\mathrm{N} &= \mathrm{FN} + k_0 x\ , \qquad (8.49)
\end{aligned}
$$

and in the reverse calculation

$$
\begin{aligned}
x &= \frac{\mathrm{N} - \mathrm{FN}}{k_0}\ , \qquad (8.50)\\
y &= \frac{\mathrm{E} - (\mathrm{FE} + Z \times 106)}{k_0}\ . \qquad (8.51)
\end{aligned}
$$

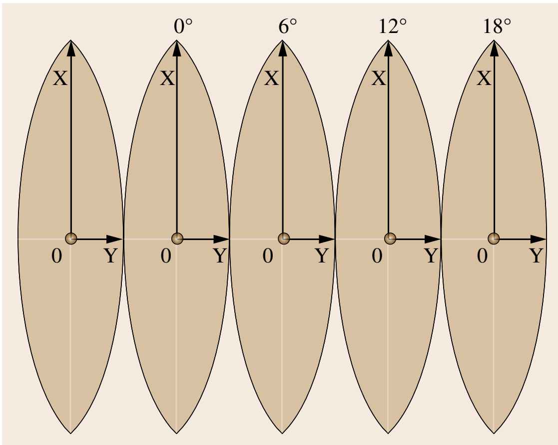

Fig. 8.18 Meridian strip projection scheme for UTM coordinates

For the UTM coordinates introduced by the US military and now being used as standard in many countries, the Earth is divided into 60 zones, each having a longitudinal extension of 6° (Fig. 8.18). The zones are numbered from 1 to 60 beginning with the zone between 180° and 174° West of Greenwich and progressing eastward. The central meridians λ_0 are $-177°$, $-171°$, $-165°$, and so forth. The defining parameters are $\lambda_I = 177°$, $k_0 = 0.9996$, FE = 500 000 m, FN = 0 m North of equator or FN = 10 000 000.0 m North of 80° Southern latitude to equator, and $W = 6°$.

By use of the scale factor k_0, isometry at the CM is lost, but now occurs for two curves parallel to the CM at a distance of about ±180 km. Note that within these two curves distortions are negative, whereas outside of them distortions are positive.

The UTM system is used only up to latitudes of 80° North and 80° South. This is due to the convergence of the meridians causing rapidly varying zones when approaching the poles. For those regions the polar stereographic projection, which is also conformal, is used. This is a projection onto a plane tangent to the ellipsoid at either one of the poles. The origin of the plane system is the pole, and the images of the meridians are straight lines through the pole. Parallels of equal latitude are concentric circles.

An often-used projection is the Lambert conformal conic projection, which is also used in the international aeronautical charts.

8.8.7 Datum Transformation

The geodetic datum includes all information necessary to define a geodetic system. Considering dimension of the reference ellipsoid, the seven parameters for the transformation to the ITRF and the height reference, ten parameters in total are required to define the geodetic datum unambiguously. These parameters are the three translations of the origin, three rotations around the Cartesian axes, a scale factor, the two reference ellipsoid parameters a and f. In addition, for the vertical reference system of the heights, the potential of the geoid W_0 has to be specified. In the classical case of separated position and height coordinates, such as with two-dimensional triangulation networks, the datum had a slightly different definition. In that case it comprised a reference surface consisting of the parameters: the latitude and longitude of an initial point (origin), the orientation of the network, and the two parameters of a reference ellipsoid. A map projection needs information about a reference surface that is fixed by a geodetic datum in space (superior coordinate reference system). National mapping agencies define their own reference system. Different countries may use the same parameters of the reference surface but with different position and orientation. National mapping agencies may use different map projections, based on the same reference system. Changing map projection between countries normally needs a geodetic datum transformation, and different map projections depend on scale, area and anticipated distortion characteristics.

Plane coordinates for a particular ground location and its height will vary based on the datum used to produce a particular map or chart. Therefore, it is essential that the datum used to derive the coordinates be included when reporting positions. ITRS and the GRS80 (or equivalently but less accurately the WGS 84) now provide the single standard reference datum for geographic reference system worldwide. They should be used in all new coordinate determinations and GIS applications. For the unification of older data, respective transformations of the underlying reference systems and datums have to be performed based on the formulas described above. The general procedure of transformation and conversion is recommended as follows.

The change of coordinates from one CRS to another often involves several steps. A scheme that involves a series of operations consisting of one or more transformations and one or more conversions is given in Fig. 8.19. If the necessary information is available, the approach from plane coordinates in the source system CRS1 to plane coordinates in another target system CRS6 should ideally be made by using the Cartesian ITRF coordinates to avoid distortions. For a rigorous transformation, the use of ellipsoidal coordinates of CRS2 and CRS5 is required. Databases of the parameters are compiled by various organizations [8.25]. However, for historical data the ten parameters required are not available in many cases. Here, it may be appropriate to use two- or three-dimensional similarity transformations based on adjustments performed over identical points. Missing height values, missing information on the type or reference of the heights, and missing information on the ellipsoid and the datum of the triangulation are potential sources for major distortions and errors in the transformed coordinates, and sophisticated methods have to be applied.

Table 8.3 presents some examples of datum transformation parameters for illustration purposes. The origins of the historical systems prior to satellite geodesy all have large offsets and differences in the

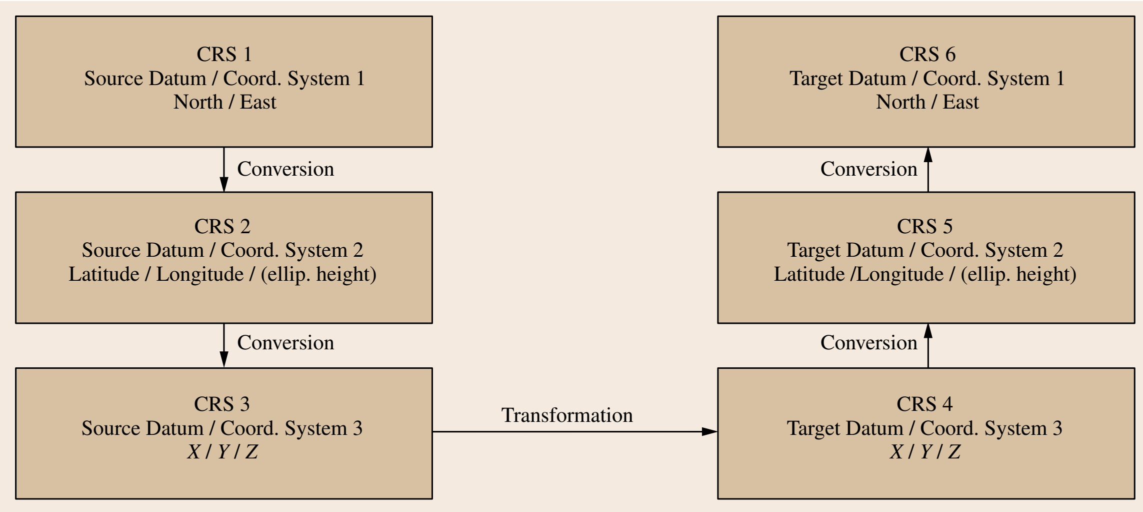

Fig. 8.19 Example of combined coordinate conversion and transformation (after [8.29])

Table 8.3 Examples of historical and actual datum transformation parameters to ITRF2007 and GRS80 ellipsoid

Datum	Ellipsoid	da (m)	df (10^{-3})	dX (m)	dY (m)	dZ (m)	M (10^{-9})	ω_x (″)	ω_y (″)	ω_z (″)
European 1950	International 1924	251	141.92702	−87	−96	−120	–	–	–	–
North American 1927	Clarke 1866	69.4	372.64639	−9	161	179	–	–	–	–
Pulkovo 1942	Krassovsky 1940	108	4.80795	28	−121	−77	–	–	–	–
ITRF 2005	GRS 80	–	–	0.03	0.04	−0.08	0.59	0.0	0.0	0.0
ITRF 97	GRS 80	–	–	0.02	0.03	−0.06	0.43	–	–	–
WGS 84 since 1994	WGS 84	–	0.00016	0.01	−0.01	0.02	7.7	0.003	0.000	−0.003
WGS 84 before 1994	WGS 84	–	0.00016	−0.060	0.52	0.22	0.01	−0.0183	0.0003	−0.007

semimajor axis a and flattening f. Modern CRS such as the WGS 84 are geocentric, and the parameters depend on the precision of their coordinate determination. For the most recent WGS 84 realization and the various ITRF realizations evolving with time, the transformation parameters just reflect improvements in reference frame determination.

Actual parameters should be taken from [8.4, 12].

8.9 Coordinate Determination

The geocoding of objects or features is done by geodetic techniques. The two basically different approaches are space-based coordinate determination by GPS observations and terrestrial geodetic methods such as tacheometry. Terrestrial methods primarily result in coordinates in a local topocentric system and relative to an existing geodetic reference point. GPS coordinate results are in absolute global coordinates, e.g., in the WGS 84, the ITRF or a national datum. The principles of these two different approaches are explained in the next two subsections.

8.9.1 GPS Coordinate Determination

GPS can be used as a proxy for all future GNSS, such as the Russian GLONASS, the European Galileo system, and the Chinese COMPASS system, which are to become fully operational by 2014. Although built for military purposes in the first place, civil use is guaranteed by the operating governments. The space segment of each of the systems, a set of 24–30 satellites, orbits the Earth at an average height of 20 000 km in a time of about 12 h. These orbits are very precisely known from tracking network observations. Each of the satellites transmits the information on its position on at least two microwave frequencies; for GPS these are 1575.42 MHz (L1) and 1227.6 MHz (L2). The carrier frequency is modulated by codes so that each satellite can be identified and the distance between the satellite and the user receiver–antenna can be calculated. This distance is called the *pseudorange* as it is derived from the travel time computed from the precise timing signals included in the data message. Based on the known position of at least four satellites and the respective ranges, the user's position is computed. As the orbits are given in the ITRF, the user coordinates are also in this global geocentric reference frame. In practice, most of today's receivers offer a coordinate conversion, e.g., to ellipsoidal coordinates and height, optionally ellipsoidal or above sea level if the EGM model is considered. Alternatively, projected plane coordinates in a specified datum often are implemented.

There are different levels of accuracy that can be achieved by GPS observations depending on the observables used. Basically four levels may be distinguished, as given in Table 8.4.

The respective receivers needed for the different levels are quite different in price; the dual-frequency and the phase observables options usually associated with geodetic receivers are the most expensive. The use of phase observables is the most advanced technique and requires sophisticated software for evaluation. The reason is the ambiguity inherent in the phase observations. The distance is composed of an integer number of full-wavelength cycles of the carrier wave plus an actually observed phase value. The integer number of cycles cannot be determined directly by the receiver, and therefore longer observing times or advanced filtering techniques are needed to resolve these ambiguities. Only after these procedures can accuracy of a few centimeters be obtained. In case the phase observables are used in combination with the code observations, the point positioning is named *precise point positioning* (PPP). For further information on the use of GNSS, refer to [8.30].

The resultant positional accuracy is determined essentially by disturbances in the satellite clock, the atmosphere, and the geometrical constellation of the satellites used. The accuracy to be expected from GPS observations can be estimated in advance by use of the dilution of precision value (DOP). There are several versions of this; for point positioning the geometrical dilution of precision (GDOP) and the position dilution of precision (PDOP) are the most relevant. The GDOP describes the quality of the solution; it is proportional to the volume generated by the polyhedron formed by the satellites and the user's antenna. Both PDOP and GDOP can be computed in advance without actual observations from approximate coordinates and predicted satellite ephemeris, e.g., for planning the optimal time for a measurement. The PDOP is computed from the trace of the cofactor matrix of the adjustment of the position. Having conducted observations, the accuracy is obtained from the covariance matrix of the final adjustment results of the position computation. GDOP values less than or equal to 4 are well suited for positioning, whereas values larger than 8 should be avoided. For actual surveys it should be observed that the visibility to the sky is such that a homogeneous distribution of more than four satellites at elevations larger than 10° are available.

Relative Positioning

Due to the high spatial correlations of the main GPS error sources, differential observations are often used to mitigate these influences. This mode of operation is

Table 8.4 GPS observables and typical accuracy for observations and coordinates

Observation type	Range accuracy (m)	Typical coordinate accuracy (m)
Code observations at one frequency	10–30	10
Code observations at one frequency supplemented by augmentation system information such as EGNOS in Europe or WAAS in the USA	5–10	5
Dual-frequency code observations	2–5	2–3
Carrier phase observations at two frequencies	0.003 m	0.01–0.03

called differential GPS (DGPS). A reference receiver is placed at a base station with known position and logs the GPS data. One or more roving receivers are used for new points. The difference of the known and the actual determined position at the base station is used to compute corrections to the observables. These corrections – or alternatively the raw observation data of the base receiver – can be transmitted by radio link or via the Internet to the rover and be evaluated in real time or in postprocessing. There are three basic types of differential corrections.

1. Corrections to the coordinates determined by the rover. These are associated with the National Marine Electronics Association (NMEA) format standard NMEA-183.
2. Corrections to the pseudoranges observed at the rover These are associated with the Radio Technical Commission for the Maritime Services (RTCM) or Radio Technical Commission for Aeronautics (RTCA) standards like RTCM-10403.1.
3. Raw phase data or corrections to phase data. These are associated with RTCM version 2.3 or version 3 standards (REF RTCM).

The improvement in accuracy by differential GPS depends on the distance between base and rover. In many countries commercial services are available that offer the various corrections. They are based on dense networks of reference stations like the German Satellitenpositionierungsdienst der deutschen Landesvermessung (SAPOS) or the National Geodetic Survey (NGS) Continuously Operating Reference Stations (CORS) in the US [8.31]. The accuracy is in the range of 3–6, 0.6–2 m, or at best for the differential phase data usage at the 0.01–0.03 m level.

GNSS Outlook

The perspective for GNSS positioning is bright. The *system of systems*, with four GNSS and more than 100 satellites, 20 of them above the horizon at a time, will allow centimeter accuracy in real time. Furthermore, the new signal types and new frequencies of Galileo and GPS will also allow use of GNSS in areas that are partly shaded and to a certain degree for applications at the meter level of accuracy indoors.

8.9.2 Terrestrial and Local Coordinate Determination

Terrestrial measurements are made in the local topocentric coordinate system. Terrestrial techniques are used in areas with low GPS observability and short distances in the kilometer range. Instruments are oriented along the local vertical and so related to the actual gravity field at the observation site. Primarily slant distances, and horizontal and vertical angles are observed by tacheometers to allow determination of polar coordinates and the attachment of new points to known reference points. Often the combination of GPS with a tacheometer is practical. At least one known point for the setup of the

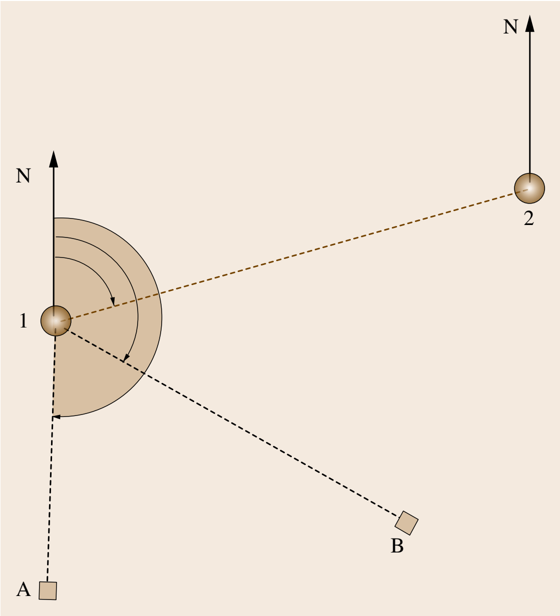

Fig. 8.20 Polar coordinate determination; *rectangles* are known points, *circles* are new points

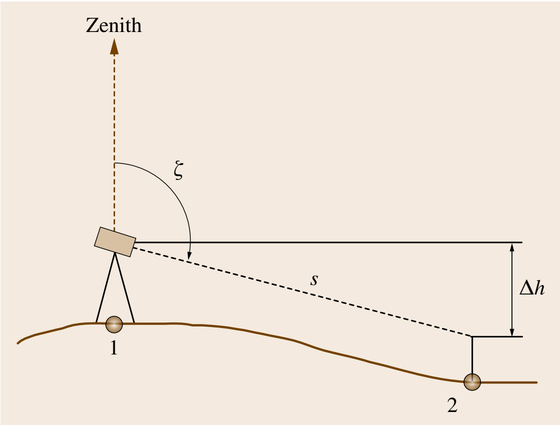

Fig. 8.21 Measurement of height difference by tachometer: distance and zenith angle

instrument and one other for the direction reference to the grid coordinates is needed (Fig. 8.20). For height determination, the distance and the vertical angles (zenith angle) are used (Fig. 8.21). The local observations have to be corrected to consider the local distortion and then allow the computation of grid coordinates, e.g., UTM northing and easting. Alternatively, the local 3-D vector can be transformed into geocentric ITRF Cartesian coordinates by using (8.31).

Modern tacheometers are available with several levels of accuracy; typical instruments measure distance at the level of $0.003 \pm 2 \times (10^{-6} \times \text{distance})$ in meters. In combination with uncertainty of angular measurements of less than 4 seconds of arc, coordinates can be determined with uncertainties of 0.01–0.03 m over 1 km. Tacheometers comprise the full integration in a software suite for combination with GPS; object coding for GIS by an electronic field book and a database connection. Mobile GIS applications allow the formation of objects from line and area data and the possibility of object classification with descriptive data in the field.

References

8.1 G. Beutler, M. Rothacher, S. Schaer, T.A. Springer, J. Kouba, R.E. Neilan: The International GPS Service (IGS): An interdisciplinary service in support of Earth sciences, Adv. Space Res. **23**, 631–653 (1999)

8.2 R. Rummel: Integrated global geodetic observing system (IGGOS), J. Geodyn. **40**, 357–362 (2005)

8.3 IERS: *International Earth Rotation and Reference Frames Service* (Agency for Cartography and Geodesy, Frankfurt 2010), http://www.iers.org/ (last accessed 03/10/2010)

8.4 ITRF: *International Terrestrial Reference Frame* (Institute Geographique National, Paris 2011), http://itrf.ign.fr/

8.5 E. Ma: The international celestial reference frame as realized by very long baseline interferometry, Astron. J. **116**, 516–546 (1998)

8.6 Eurogeographics: http://www.eurogeographics.org (Champs-sur-Marne 2010)

8.7 IERS: Conventional Terrestrial Reference System and Frame, IERS Conventions (IERS, Sevres 2007), Chap. 4

8.8 Z. Altamimi, X. Collilieux, J. Legrand, B. Garayt, C. Boucher: ITRF2005: A new release of the International Terrestrial Reference Frame based on time series of station positions and Earth orientation parameters, J. Geophys. Res. **112**, B09401 (2007), doi:10.1029/2007JB004949

8.9 C. DeMets, R.G. Gordon, D.F. Argus, S. Stein: Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, Geophys. Res. Lett. **21**(20), 2191–2194 (1994)

8.10 J. Ihde, J. Luthardt, C. Boucer, P. Dunkley, B. Farrell, E. Gubler, J. Torres (Eds.): *European Spatial Reference Systems, Frames for Geoinformation Systems* (CRS-EU Operator/Bundesamt für Kartographie und Geodäsie, Leipzig 1999)

8.11 M.B. Pearse: *Realisation of the New Zealand Geodetic Datum 2000*, Office of the Surveyor-General Technical Report No. 5 (Land Information New Zealand, Wellington 2000), http://www.linz.govt.nz/geodetic/standards-publications/technical-reports/index.aspx (last accessed 03/10/2010)

8.12 NIMA: *Department of Defense World Geodetic System 1984: Its defnition and relationsships with local geodetic systems*, Technical Rep. TR8350.2 (National Imagery and Mapping Agency NIMA, Bethesda 1997)

8.13 H. Moritz: Geodetic Reference System of 1980 (GRS80) Bulletin Géodésique. In: *The Geodesists Handbook 1969*, ed. by O.B. Andersen (International Union of Geodesy and Geophysics, Karlsruhe 1988), http://www.gfy.ku.dk/~iag/handbook/geodeti.htm (last accessed 03/10/2010)

8.14 B. Hofmann-Wellenhof, H. Moritz: *Physical Geodesy*, 2nd edn. (Springer, Berlin, Heidelberg 2006)

8.15 K. Bretterbauer: *A Primer of Geodesy*, Report. Inst. for Geodesy and Geophysics; Section for Advanced Geodesy (Technical University, Vienna 2003)

8.16 M. Sacher, G. Liebsch, J. Ihde, J. Luthardt: Steps on the way to UELN05 and enhancements of the web-based Geodetic Information and Service System (CRS-EU), Mitt. Bundesamt. Kartogr. Geodäsie **38**, 158–165 (2006)

8.17 F.G. Lemoine, S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, T.R. Olson: *The Development of the Joint NASA GSFC and National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*, NASA/TP-1998-206861 (Goddard Space Flight Centre, Greenbelt 1998)

8.18 NGA: *Earth Gravitational Model EGM2008* (National Geospatial Agency, Bethesda 2008), http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/index.html (last accessed 03/10/2010)

8.19 W. Torge: *Geodesy* (deGruyter, Berlin 2003)

8.20 NGA-GIS: *Earth Gravitational Model EGM2008-GIS* (National Geospatial Agency, Bethesda 2008), http://earth-info.nima.mil/GandG/wgs84/gravitymod/egm2008/egm08_gis.html (last accessed 03/10/2010)
8.21 B. Guinot, E.F. Arias: Atomic time-keeping from 1955 to the present, Metrologia **42**, 20–29 (2005)
8.22 A. Moudrak, B. Eissfeller, G. Hein, H. Klein: Future time opportunities for using optical clocks in GNSS systems, Inside GNSS **9/10**, 45–50 (2008), http://www.insidegnss.com/node/768 (last accessed 03/10/2010)
8.23 M. Hooijberg: *Geometrical Geodesy* (Springer, Berlin, Heidelberg 2008) p. 439
8.24 J.P. Snyder: *Map Projections – A Working Manual*, Professional Paper No. 1395 (US Geological Survey, Reston 1987), http://pubs.er.usgs.gov/publication/pp1395 (last accessed 03/10/2010)
8.25 OGP: Coordinate conversions and transformations including formulas. In: *Geomatics Guidance Note*, No. 7, Part 2, OGP Publication 373-7-2 (OGP Geomatics Committee, London 2011)
8.26 M. Hooijberg: *Practical Geodesy* (Springer, Berlin, Heidelberg 1997)
8.27 J. Klotz: Eine analytische Lösung der Gauß-Krüger-Abbildung, Z. Verm.-wesen **118**(3), 106–116 (1993)
8.28 N. Stuifbergen: *Wide Zone Transverse Mercator Projection*, Can. Tech. Rep. Hydrogr. Ocean Sci., Vol. 262 (Canadian Hydrographic Service, Dartmouth 2009)
8.29 CRS: *Coordinate Reference Systems in Europe* (CRS-EU Operator/Bundesamt für Kartographie und Geodäsie, Leipzig 2010), http://www.crs-geo.eu (last accessed 03/10/2010)
8.30 B. Hofmann-Wellenhof, H. Lichtenegger, E. Wasle: *GNSS Global Navigation Satellite Systems* (Springer, Vienna, Heidelberg 2008)
8.31 CORS: *Continuously Operating Reference Stations* (National Geodetic Survey, Bethesda 2010), http://www.ngs.noaa.gov/CORS/cors-data.html (last accessed 03/10/2010)