Lecture notes to the courses:
19770 Referenzsysteme
44700 Koordinaten- und Zeitsysteme in der Geodäsie, Luft- und Raumfahrt

Reference Systems

Nico Sneeuw

Geodätisches Institut
Universität Stuttgart

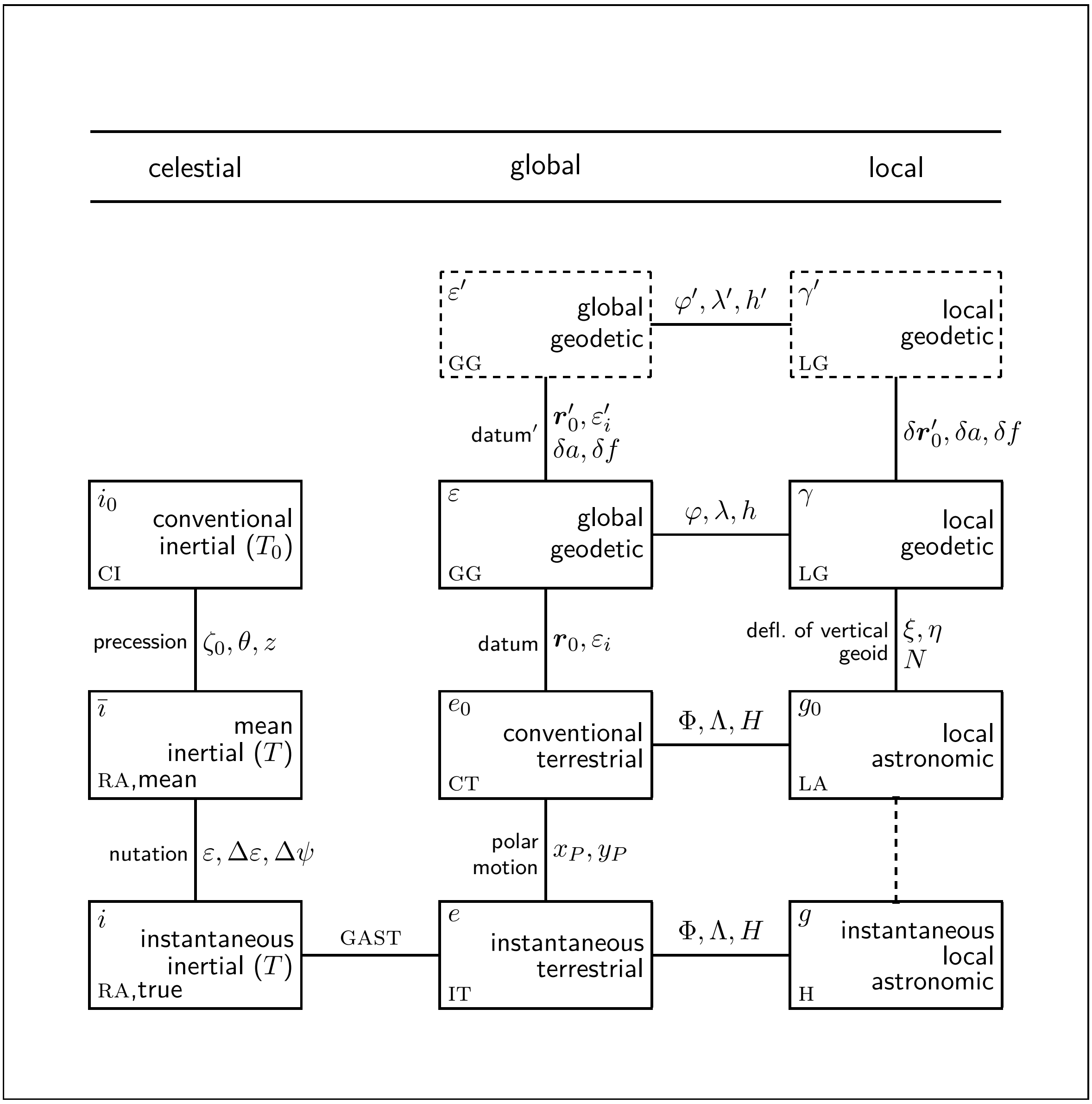
www.gis.uni-stuttgart.de

May 29, 2018

These are lecture notes in progress. They borrow strongly from the lecture notes on the "Fundamentals of Geodesy" by the late K. P. Schwarz, University of Calgary, Department of Geomatics Engineering.
Text and graphics were developed with the support of Matthias Roth and Philipp Hirt, which is gratefully acknowledged.
Please contact me (sneeuw@gis.uni-stuttgart.de) for remarks, errors, suggestions, etc.

Contents

Hierarchy of reference systems



1 Coordinate systems on the ellipsoid

1.1 Basic ellipsoidal geometry

The ellipse is defined as the set of points whose sum of distances to two *foci* is constant. This definition provides a curve in two-dimensional space. The *bi-axial* ellipsoid in 3D space is the result of rotating the ellipse around one of its axes. Brennpunkte zweiachsig

Inspection of fig. 1.1, in which we choose a point on the major axis (left panel), tells us that this sum must be $(a + x) + (a - x) = 2a$, the length of the major axis. The quantity a is called the *semi-major axis*. lange Halbachse

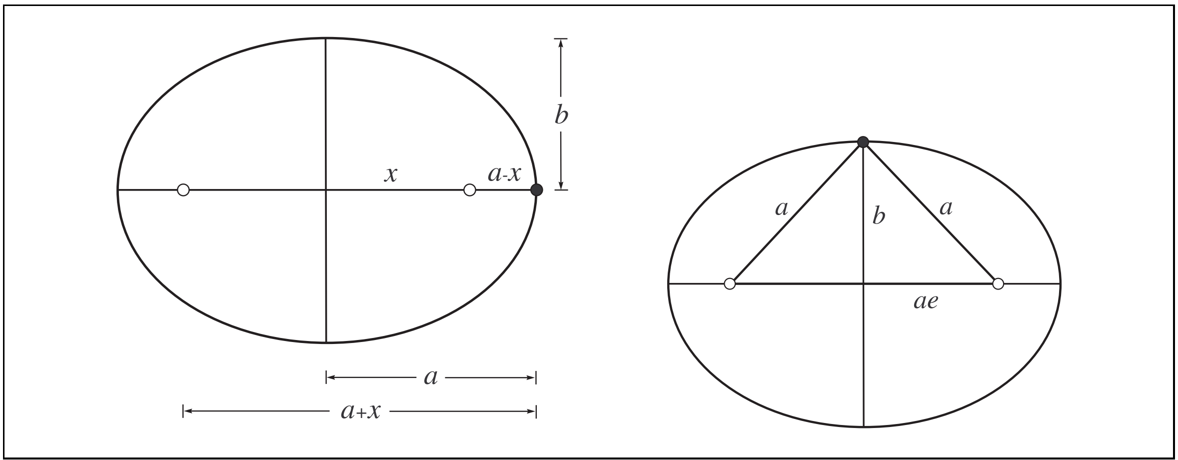


Figure 1.1: Planar geometry of the ellipse.

But then, for a point on the minor axis, see right panel, we have a symmetrical configuration. The distance from this point to each of the foci is a. The length b is called the *semi-minor axis*. Knowing both axes, we can express the distance to focus and centre of the ellipse. It is $\sqrt{a^2 - b^2}$. Usually it is expressed as a proportion e of the semi-major axis a: kurze Halbachse

$$(ae)^2 + b^2 = a^2 \Longrightarrow e^2 = \frac{a^2 - b^2}{a^2}, \text{ or } b = \sqrt{1 - e^2}\, a\,.$$

The proportionality factor e is called the *eccentricity*; the out-of-centre distance ae is known as the *linear eccentricity*, often denoted by the parameter E. Exzentrizität

A point on the ellipsoidal surface can be defined either by curvilinear coordinates, ellipsoidal coordinates or by Cartesian coordinates. The use of curvilinear coordinates appeals more to our intuition and has for a long time been the standard method of representing points in space. Ellipsoidal coordinates will lead to the geodetic coordinates discussed in Chapter 1.2.

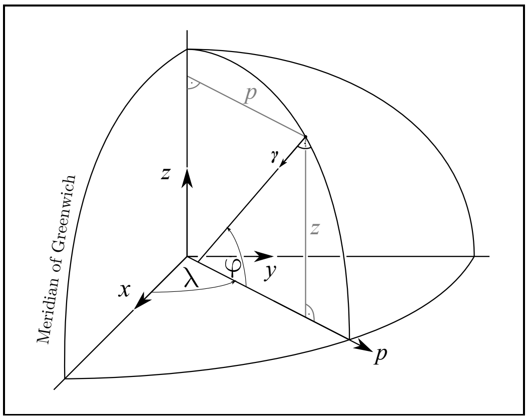


Figure 1.2: Geodetic Coordinates $\{\varphi, \lambda\}$.

Three-dimensional Cartesian coordinates have gained in importance after measurements to satellites became a major component of geodetic operations.

Curvilinear coordinates The curvilinear representation makes use of the fact that the point has to be on the surface of the ellipsoid, and position can thus be defined by two coordinates. This will be the approach taken in this chapter. In the Cartesian representation, three coordinates are needed because the surface is embedded in three-dimensional space. In practice, this difference is of little significance because measurements take place in 3D space and are not confined to the surface of an ellipsoid. Thus, the curvilinear as well as the Cartesian representation requires a three-tupel of numbers to define a position in space. Therefore, a third coordinate, the ellipsoidal height h, has to be defined to represent measurements by curvilinear coordinates. The concept is given in Figure 1.2.

Since the ellipsoid is a close approximation of the geoid, the concepts outlined in Chapters 2.1 and 2.2 can easily be transferred to the ellipsoid. Thus, natural coordinates and the local astronomic system have their equivalents in ellipsoidal geometry. The main simplification is due to the rotational symmetry of the biaxial ellipsoid. Thus, in contrast to the situation on the geoid, each meridian plane has the same geometric properties. The longitude λ is the same for all curvilinear systems on the ellipsoid and it is therefore possible to restrict the discussion to an arbitrary meridian plane. In this plane, the x and y axes in the equatorial plane can be replaced by the p-axis which is defined as the intersection of the meridian plane with the equatorial plane, see Figure 1.2. Thus, a point in the meridian plane is either defined by the latitude or by the coordinates p and z. Different definitions of the latitude are possible and three of them will be discussed in this chapter.

Once p and z have been defined, the $\{x, y\}$-coordinates in a suitably defined Cartesian system, as for instance the CT-system, can be obtained from

$$\begin{aligned} x &= p\cos\lambda, \\ y &= p\sin\lambda. \end{aligned} \tag{1.1}$$

Thus, the transformation into a three-dimensional Cartesian system is given if λ is known.

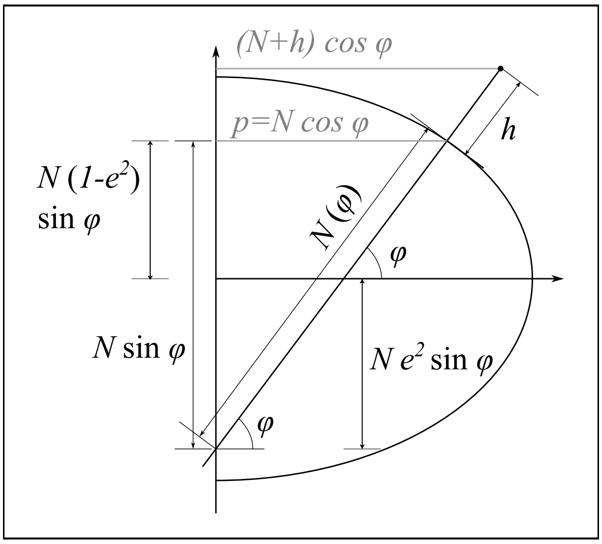


Figure 1.3: Ellipsoidal geometry.

Ellipsoidal coordinates Using a parametrical formulation of ellipsoidal coordinates $\{\varphi, \lambda\}$, for points *on* the ellipsoid the transformation from ellipsoidal to Cartesian coordinates reads:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} N(\varphi)\cos\varphi\cos\lambda \\ N(\varphi)\cos\varphi\sin\lambda \\ N(\varphi)(1-e^2)\sin\varphi \end{pmatrix} \quad \text{, with:} \quad N(\varphi) = \frac{a}{\sqrt{1-e^2\sin^2\varphi}} \tag{1.2a}$$

For points above the ellipsoidal surface, we have to add the ellipsoidal height h in normal direction as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (N+h)\cos\varphi\cos\lambda \\ (N+h)\cos\varphi\sin\lambda \\ (N(1-e^2)+h)\sin\varphi \end{pmatrix} \tag{1.2b}$$

A closed analytical solution for the reverse transformation from Cartesian to geodetic coordinates does exist. Here, however, we will simply apply an iteration. First off, longitude can be determined by: $\tan\lambda = \frac{y}{x}$. But geodetic latitude and height must be solved iteratively together. To that end we introduce the coordinate p again (distance to z-axis):

$$p = \sqrt{x^2+y^2} = (N+h)\cos\varphi$$

$$\text{iteration equation 1:} \qquad h = \frac{p}{\cos\varphi} - N(\varphi)$$

$$z = (N(1-e^2)+h)\sin\varphi \quad \Longrightarrow \quad \frac{z}{p} = \frac{N(1-e^2)+h}{N+h}\tan\varphi$$

$$\text{iteration equation 2:} \qquad \varphi = \arctan\left(\frac{z}{p}\frac{N+h}{N(1-e^2)+h}\right)$$

With the two equations above, the iteration runs as shown in Figure 1.4.

i) Starting value $i = 0$: $h_0 = 0$ (just assume point on surface, if no better information available).

ii) Starting latitude: $\varphi_0 = \arctan(\frac{z}{p}\frac{1}{(1-e^2)})$ from iteration equation 2.

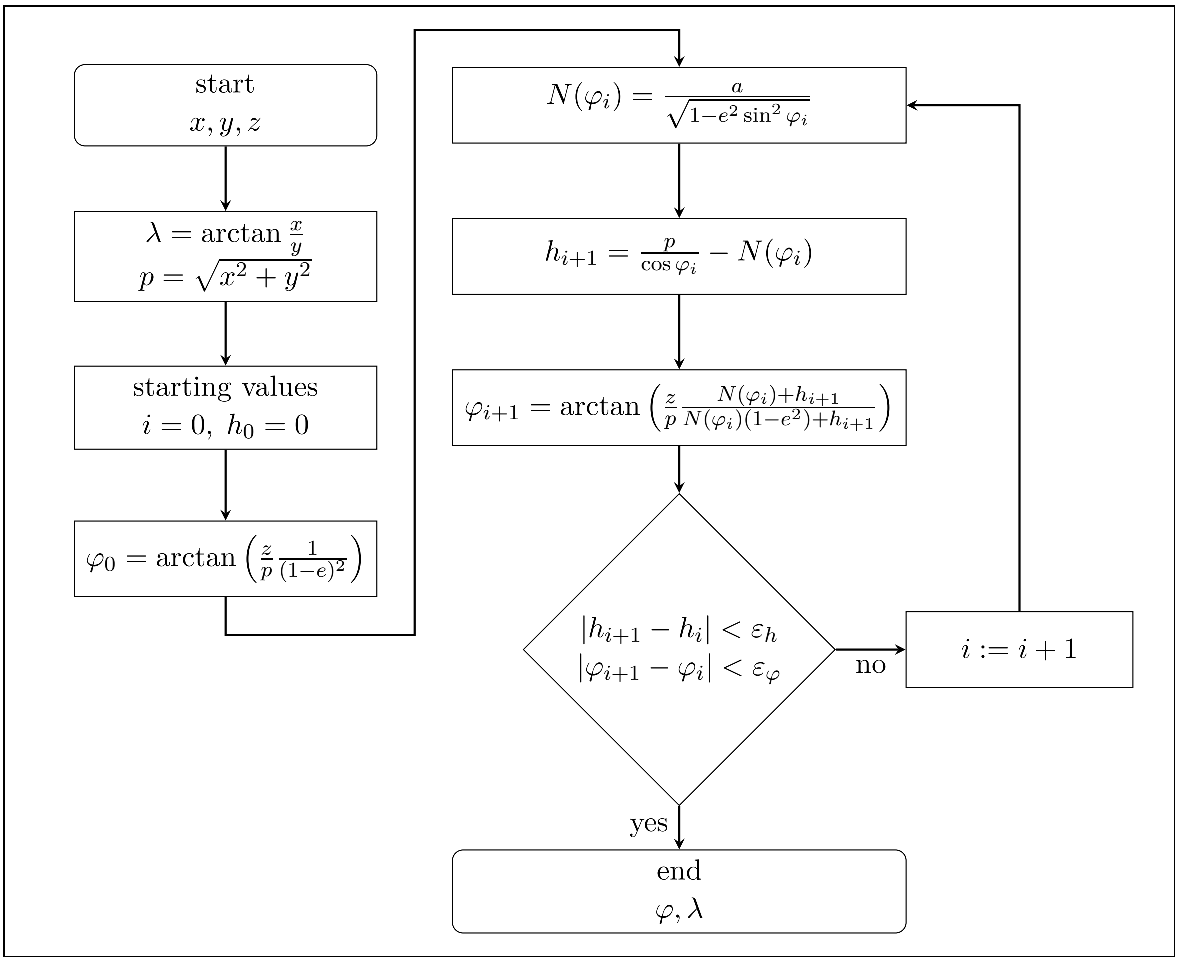


Figure 1.4: Iterative transformation from Cartesian to ellipsoidal coordinates.

iii) $N(\varphi_0) = \ldots$

iv) $h_{i+1} = \frac{p}{\cos \varphi_i} - N(\varphi_i)$ from iteration equation 1.

v) $\varphi_{i+1} = \arctan\left(\frac{z}{p}\frac{N(\varphi_i)+h_{i+1}}{N(\varphi_i)(1-e^2)+h_{i+1}}\right)$ from iteration equation 2 again.

vi) $N(\varphi_{i+1}) =$ and so on.

vii) Iteration until convergence is achieved
$|h_{i+1} - h_i| < \varepsilon_h$
$|\varphi_{i+1} - \varphi_i| < \varepsilon_\varphi$

It ends with achieved convergence when the differences between the values of h and φ from last and current iteration steps are beneath a certain threshold ε_h or ε_φ.

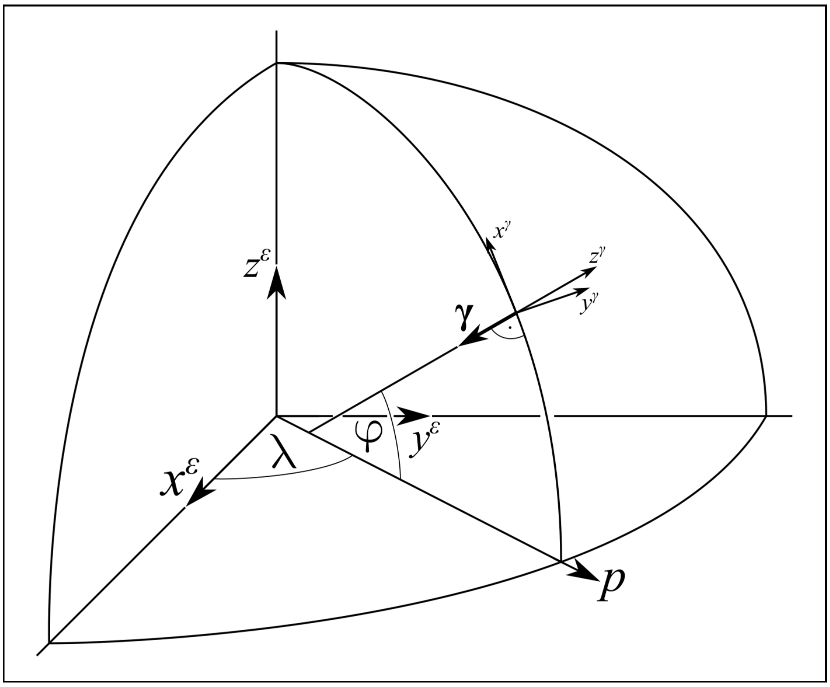


Figure 1.5: Local Geodetic System.

1.2 Geodetic coordinates

In Chapter 2.2, astronomical coordinates $\{\Phi, \Lambda\}$ will be defined. They are given by the direction of the gravity vector in the CT-system. This vector is normal to the equipotential surface passing through the point under consideration.

Geodetic coordinates $\{\varphi, \lambda\}$ are defined in a similar way. They are given by the direction of the normal to the ellipsoid in the CT-system. This vector corresponds to the normal gravity vector if the ellipsoidal surface is the equipotential surface defined by the reference potential U. Geodetic coordinates are important because they are close approximations of the 'observables' Φ and Λ.

The local geodetic system is defined in the same way as the local astronomic system, except that the reference figure is the ellipsoid, not the geoid, see Figure 1.5. We therefore have:

Local Geodetic Frame (LG)	
origin:	at point P
primary axis (z):	orthogonal to ellipsoid at P
secondary axis (x):	tangent to geodetic meridian pointing north
tertiary axis (y):	orthogonal in a left-handed system

Note, that the z-axis coincides with the ellipsoidal normal at P and points in the direction opposite to the normal gravity vector γ at P if the normal ellipsoid is used. The (x, y)-plane coincides with the plane tangent to the ellipsoid at P and normal to γ. The x-axis lies in the meridian plane and is oriented to north. The y-axis is oriented eastward. Thus the local geodetic system is left-handed.

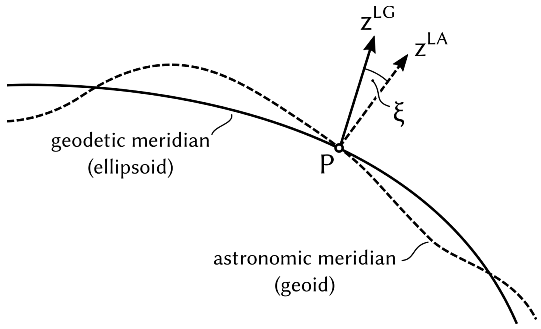


Figure 1.6: Deflection of the vertical in north-south direction (ξ).

The vector $\boldsymbol{n}$ normal to the ellipsoid is

$$\boldsymbol{n}^{\varepsilon} = \begin{pmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ \sin\varphi \end{pmatrix} \tag{1.3}$$

in complete analogy to equation (2.8). Thus, as before, φ and λ define the direction of the surface normal in space.

Geodetic coordinates $\{\varphi, \lambda\}$ are very close to astronomic coordinates $\{\Phi, \Lambda\}$. The difference does usually not exceed $\pm 1'$ and more typically is $15''$ or less.

The quantities

$$\begin{aligned} \xi &= \Phi - \varphi, \\ \eta &= (\Lambda - \lambda)\cos\varphi. \end{aligned} \tag{1.4}$$

are the deflections of the vertical, as described in Chapter 2.3. Here they are geometrically interpreted as changes in direction between the astronomic and the geodetic coordinate systems. The north-south deflection ξ is shown in Figure 1.6. Deflections of the vertical will be further discussed in Chapter 2.3.

The parameters $\{\varphi, \lambda\}$ can also be considered as coordinates on the ellipsoidal surface. The coordinate lines of the system are:

$$\begin{aligned} \text{parallels:} \quad & \varphi = \text{const.}, \\ \text{meridians:} \quad & \lambda = \text{const.} \end{aligned}$$

The transformation of geodetic surface coordinates into p and z will be discussed in chapter 1.3. Transformation of three-dimensional geodetic coordinates into Cartesian coordinates and vice versa have been shown in the introduction of this chapter in the point *ellipsoidal coordinates*.

It should be noted that rotational symmetry which is true for geodetic coordinates does not hold for astronomic coordinates. They refer to the geoid which is not rotationally symmetric. In the literature, geodetic coordinates are frequently called geographical coordinates.

Global Geodetic coordinate systems may refer to ellipsoids of different dimensions and different origins. Although all of them are 'close' to the CT-system in some sense, they are not always identical to it. Such systems are called Global Geodetic Systems and are defined by

Global Geodetic System (G)	
origin:	geometrical centre of the ellipsoid
primary axis (z):	rotation axis of the ellipsoid
secondary axis (x):	meridian of Greenwich
tertiary axis (y):	orthogonal in a right-handed system

Global Geodetic Systems are widely used in geodesy to link the local coordinate systems to a global reference. They may differ in origin, orientation, and ellipsoid dimensions from the CT-system. The reasons for these differences are historical and major efforts are currently made to refer all coordinates to a unique global reference system such as the CT-system. The practical implementation of such a system via GPS satellites (IGS) will be discussed in a later chapter.

1.3 The geocentric latitude

Another way to descpribe latitude is the geocentric latitude φ^z, which refers to the center of the ellipsoid (fig. 1.7). The main difference between geodetic and geocentric latitude is their behaviour concerning the surface of the ellipsoid. From the implicit formulation of the ellipsoid, we can derive the surface normal vector simply by taking the gradient:

3D	**2D**
$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1 = f(x, y, z)$	$\frac{p^2}{a^2} + \frac{z^2}{b^2} = 1 = f(p, z)$
$\nabla f = 2 \begin{pmatrix} \frac{x}{a^2} \\ \frac{y}{a^2} \\ \frac{z}{b^2} \end{pmatrix}$	$\nabla f = 2 \begin{pmatrix} \frac{p}{a^2} \\ \frac{z}{b^2} \end{pmatrix}$
$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} N \cos\varphi \cos\lambda \\ N \cos\varphi \sin\lambda \\ N(1-e^2) \sin\varphi \end{pmatrix}$	$\begin{pmatrix} p \\ z \end{pmatrix} = \begin{pmatrix} N \cos\varphi \\ N(1-e^2) \sin\varphi \end{pmatrix}$

From fig. 1.7 the link between geocentric and geodetic latitude becomes clear:

$$\left.\begin{array}{l} \tan\varphi^z = \frac{z}{p} \text{ (see figure)} \\ \tan\varphi = \frac{z}{b^2} : \frac{p}{a^2} = \frac{a^2 z}{b^2 p} \text{ (from } \nabla f) \end{array}\right\} \Longrightarrow \tan\varphi^z = \frac{b^2}{a^2} \tan\varphi = (1-e^2)\tan\varphi \equiv \frac{z}{p} \,.$$

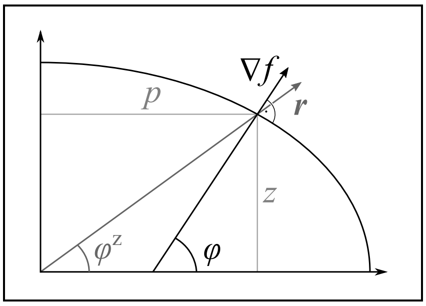


Figure 1.7: Normal vector ∇f vs. radial direction $\boldsymbol{r}$ and link between geodetic latitude φ and geocentric latitude φ^z.

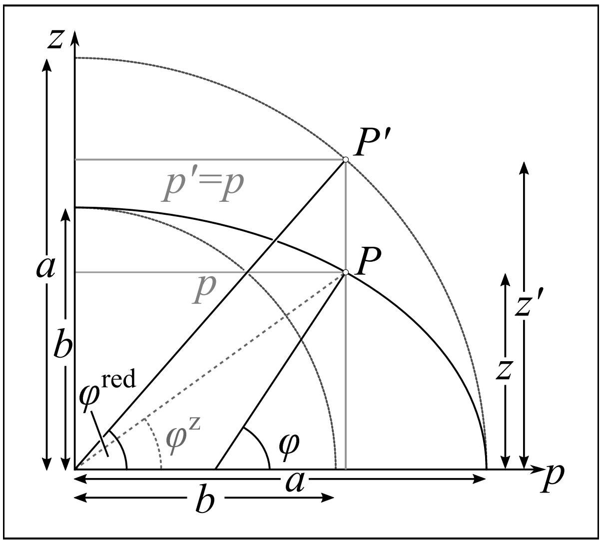


Figure 1.8: Geodetic latitude φ, reduced latitude φ^{red} and geocentric latitude φ^{z}, connected by

$\tan \varphi^{\mathrm{z}} = \frac{b}{a} \tan \varphi^{\mathrm{red}} = \frac{b^2}{a^2} \tan \varphi$.

1.4 The reduced latitude

The concept of the reduced latitude φ^{red} of the point P on the surface of the ellipsoid with semiaxes a and b is given in Fig. 1.8. It shows a section of the ellipsoid along the meridian plane (z, p). The distance p of the point P from the z-axis is the radius of the parallel

$$p = \sqrt{x^2 + y^2}.$$

The parametric representation of the meridian ellipse can be derived from Fig. 1.8. From the circle with radius a we get

$$p' = p = a \cos \varphi^{\mathrm{red}} \qquad \text{and} \qquad z' = a \sin \varphi^{\mathrm{red}}. \tag{1.5}$$

From the circle with radius b we obtain

$$z = z' = b \sin \varphi^{\mathrm{red}}. \tag{1.6}$$

The reduced latitude is obtained by projecting the ellipse on the concentric circle with radius a. It is convenient to use in derivations and is important in the theory of ellipsoidal mappings.

geodetic latitude:	$\begin{pmatrix} x \\ y \\ z \end{pmatrix}^{\varepsilon} = N \begin{pmatrix} \cos\varphi\cos\lambda \\ \cos\varphi\sin\lambda \\ (1-e^2)\sin\varphi \end{pmatrix}$,	$N = N(\varphi) = \frac{a}{\sqrt{1-e^2\sin^2\varphi}}$
reduced latitude:	$\begin{pmatrix} x \\ y \\ z \end{pmatrix}^{\varepsilon} = \begin{pmatrix} a\cos\varphi^{\mathrm{red}}\cos\lambda \\ a\cos\varphi^{\mathrm{red}}\sin\lambda \\ b\sin\varphi^{\mathrm{red}} \end{pmatrix}$	
geocentric latitude:	$\begin{pmatrix} x \\ y \\ z \end{pmatrix}^{\varepsilon} = r \begin{pmatrix} \cos\varphi^{\mathrm{z}}\cos\lambda \\ \cos\varphi^{\mathrm{z}}\sin\lambda \\ \sin\varphi^{\mathrm{z}} \end{pmatrix}$,	$r = r(\varphi^{\mathrm{z}}) = \frac{b}{\sqrt{1-e^2\cos^2\varphi^{\mathrm{z}}}}$

Table 1.1: Transformation of ellipsoidal surface coordinates into cartesian coordinates.

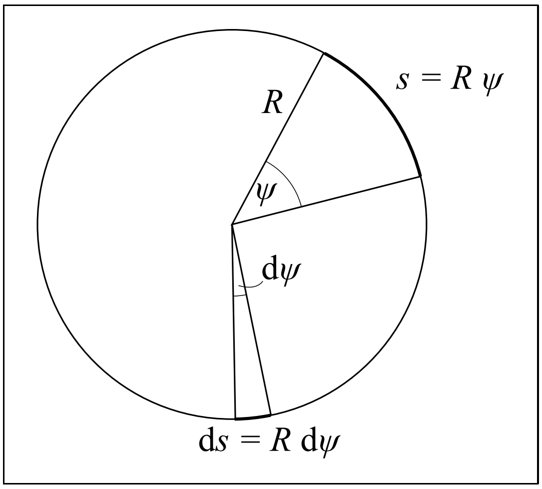


Figure 1.9: Finite and infinitesimal arc length on the sphere.

1.5 Curvature

Sphere An infinitesimal arc length $\mathrm{d}s$ on the sphere is related to its infinitesimal central angle simply by multiplying by the sphere's radius R, see fig. 1.9:

$$\mathrm{d}s = R\,\mathrm{d}\psi\,.$$

This is more or less the translation of $\mathrm{d}\psi$ in angular measure into linear measure. However, it leads to a more fundamental concept, as the quantity

$$\rho = \frac{1}{R} = \frac{\mathrm{d}\psi}{\mathrm{d}s}$$

is called the *curvature*. The radius R is known as the *radius of curvature*. In general, the curvature of a surface is a local quantity, that is, it depends on position. On the sphere, though, curvature is constant. Thus, *surface of constant curvature* can be added as a definition of the sphere.

Krümmung
Krümmungsradius

Ellipsoid On the ellipsoid, on the other hand, the curvature is a local measure. To be more precise:

$$\rho = \rho(\varphi, \alpha)\,,$$

that is, the curvature is latitude and direction dependent. It is a function of latitude φ and on *azimuth* α. At every point on the ellipsoid there will be a direction in which the curvature is maximal and a direction in which it is minimal. Each direction spans up a surface through the local normal vector. Such surfaces are called *normal sections*, see fig. 1.11.

Azimut

Normalschnitte

As might be expected, the two extremes in curvature take place

Meridian-schnitt

i) in the *meridian section*, and

ii) in the prime vertical normal section, which is perpendicular to the meridian section and tangent to the local latitude circle.

Note that the plane through a latitude circle by itself is not a normal section.

Let us consider the curvature and its variations in the meridian and in the equator. The latitude dependence is obvious from fig. 1.10 (left panel). At the equator, the smaller circle fits the ellipse in an optimal way. Its radius is the radius of curvature. It is clear that this radius of curvature is smaller than the semi-major axis a. At the pole, though, the best fitting circle has the largest possible radius, larger than a. Thus the curvature at the pole, $\rho(\varphi = 90°)$, is minimum.

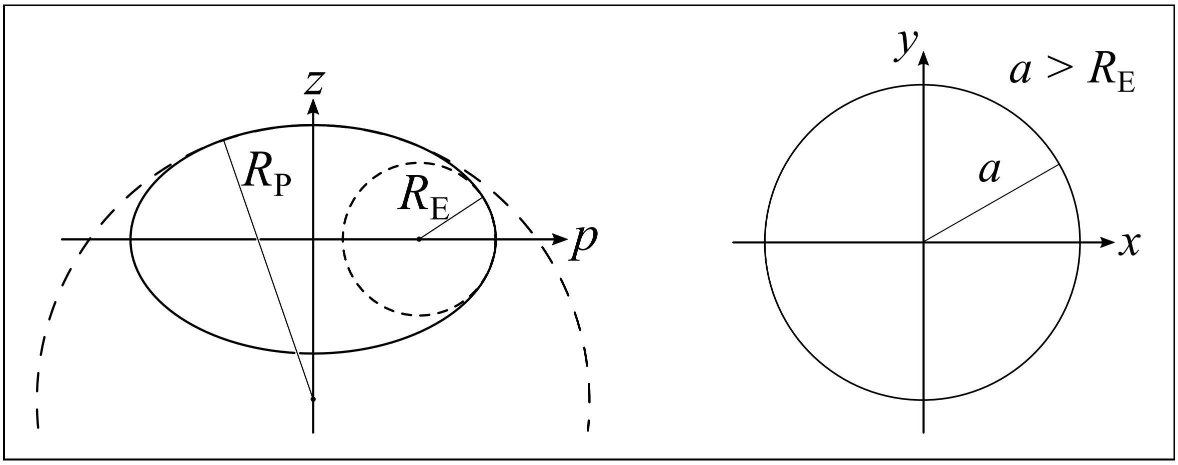
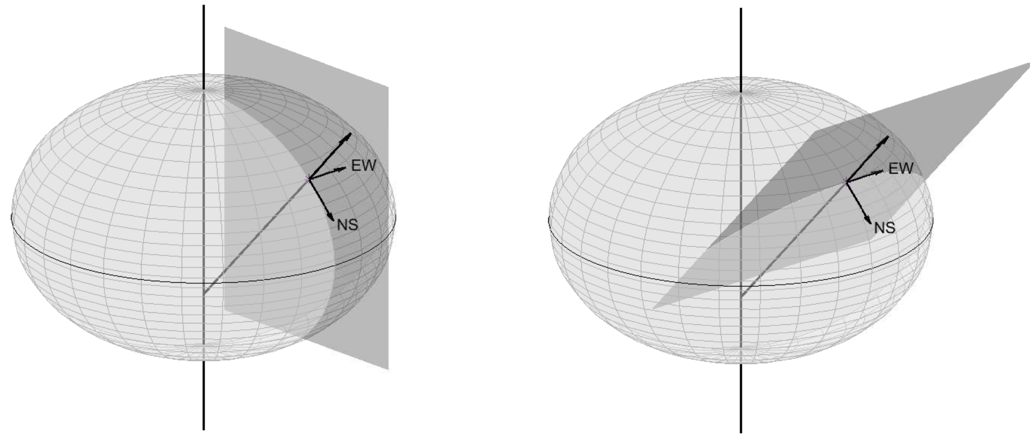


Figure 1.10: Latitude dependence of curvature in the meridian plane (left) and azimuth dependence at the equator (right).

At the pole, no directional dependence can exist, as all meridian planes are normal sections. At the equator, though, there will be a clear difference in curvature between meridian plane (as discussed above) and in the equator plane. The equatorial normal section of the ellipsoid is a circle, see fig. 1.10. The radius of curvature at the equator in East-West direction is therefore a and the curvature $\rho(\varphi = 0°, \alpha = 90°) = 1/a$. In the previous paragraph, we already concluded that the radius of curvature at the equator in North-South direction was smaller than a.



North-South – meridian section

$$\rightarrow M(\varphi) = a\frac{1-e^2}{(1-e^2\sin^2\varphi)^{\frac{3}{2}}}$$

East-West – normal section

$$\rightarrow N(\varphi) = a\frac{1}{(1-e^2\sin^2\varphi)^{\frac{1}{2}}}$$

Figure 1.11: Normal and meridian section with corresponding radii of curvature.

Main radii of curvature This behaviour is not only valid at the equator. At every latitude we will see the minimum radius of curvature (and hence the maximum curvature) in the meridian plane and the maximum radius of curvature in the prime vertical normal section. They are known, respectively, as the *meridian radius of curvature* $M(\varphi)$ and *normal radius of curvature* $N(\varphi)$. The latter radius is exactly the quantity that we know already from (1.2). The corresponding equations and some examples are given in the following table.

Meridian-krümmungsradius

Normal-krümmungsradius

	in meridian	in prime vertical
general	$M(\varphi) = a\dfrac{1-e^2}{(1-e^2\sin^2\varphi)^{3/2}}$	$N(\varphi) = a\dfrac{1}{(1-e^2\sin^2\varphi)^{1/2}}$
at equator	$M(0°) = a(1-e^2)$	$N(0°) = a$
at pole	$M(90°) = \dfrac{a}{\sqrt{1-e^2}}$	$N(90°) = \dfrac{a}{\sqrt{1-e^2}}$

The table indeed confirms that the smallest radius of curvature is in North-South direction: $M(0°) < N(0°)$. Moreover, at the poles there is no azimuth dependence: $M(90°) = N(90°)$.

Gauss curvature The radius of a best fitting sphere at a certain latitude is the Gauss radius of curvature:

$$R_{\mathrm{G}} = \sqrt{MN} = \frac{a\sqrt{1-e^2}}{1-e^2\sin^2\varphi}\,.$$

Mean curvature The mean curvature is defined by:

$$\rho_{\mathrm{M}} = \frac{1}{R_{\mathrm{M}}} = \frac{1}{2}\left(\frac{1}{M} + \frac{1}{N}\right).$$

Curvature in arbitrary direction The mathematician Euler developed a formula that relates the curvatures in North-South direction $\rho(\alpha = 0°)$ and in East-West direction $\rho(\alpha = 90°)$ to the curvature in arbitrary direction:

$$\rho(\alpha) = \frac{1}{R_\alpha} = \frac{\sin^2\alpha}{N} + \frac{\cos^2\alpha}{M}. \tag{1.7}$$

1.6 The direct and inverse geodetic problem on the ellipsoid

geodätische Linie

The shortest path between two points on a curved surface is called a *geodesic*. Solving the direct and inverse geodetic problem on the ellipsoid would require finding and describing geodesics on the ellipsoid. This is a mathematically demanding topic, particularly if analytical solutions are attempted. To exemplify the level of complexity on the ellipsoid, it is remarked that a geodesic is in general not a closed curve, like the great circle on the sphere. It suffices to say that the geodesic is described by a set of three coupled ordinary differential equations, that may be solved numerically.

Meridianbogen

Meridian arc A *meridian arc* s is a special geodesic. It is described by a single differential equation:

$$\frac{\mathrm{d}s}{\mathrm{d}\varphi} = M(\varphi),$$

which is of course the reverse of the definition of a differential arc length (compare the spherical case):

$$\mathrm{d}s = M(\varphi)\,\mathrm{d}\varphi.$$

Therefore, the meridian arc length between two points at different latitudes is

$$s_{1,2} = \int_1^2 \mathrm{d}s = \int_{\varphi_1}^{\varphi_2} M(\varphi)\,\mathrm{d}\varphi,$$

which can be evaluated by numerical quadrature.

1.7 Example: geodetic and geocentric latitude

Assuming the GRS80-ellipsoid with semi-major axis $a_E = 6\,378\,137$ m and semi-minor axis $b_E = 6\,356\,752.3$ m, we derive the eccentricity e and linear eccentricity E.

$$e^2 = \frac{a_E^2 - b_E^2}{a_E^2} = 0.006694$$

$$E = a_E \cdot e = 521\,854\ \mathrm{m}$$

From 1.3 we obtain the link between geodetic and geocentric latitude (assuming $h = 0\,\text{m}$).

$$\left.\begin{aligned} p &= \quad N\cos\varphi \\ z &= N(1-e^2)\sin\varphi \end{aligned}\right\} \Longrightarrow \tan\varphi^{\mathrm{z}} = \frac{z}{p} = (1-e^2)\tan\varphi = \frac{b_E^2}{a_E^2}\tan\varphi$$

With a geodetic latitude of $\varphi = 45°$ we compute the difference in comparison to the geocentric latitude.

$$\varphi = 45° \Rightarrow \tan\varphi = 1 \Rightarrow \tan\varphi^{\mathrm{z}} = \frac{b_E^2}{a_E^2} \Rightarrow \varphi^{\mathrm{z}} = 44°.81 \Rightarrow \Delta\varphi \approx 0°.19 \approx 11'$$

This latitude difference seems small, but its influence on the earth's surface certainly is not.

$$0°.19\ \frac{\pi}{180°}\ a_E \approx 21.4\,\text{km}\ !$$

2 Natural coordinates (e, g)

2.1 Level surfaces and plumb lines (H, C)

In this section, the direction of the *gravity vector* and the characteristics of the surface orthogonal to it will be discussed. Schwerevektor

Let us start by setting

$$W(x, y, z) = W_{\mathrm{P}} = \text{const.}. \tag{2.1}$$

The surfaces defined in this way are surfaces of constant potential, called *equipotential surfaces* or in case of the gravity potential, level surfaces. They coincide with the surface of a homogeneous fluid in equilibrium, which explains the term level surface. In first approximation the idealized surfaces of lakes can be considered as such level surfaces. They approximate $W = \text{const.}$ for a specific value W_{P}. Äquipotenzialflächen

Differentiating $W = W(x, y, z)$ with respect to x, y, z gives

$$\mathrm{d}W = \frac{\partial W}{\partial x}\,\mathrm{d}x + \frac{\partial W}{\partial y}\,\mathrm{d}y + \frac{\partial W}{\partial z}\,\mathrm{d}z$$

or in vector notation

$$\mathrm{d}W = \operatorname{grad} W \cdot \mathrm{d}\boldsymbol{r} = \nabla W \cdot \mathrm{d}\boldsymbol{r} \tag{2.2}$$

where $\mathrm{d}\boldsymbol{r}^{\mathsf{T}} = (\,\mathrm{d}x,\,\mathrm{d}y,\,\mathrm{d}z)$ is a displacement vector and

$$\nabla W = \operatorname{grad} W = \left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}, \frac{\partial W}{\partial z}\right).$$

Let $\mathrm{d}\boldsymbol{r}$ lie in the equipotential surface $W(x, y, z) = W_{\mathrm{P}}$, then $\mathrm{d}W = 0$ and (2.2) becomes

$$\nabla W \cdot \mathrm{d}\boldsymbol{r} = 0,$$

or using the gravity vector $\boldsymbol{g} = \nabla W$

$$\boldsymbol{g} \cdot \mathrm{d}\boldsymbol{r} = 0 \quad \text{on } W(x, y, z) = W_{\mathrm{P}}. \tag{2.3}$$

If the dot product of two non-zero vectors is equal to zero, then the vectors are orthogonal to one another. Since $\boldsymbol{g} \neq \boldsymbol{0}$ and $\mathrm{d}\boldsymbol{r} \neq \boldsymbol{0}$, the gravity vector $\boldsymbol{g}$ must be orthogonal to $\mathrm{d}\boldsymbol{r}$, i. e.

$$\boldsymbol{g} \perp \mathrm{d}\boldsymbol{r}.$$

This means that the gravity vector is normal to the equipotential surface passing through the point P. It is therefore simple to find the direction of the gravity vector on the surface of the Earth. It is orthogonal to the surface established by a level bubble, or, in other words, the

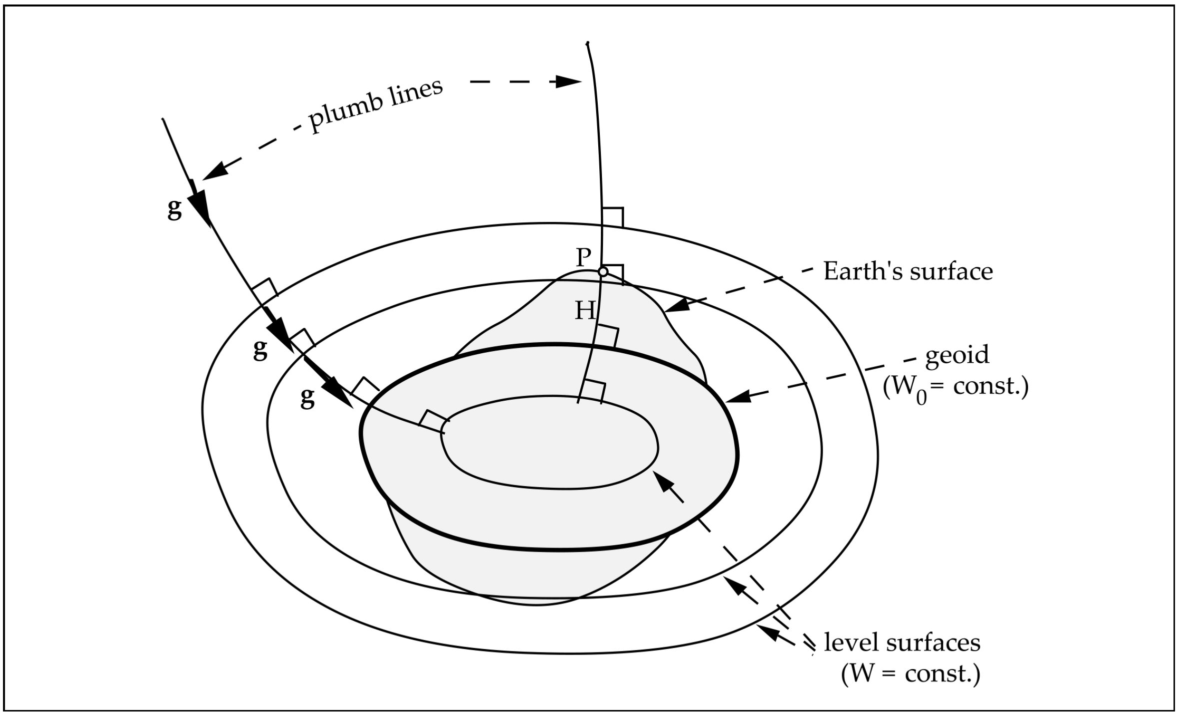


Figure 2.1: Level Surfaces, Plumb Lines, Geoid.

bubble represents the level surface in that specific point. This principle is used extensively when levelling geodetic instruments. The lines which intersect all level surfaces of the Earth orthogonally are called *plumb lines*. They are curved lines and the gravity vector is obviously tangent to the plumb line at the points of intersection. A good approximation of such a tangent, and therefore of the direction of gravity, is the string holding a plumb bob.

Lotlinien

Each specific W_P defines a different equipotential surface. The particular equipotential surface which coincides with the idealized surface of the oceans is called the geoid. The name was proposed by Listing to describe the figure of the Earth. The geoid is used as a reference surface for the orthometric height system. It defines the height H of a point at the physical surface of the Earth by its distance from the geoid measured along the plumb line, see Fig. 2.1.

The following properties of the Earth's equipotential surfaces are of importance in geodesy:

- they are continuous surfaces,
- they never cross each other,
- they are not parallel to one another,
- their curvature changes smoothly from point to point.

To define H mathematically, let us use equation (2.2) again

$$\mathrm{d}W = \boldsymbol{g} \cdot \mathrm{d}\boldsymbol{r}.$$

Let $\mathrm{d}\boldsymbol{r}$ point upward along the plumb line, see Figure 2.2, i.e. $|\mathrm{d}\boldsymbol{r}| = \mathrm{d}H$.

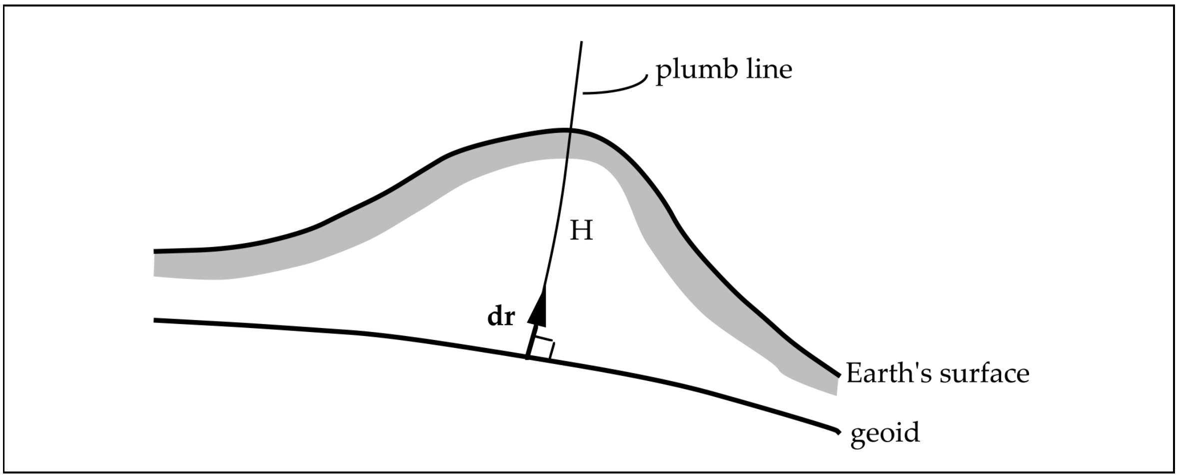


Figure 2.2: Definition of Orthometric Height.

Thus,

$$\mathrm{d}W = \boldsymbol{g} \cdot \mathrm{d}\boldsymbol{r} = |\boldsymbol{g}| \cdot |\,\mathrm{d}\boldsymbol{r}|\cos(\boldsymbol{g}, \mathrm{d}\boldsymbol{r}) = g\,\mathrm{d}H\cos(180°)$$

and therefore

$$\mathrm{d}W = -g\,\mathrm{d}H. \tag{2.4}$$

This is the fundamental equation for height definition. It is often rewritten in the form

$$\mathrm{d}H = -\frac{\mathrm{d}W}{g}. \tag{2.5}$$

Equation (2.5) defines the height in terms of potential differences and gravity. Since gravity cannot be measured inside the Earth, different approximations for g inside the Earth are made. Thus, different height systems are possible.

Integrating

$$H_\mathrm{P} = \int_0^\mathrm{P} \mathrm{d}H = -\int_0^\mathrm{P} \frac{1}{g}\,\mathrm{d}W = -\frac{1}{\tilde{g}}\int_0^\mathrm{P} \mathrm{d}W = -\frac{W_\mathrm{P} - W_0}{\tilde{g}} = \frac{W_0 - W_\mathrm{P}}{\tilde{g}} = \frac{C_\mathrm{P}}{\tilde{g}} \tag{2.6}$$

where W_0 gravity potential at the geoid, W_P gravity potential at P, C *geopotential number*. geopotenzielle Kote

(2.5) can be written as

$$\text{Height} = \frac{C}{\tilde{g}}, \tag{2.7}$$

where $\tilde{g}$ is the mean gravity between surface point P and its footprint at the geoid along the plumbline. The resulting height is called *orthometric height* which is the length of the curved plumbline between P and geoid.

The SI unit of geopotential numbers is $\frac{\mathrm{m}^2}{\mathrm{s}^2}$.

However, in literature sometimes the non-standard *geopotential units* (g. p. u.) are used, i. e. 1 g. p. u. = 1 kGal · m. Since globally $\tilde{g}$ is approximately 0.98 kGal, geopotential numbers in g. p. u. are always close to heights above the geoid in meters.

The system of natural coordinates can therefore be expressed by either one of the two coordinate triples

$$\Phi, \Lambda, H \qquad \text{or} \qquad \Phi, \Lambda, C$$

The first has a simple geometrical explanation but introduces some assumptions in the definition of H. The second is more precise in terms of defining the coordinates but lacks the intuitive geometrical meaning. In both cases, the coordinates can be uniquely expressed by potential differences and gradients of W, thus explaining Bruns's (1878) concise statement: "The task of geodesy is the determination of the potential function $W(x, y, z)$".

2.2 Astronomical latitude (Φ) and longitude (Λ)

In the introduction to this chapter, the system of natural coordinates has been described as a system whose axes are defined by directions which are physically meaningful in terrestrial space as e. g. the directions of the gravity vector and the *spin axis* of the Earth. How can these spatial directions be related to positions in an Earth-fixed coordinate system? It is obvious from the preceding sections that the gravity potential and its gradients are important in this context. They define the direction of the gravity vector by gradients of W in an Earth-fixed Cartesian $\{x, y, z\}$-system. The relationship between $\operatorname{grad} W$ and $\{x, y, z\}$ will be briefly discussed in this section.

Drehachse

The simplest representation of the gravity vector is obtained in the Local Astronomic System (LA) which is defined by:

Local Astronomic System (LA)	
Origin:	At point P
Primary axis (z):	Orthogonal to level surface at P (W_{P} = const.)
Secondary axis (x):	Tangent to astronomic meridian pointing north
Tertiary axis (y):	Orthogonal in a left-handed system

In this system, which is shown in Figure 2.3, the gravity vector has the coordinates

$$\boldsymbol{g}^g = \operatorname{grad} W = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix}.$$

To relate this representation to that in an Earth-fixed Cartesian system, let us first define such a system:

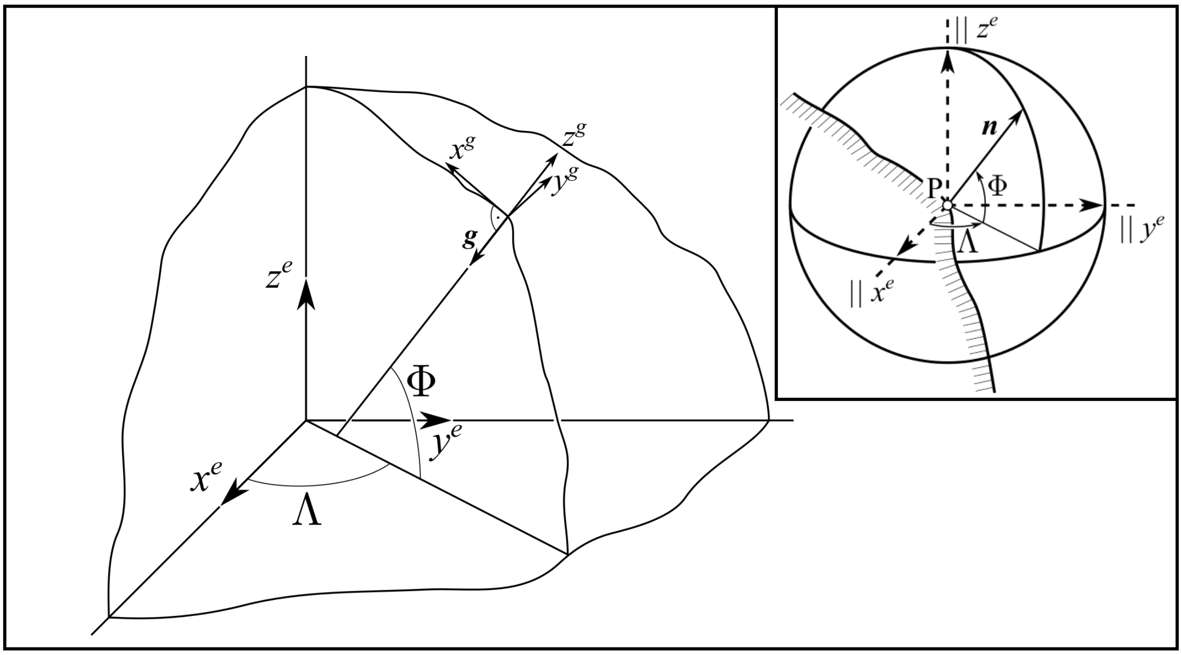


Figure 2.3: Local Astronomic System and Astronomic Coordinates (bottom left) and direction of normal vector in terms of Φ and Λ on the celestial sphere (top right).

Conventional Terrestrial System (CT)	
Origin:	Centre of mass of the Earth
Primary axis (z):	Conventional (or mean) spin axis of the Earth
Secondary axis (x):	Intersection of the conventional (or mean) equator plane and the mean meridian plane of Greenwich
Tertiary axis (y):	Orthogonal in a right-handed system

This coordinate system is of fundamental importance in geodesy. It will be more rigorously defined in Chapter 4, where the terms mean or conventional get a proper definition. The plane orthogonal to the conventional spin axis is called conventional *equator plane*. Äquatorebene

The direction of $\boldsymbol{g}$ in the CT-system is given by two angles, astronomic latitude Φ and astronomic longitude Λ. The definition of Λ is tied to the definition of the astronomic *meridian plane*. Meridianebene

The astronomic meridian plane of a point P is the plane containing the gravity vector at P and the parallel to the conventional rotation axis of the Earth through P. Thus, it is orthogonal to the conventional equator plane. The astronomic longitude Λ is the angle between the astronomic meridian planes of two points. The convention is that the angle Λ is counted counterclockwise from the mean astronomic meridian plane of Greenwich.

The astronomic latitude Φ of a point P is the smallest angle between the conventional equator plane and the vector normal to the level surface in P measured in the meridian plane of P. The normal vector is opposite in direction to the gravity vector in P. The angle Φ is conventionally counted from the mean equator plane positive towards the north pole and negative towards the south pole.

The normal vector $\boldsymbol{n}$ has the form

$$\boldsymbol{n}^e = \begin{pmatrix} \cos\Phi\cos\Lambda \\ \cos\Phi\sin\Lambda \\ \sin\Phi \end{pmatrix}. \tag{2.8}$$

Using this definition of the normal vector $\boldsymbol{n}$, the gravity vector $\boldsymbol{g}$ can be expressed as

$$\boldsymbol{g}^e = -g\,\boldsymbol{n} \tag{2.9}$$

and substituting (2.8) into (2.9), we get

$$\boldsymbol{g}^e = \operatorname{grad} W = \begin{pmatrix} \partial W/\partial x \\ \partial W/\partial y \\ \partial W/\partial z \end{pmatrix} = \begin{pmatrix} W_x \\ W_y \\ W_z \end{pmatrix} = -g \begin{pmatrix} \cos\Phi\cos\Lambda \\ \cos\Phi\sin\Lambda \\ \sin\Phi \end{pmatrix}. \tag{2.10}$$

Formula (2.10) shows an interesting connection between physics and geometry. Starting from physics (gravity potential), the astronomic coordinates $\{\Phi, \Lambda\}$ can be derived and the geometry of space (e. g. curvature of the Earth's surface) can be determined.

Formula (2.10) defines the gradients of the gravity potential in terms of astronomic coordinates Φ and Λ. The reverse formulas expressing Φ and Λ as gradients of the gravity potential, can also be obtained.

From (2.10) we have

$$W_x^2 + W_y^2 = g^2\cos^2\Phi(\cos^2\Lambda + \sin^2\Lambda) = g^2\cos^2\Phi$$

and

$$\frac{W_z}{\sqrt{W_x^2 + W_y^2}} = \frac{-g\sin\Phi}{g\cos\Phi} = -\tan\Phi.$$

Also

$$\frac{W_y}{W_x} = \tan\Lambda.$$

Thus,

$$\begin{aligned} \Phi &= \arctan\frac{-W_z}{\sqrt{W_x^2 + W_y^2}} \\ \Lambda &= \arctan\frac{W_y}{W_x} \end{aligned} \tag{2.11}$$

Formula (2.11) shows that, if the gravity potential $W(x, y, z)$ is given, the coordinates Φ and Λ can always be determined.

To describe the position of a point in three-dimensional space, three coordinates are needed. krummlinig They can be Cartesian $\{x, y, z\}$, *curvilinear* $\{j, l, h\}$, or some other coordinate triple. Φ and Λ give the position of a point on an equipotential surface. It makes sense, therefore, to define the third coordinate as being orthogonal to this surface. It has been mentioned before that this coordinate is called the orthometric height H if the reference surface is the geoid.

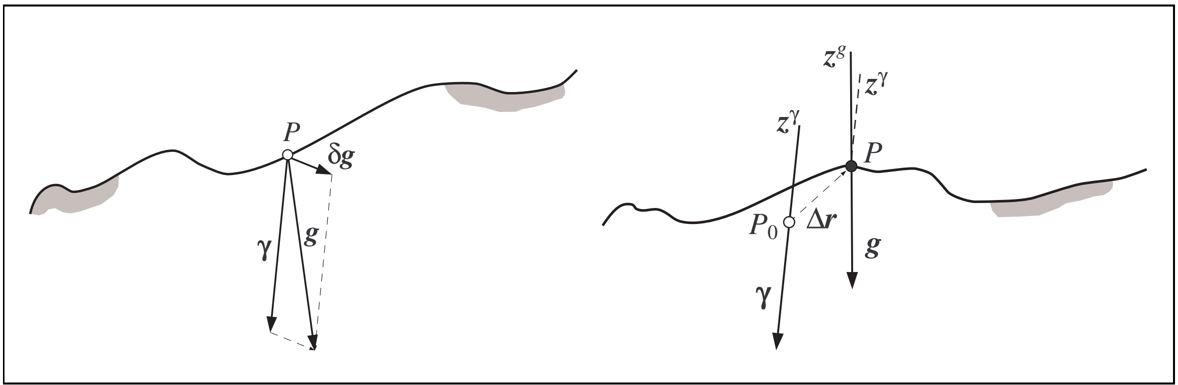


Figure 2.4: Definition of gravity disturbance (left) and gravity anomaly (right).

2.3 Gravity disturbance

The scalar *gravity disturbance* is defined as: Schwere-störung

$$\delta g = g(\boldsymbol{r}) - \gamma(\boldsymbol{r}) \quad \text{or} \quad \delta g = g_{\mathrm{P}} - \gamma_{\mathrm{P}} \,. \tag{2.12}$$

The vectorial version would be something like $\delta \boldsymbol{g}_{\mathrm{P}} = \boldsymbol{g}_{\mathrm{P}} - \boldsymbol{\gamma}_{\mathrm{P}}$, which is visualized in fig. 2.4. However, before doing the subtraction both vectors $\boldsymbol{g}$ and $\boldsymbol{\gamma}$ must be in the same coordinate system. This is usually not the case. Normal gravity is usually expressed in the local geodetic γ-frame, whereas $\boldsymbol{g}$ is usually expressed in the local astronomic g-frame. The coordinate system is indicated by a superindex.

$$\boldsymbol{g}^g = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \quad \text{and} \quad \boldsymbol{\gamma}^\gamma = \begin{pmatrix} 0 \\ 0 \\ -\gamma \end{pmatrix} .$$

Before subtraction one of these vectors should be expressed in the coordinate system of the other. This is achieved through the following detour over global coordinate systems:

$$\boldsymbol{r}^\gamma = S_1 R_2(\tfrac{1}{2}\pi - \varphi) R_3(\lambda) \boldsymbol{r}^\varepsilon \tag{2.13a}$$
$$\boldsymbol{r}^g = S_1 R_2(\tfrac{1}{2}\pi - \Phi) R_3(\Lambda) \boldsymbol{r}^e \tag{2.13b}$$
$$\boldsymbol{r}^\varepsilon = R_1(\varepsilon_1) R_2(\varepsilon_2) R_3(\varepsilon_3) \boldsymbol{r}^e \tag{2.13c}$$

in which the e-frame denotes the conventional terrestrial system and the ε-frame denotes the global geodetic one. The angles ε_i represent a orientation difference between these two global systems. It is assumed that their origins coincide. Combination of these transformations gives us the transformation between the two local frames:

$$\boldsymbol{r}^\gamma = S_1 R_2(\tfrac{1}{2}\pi - \varphi) R_3(\lambda) R_1(\varepsilon_1) R_2(\varepsilon_2) R_3(\varepsilon_3) R_3(-\Lambda) R_2(\Phi - \tfrac{1}{2}\pi) S_1 \boldsymbol{r}^g$$

- neglect ε_i
- lot of calculus
- small angle approximation

$$= \begin{pmatrix} 1 & (\Lambda - \lambda)\sin\varphi & (\Phi - \varphi) \\ -(\Lambda - \lambda)\sin\varphi & 1 & (\Lambda - \lambda)\cos\varphi \\ -(\Phi - \varphi) & -(\Lambda - \lambda)\cos\varphi & 1 \end{pmatrix} \boldsymbol{r}^g \tag{2.14a}$$

$$= \begin{pmatrix} 1 & \delta\Lambda\sin\varphi & \delta\Phi \\ -\delta\Lambda\sin\varphi & 1 & \delta\Lambda\cos\varphi \\ -\delta\Phi & -\delta\Lambda\cos\varphi & 1 \end{pmatrix} \boldsymbol{r}^g \tag{2.14b}$$

$$= \begin{pmatrix} 1 & \psi & \xi \\ -\psi & 1 & \eta \\ -\xi & -\eta & 1 \end{pmatrix} \boldsymbol{r}^g = R_1(\eta) R_2(-\xi) R_3(\psi)\, \boldsymbol{r}^g\,. \tag{2.14c}$$

Between (2.14a) and (2.14b) we made use of the following definitions:

$$\text{Latitude disturbance:} \qquad \delta\Phi = \Phi_{\mathrm{P}} - \varphi_{\mathrm{P}}\,, \tag{2.15a}$$

$$\text{Longitude disturbance:} \qquad \delta\Lambda = \Lambda_{\mathrm{P}} - \lambda_{\mathrm{P}}\,. \tag{2.15b}$$

The matrix element $\psi = \delta\Lambda\sin\varphi = \eta\tan\varphi$ is actually one of the contributions to the azimuth disturbance (see below). The step from (2.14b) to (2.14c) made use of the definition of the *deflection of the vertical*:

$$\text{Deflection in N-S:} \qquad \xi = \delta\Phi\,, \tag{2.16a}$$

$$\text{Deflection in E-W:} \qquad \eta = \delta\Lambda\cos\varphi\,. \tag{2.16b}$$

Remark 2.1 *Remind that the orientation difference between the two global frames e and ε, represented by the angular datum parameters ε_i, has been neglected. The corresponding full transformation would become somewhat more elaborate. The equations are still manageable, though, since ε_i are small angles.*

We know our gravity vector $\boldsymbol{g}$ in the g-frame. Using (2.14c) it is easily transformed into the γ-frame now:

$$\boldsymbol{g}^\gamma = -g \begin{pmatrix} \xi \\ \eta \\ 1 \end{pmatrix}\,.$$

Finally, we are able to subtract the normal gravity vector from the gravity vector to get the *vector gravity disturbance*, see also fig. 2.5:

$$\delta\boldsymbol{g}^\gamma = \boldsymbol{g}^\gamma - \boldsymbol{\gamma}^\gamma = \begin{pmatrix} -g\xi \\ -g\eta \\ -g \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ -\gamma \end{pmatrix} = \begin{pmatrix} -g\xi \\ -g\eta \\ -\delta g \end{pmatrix} = \begin{pmatrix} -\gamma\xi \\ -\gamma\eta \\ -\delta g \end{pmatrix}\,. \tag{2.17}$$

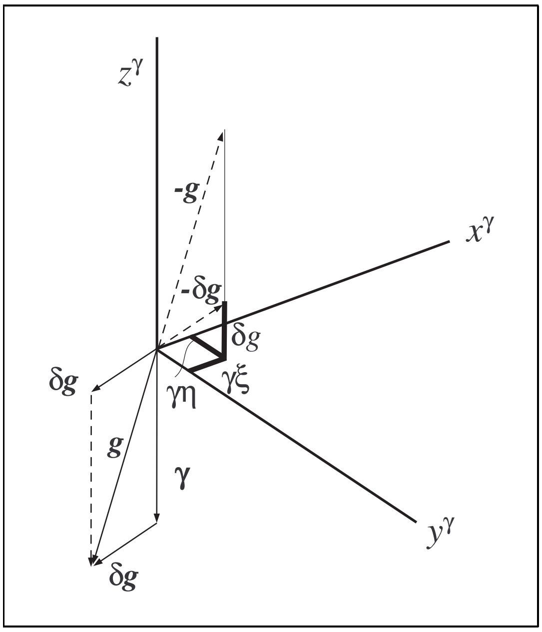


Figure 2.5: The gravity disturbance $\delta\boldsymbol{g}$ projected into the local geodetic frame is decomposed into deflections of the vertical $(\gamma\xi, \gamma\eta)$ and scalar gravity disturbance δg.

The latter change from g into γ is allowed because the deflection of the vertical is such a small quantity that the precision of the quantity by which it is multiplied doesn't matter.

The gravity vector is the gradient of the gravity potential. Correspondingly, the normal gravity vector is the gradient of the normal potential. Thus we have:

$$\delta\boldsymbol{g} = \nabla W - \nabla U = \nabla(W - U) = \nabla T\,,$$

i.e. the gravity disturbance vector is the gradient of the disturbing potential. We can write the gradient for instance in local Cartesian or in spherical coordinates. Written out in components:

$$\boxed{\text{Local Cartesian: } \begin{cases} \dfrac{\partial T}{\partial x} = -\gamma\xi \\ \dfrac{\partial T}{\partial y} = -\gamma\eta \\ \dfrac{\partial T}{\partial z} = -\delta g \end{cases} \qquad \text{Spherical: } \begin{cases} \dfrac{1}{r}\dfrac{\partial T}{\partial \varphi} = -\gamma\delta\Phi \\ \dfrac{1}{r\cos\varphi}\dfrac{\partial T}{\partial \lambda} = -\gamma\delta\Lambda\cos\varphi \\ \dfrac{\partial T}{\partial r} = -\delta g \end{cases}} \tag{2.18}$$

The RHS of each of these 6 equations represent the observable, the LHS the unknowns (derivatives of T).

Zenith and azimuth disturbances. Without going into too much detail the derivation of zenith and azimuth disturbances is straightforward. We can write any position vector $\boldsymbol{r}^\gamma$ in geodetic azimuth (α) and zenith (z). Similarly, any $\boldsymbol{r}^g$ can be written in astronomic azimuth (A) and zenith (Z). The transformation between both is known from (2.14c):

$$\begin{pmatrix} \sin z\cos\alpha \\ \sin z\sin\alpha \\ \cos z \end{pmatrix} = \begin{pmatrix} 1 & \psi & \xi \\ -\psi & 1 & \eta \\ -\xi & -\eta & 1 \end{pmatrix} \begin{pmatrix} \sin Z\cos A \\ \sin Z\sin A \\ \cos Z \end{pmatrix}.$$

After some manipulation one can derive:

$$\text{Azimuth disturbance:} \qquad \delta A = A_{\mathrm{P}} - \alpha_{\mathrm{P}} = \psi + \cot z(\xi \sin\alpha - \eta\cos\alpha)\,, \tag{2.19a}$$

$$\text{Zenith disturbance:} \qquad \delta Z = Z_{\mathrm{P}} - z_{\mathrm{P}} = -\xi\cos\alpha - \eta\sin\alpha\,. \tag{2.19b}$$

Recommended additional reading

- Torge (1991): Chapter 2,
- Vanicek and Krakiwsky (1982): Chapter 6,
- Heiskanen and Moritz (1967): Chapters 1.1, 1.2, 2.1–2.4.

Test your knowledge

i) How would you build a sensor to measure gravity on the surface of the Earth?

ii) A plumbline is defined as the line intersecting all level surfaces orthogonally. If you determine astronomic coordinates at the geoid and at a point at height H above the geoid measured along the plumbline, will the astronomic coordinates be the same?

iii) If the level surfaces of the Earth were all parallel, what could you say about the figure of the Earth, its density, and its angular rotation.

iv) The gravity potential W for a small region of the Earth is approximated by

$$W(x,y,z) = \frac{GM}{R}\left\{1 - a_1\left(\frac{z}{R}\right)^2 + a_2 - a_3\frac{xy}{R^2}\right\} + \frac{1}{2}\omega^2(x^2+y^2).$$

Determine the astronomic coordinates of the point with coordinates

$$x = 2800\,\mathrm{km}, \quad y = 2900\,\mathrm{km}, \quad z = 5000\,\mathrm{km}$$

if

$$a_1 = 1.6{\cdot}10^{-3}, \quad a_2 = 0.65{\cdot}10^{-3}, \quad a_3 = 0.6{\cdot}10^{-5}, \quad R = 6371\,\mathrm{km}, \quad \omega = 0.7292{\cdot}10^{-4}\,\mathrm{s}^{-1}.$$

v) The system of natural coordinates can either be expressed by the coordinates $\{\Phi, \Lambda, H\}$ or by $\{\Phi, \Lambda, C\}$. Explain the difference between them and discuss situations where you would use one or the other.

vi) Which advantages do you see in using a CT-system instead of an LA-system?

3 Rotation

kinematics Gravity related measurements take generally place on non-static platforms: sea-gravimetry, airborne gravimetry, satellite gravity gradiometry, inertial navigation. Even measurements on a fixed point on Earth belong to this category because of the Earth's rotation. Accelerated motion of the reference frame induces inertial accelerations, which must be taken into account in physical geodesy. The rotation of the Earth causes a centrifugal acceleration which is combined with the gravitational attraction into a new quantity: *gravity*. Other inertial accelerations are usually accounted for by correcting the gravity related measurements, e.g. the Eötvös correction. For these and other purposes we will start this chapter by investigating velocity and acceleration in a rotating frame.

dynamics One of geodesy's core areas is determining the orientation of Earth in space. This goes to the heart of the transformation between inertial and Earth-fixed reference systems. The solar and lunar gravitational fields exert a torque on the flattened Earth, resulting in changes of the polar axis. We need to elaborate on the dynamics of solid body rotation to understand how the polar axis behaves in inertial and in Earth-fixed space.

geometry Newton's laws of motion are valid in inertial space. If we have to deal with satellite techniques, for instance, the satellite's ephemeris is most probably given in inertial coordinates. Star coordinates are by default given in inertial coordinates: right ascension α and declination δ. Moreover, the law of gravitation is defined in inertial space. Therefore, after understanding the kinematics and dynamics of rotation, we will discuss the definition of inertial reference systems and their realizations. An overview will be presented relating the conventional inertial reference system to the conventional terrestrial one.

3.1 Geometry: rotation basics

properties of rotation matrices Rotations are performed by rotation matrices R_i. Their dimension is 2×2 in two-dimensional and 3×3 in three-dimensional coordinate systems. The lower index defines the axis around which the rotation is carried out, e.g. $R_1(\alpha)$ is a rotation around the x-axis by the angle α. The matrices have certain properties.

i) Rotated vectors are invariant, i.e. their length doesn't change.

ii) Rotations are not commutative:

$$R_i(\alpha)R_j(\beta) \neq R_j(\beta)R_i(\alpha)\,.$$

iii) Rotations are associative:

$$R_i(\alpha)[R_j(\beta)R_k(\gamma)] = [R_i(\alpha)R_j(\beta)]R_k(\gamma)\,.$$

iv) Rotations around the same axis are additive:

$$R_i(\alpha)R_i(\beta) = R_i(\alpha+\beta)\,.$$

v) The inverse equals the transposed matrix:

$$R^{\mathsf{T}} = R^{-1} \Longleftrightarrow R^{\mathsf{T}}R = I\,.$$

passive rotation in 2D A passive rotation is a rotation of the coordinate system as shown in fig. 3.1 on the left side. The coordinates in the new system (x',y') have to be computed by a multiplication of a rotation matrix to the given points $(1,0)^{\mathsf{T}}$ and $(0,1)^{\mathsf{T}}$ on the axes.

The new coordinates are known by trigonometry, which gives one column of the unknown rotation matrix per point.

$$\boldsymbol{r}' = R(\alpha)\boldsymbol{r} \tag{3.1}$$

$$\text{Point 1}: \begin{pmatrix} \cos\alpha \\ -\sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & \dots \\ -\sin\alpha & \dots \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\text{Point 2}: \begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix} = \begin{pmatrix} \dots & \sin\alpha \\ \dots & \cos\alpha \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Downarrow$$

$$R(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{3.2}$$

Remark 3.1 *The direction of rotation is mathematically positive, i.e. counterclockwise, because the first axis (x) is turned towards the second axis (y).*

active rotation in 2D Active rotation doesn't change the coordinate system but actually moves the point itself. The rotation of the two points from above in the same coordinate system leads to different new points (cf. fig. 3.1, right).

$$\text{Point 1}: \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & \dots \\ \sin\alpha & \dots \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\text{Point 2}: \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix} = \begin{pmatrix} \dots & -\sin\alpha \\ \dots & \cos\alpha \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Downarrow$$

$$R(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{3.3}$$

The difference between active and passive rotation is the position of the negative sign. Because we are dealing with different coordinate systems, we use passive rotation.

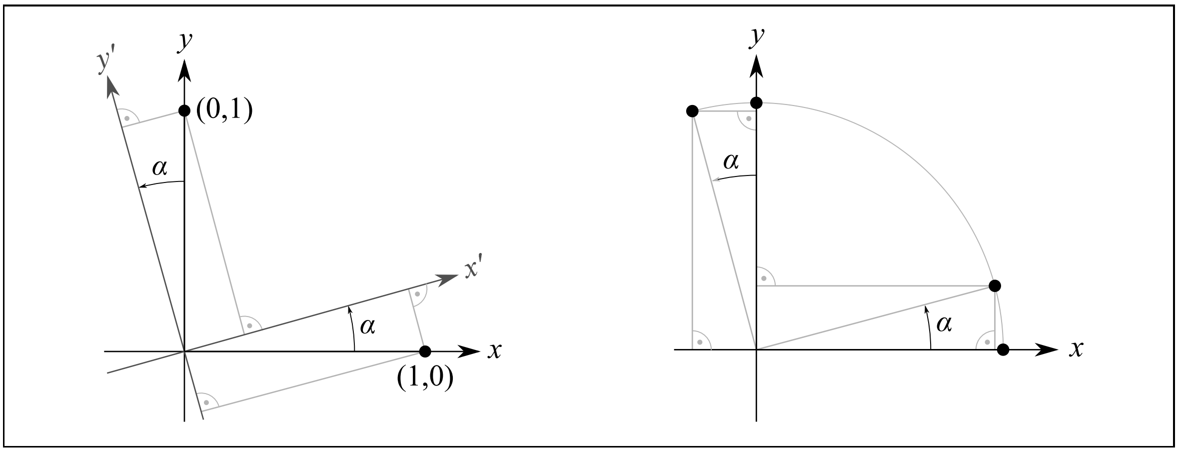


Figure 3.1: Passive rotation of the coordinate system (left) and active rotation of a point (right) in two-dimensional space.

Remark 3.2 *Active and passive rotation are linked by*

$$R_{\text{passive}} = R_{\text{active}}^{\mathsf{T}} = R_{\text{active}}^{-1} .$$

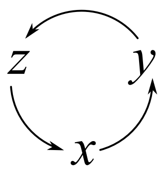
passive rotation in 3D Rotation around the first axis in three-dimensional space is a passive two-dimensional rotation of the (y,z)-plane. The two-dimensional matrix is extended by a row and a column fixing the x-axis.

$$R_1(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix} \tag{3.4}$$

Rotating around the z-axis follows the same scheme.

$$R_3(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.5}$$

Remark 3.3 *The rotation direction stays counterclockwise, because also the y-axis moves towards the z-axis. In the next step, z will move in direction of x because the positive rotation can be defined by* cyclic permutation*:* Zyklische Permutation



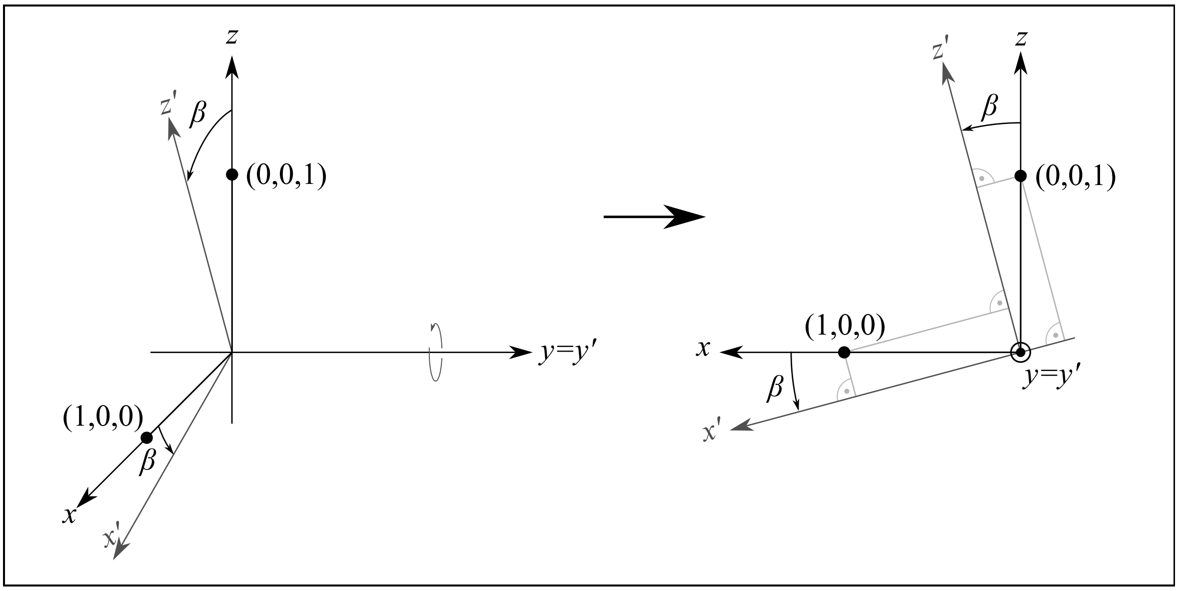


Figure 3.2: Three-dimensional rotation around y-axis simplified by directly looking on the (x,z)-plane.

The matrix looks different for the rotation axis y. With fig. 3.2, the matrix can be built like the two-dimensional ones looking on the (x,z)-plane.

$$R_2(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \tag{3.6}$$

Euler rotation Every rotation can be represented by a series of three rotations around one of the three axes each. The Euler[1] rotations use the rotation around one axis (mostly R_3) twice (cf. fig. 3.3).

$$R_{\text{Euler}}(\alpha, \beta, \gamma) = R_3(\gamma) R_1(\beta) R_3(\alpha) \tag{3.7}$$

Cardan rotation Another often used representation in praxis is the Cardan[2] rotation, which consists of three rotations around three different axes. They are also called Tait[3]-Bryan[4] rotations. There are two main applications:

i) Cardan rotation in vehicle and aircraft engineering

$$R_{\text{Cardan}}(\Psi, \Theta, \Phi) = R_1(\Phi) R_2(\Theta) R_3(\Psi)\,. \tag{3.8}$$

The angles are

[1]Leonhard Euler (1707–1783), Swiss mathematician and physicist.
[2]Gerolamo Cardano (1501–1576), Italian mathematician.
[3]Peter Guthrie Tait (1831–1901), Scottish mathematical physicist.
[4]George H. Bryan (1864–1928), British physicist.

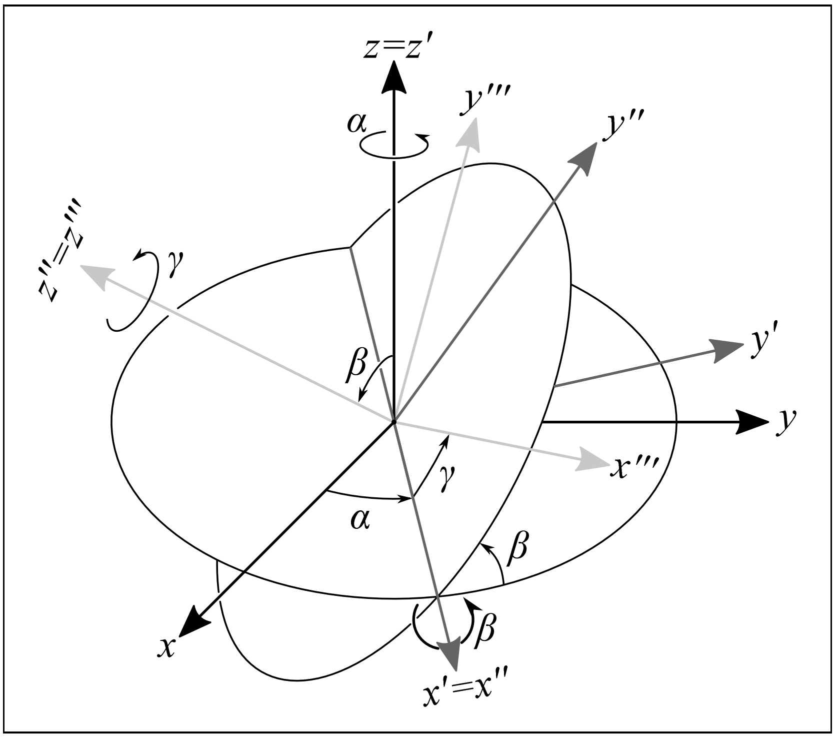


Figure 3.3: Euler rotation sequence $R_3(\gamma)R_1(\beta)R_3(\alpha)$.

- *yaw*, *heading* or azimuth Ψ Gier Steuerkurs
- *pitch* Θ Nick
- *roll* Φ . Roll

ii) Cardan rotation for the exterior orientation in photogrammetry

$$R_{\text{Cardan}}(\omega, \varphi, \kappa) = R_3(\kappa)R_2(\varphi)R_1(\omega)\,. \tag{3.9}$$

Here the order is first roll angle ω, then pitch φ and heading κ.

differential rotation The angles used for transformations between different reference systems are often small. This allows the simplification of the trigonometric functions $\sin\alpha \approx \alpha$ and $\cos\alpha \approx 1$ for small angles $\alpha \to 0$. As a consequence, the rotation matrices can be written differently.

$$R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix} \overset{\alpha\to 0}{=} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & -\alpha & 1 \end{pmatrix} = I + \alpha L_1,\; L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \tag{3.10a}$$

$$R_2 = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix} \overset{\beta\to 0}{=} \begin{pmatrix} 1 & 0 & -\beta \\ 0 & 1 & 0 \\ \beta & 0 & 1 \end{pmatrix} = I + \beta L_2,\; L_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.10b}$$

$$R_3 = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{\gamma\to 0}{=} \begin{pmatrix} 1 & \gamma & 0 \\ -\gamma & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I + \gamma L_3,\; L_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \tag{3.10c}$$

Generator-matrizen

The matrices L_1, L_2, L_3 are so-called *generator matrices*. They allow the separation of the angle and the matrix, which is especially useful when we apply Euler or Cardan rotations for differential angles.

$$\begin{aligned} R_{\text{Euler}}(\alpha,\beta,\gamma) &= (I+\gamma L_3)(I+\beta L_2)(I+\alpha L_3) \\ &= I+\gamma L_3+\beta L_1+\alpha L_3+\underbrace{\alpha\gamma L_3 L_3+\beta\gamma L_2 L_3+\alpha\beta L_2 L_3}_{=0} \\ &= I+\gamma L_3+\beta L_1+\alpha L_3 \\ &= \begin{pmatrix} 1 & \alpha+\gamma & -\beta \\ -(\alpha+\gamma) & 1 & 0 \\ \beta & 0 & 1 \end{pmatrix}. \end{aligned} \tag{3.11}$$

If written in the same additive form, the Cardan rotation for differential angles looks more intuitive and is commonly used in coordinate transformations.

$$\begin{aligned} R_{\text{Cardan}}(\alpha,\beta,\gamma) &= (I+\gamma L_3)(I+\beta L_2)(I+\alpha L_1) = I+\alpha L_1+\beta L_2+\gamma L_3 \\ &= \begin{pmatrix} 1 & \gamma & -\beta \\ -\gamma & 1 & \alpha \\ \beta & -\alpha & 1 \end{pmatrix}. \end{aligned} \tag{3.12}$$

3.2 Kinematics: acceleration in a rotating frame

Let us consider the situation of motion in a *rotating* reference frame and let us associate this rotating frame with the Earth-fixed frame. The following discussion on velocities and accelerations would be valid for any rotating frame, though.

Inertial coordinates, velocities and accelerations will be denoted with the index i. Earth-fixed quantities get the index e. Now suppose that a *time-dependent* rotation matrix $R = R(\alpha(t))$, applied to the inertial vector $\boldsymbol{r}^i$, results in the Earth-fixed vector $\boldsymbol{r}^e$. We would be interested in velocities and accelerations in the rotating frame. The time derivations must be performed in the inertial frame, though.

From $R\boldsymbol{r}^i = \boldsymbol{r}^e$ we get:

$$\boldsymbol{r}^i = R^{\mathsf{T}}\boldsymbol{r}^e \tag{3.13a}$$

$$\Downarrow \text{ time derivative}$$

$$\dot{\boldsymbol{r}}^i = R^{\mathsf{T}}\dot{\boldsymbol{r}}^e + \dot{R}^{\mathsf{T}}\boldsymbol{r}^e \tag{3.13b}$$

$$\Downarrow \text{ multiply by } R$$

$$\begin{aligned} R\dot{\boldsymbol{r}}^i &= \dot{\boldsymbol{r}}^e + R\dot{R}^{\mathsf{T}}\boldsymbol{r}^e \\ &= \dot{\boldsymbol{r}}^e + \Omega\boldsymbol{r}^e \end{aligned} \tag{3.13c}$$

The matrix $\Omega = R\dot{R}^{\mathsf{T}}$ is called *Cartan*[5] *matrix.* It describes the rotation rate, as can be seen from the following simple 2D example with $\alpha(t) = \omega t$:

$$
\begin{aligned}
R &= \begin{pmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{pmatrix} \\
\Rightarrow \quad \Omega &= \begin{pmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{pmatrix} \omega \begin{pmatrix} -\sin\omega t & -\cos\omega t \\ \cos\omega t & -\sin\omega t \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}
\end{aligned}
$$

It is useful to introduce Ω. In the next time differentiation step we can now distinguish between time dependent rotation matrices and time variable rotation rate. Let's pick up the previous derivation again:

$$
\begin{aligned}
&\Downarrow \text{ multiply by } R^{\mathsf{T}} \\
\dot{\boldsymbol{r}}^i &= R^{\mathsf{T}}\dot{\boldsymbol{r}}^e + R^{\mathsf{T}}\Omega\boldsymbol{r}^e && (3.13\text{d}) \\
&\Downarrow \text{ time derivative} \\
\ddot{\boldsymbol{r}}^i &= R^{\mathsf{T}}\ddot{\boldsymbol{r}}^e + \dot{R}^{\mathsf{T}}\dot{\boldsymbol{r}}^e + \dot{R}^{\mathsf{T}}\Omega\boldsymbol{r}^e + R^{\mathsf{T}}\dot{\Omega}\boldsymbol{r}^e + R^{\mathsf{T}}\Omega\dot{\boldsymbol{r}}^e \\
&= R^{\mathsf{T}}\ddot{\boldsymbol{r}}^e + 2\dot{R}^{\mathsf{T}}\dot{\boldsymbol{r}}^e + \dot{R}^{\mathsf{T}}\Omega\boldsymbol{r}^e + R^{\mathsf{T}}\dot{\Omega}\boldsymbol{r}^e && (3.13\text{e}) \\
&\Downarrow \text{ multiply by } R \\
R\ddot{\boldsymbol{r}}^i &= \ddot{\boldsymbol{r}}^e + 2\Omega\dot{\boldsymbol{r}}^e + \Omega\Omega\boldsymbol{r}^e + \dot{\Omega}\boldsymbol{r}^e \\
&\Downarrow \text{ or the other way around} \\
\ddot{\boldsymbol{r}}^e &= R\ddot{\boldsymbol{r}}^i - 2\Omega\dot{\boldsymbol{r}}^e - \Omega\Omega\boldsymbol{r}^e - \dot{\Omega}\boldsymbol{r}^e && (3.13\text{f})
\end{aligned}
$$

This equation tells us that acceleration in the rotating e-frame equals acceleration in the inertial i-frame—in the proper orientation, though—when 3 more terms are added. The additional terms are called *inertial* accelerations. Analyzing (3.13f) we can distinguish the four terms at the right hand side:

i) $R\ddot{\boldsymbol{r}}^i$ is the inertial acceleration vector, expressed in the orientation of the rotating frame.

ii) $2\Omega\dot{\boldsymbol{r}}^e$ is the so-called *Coriolis*[6] acceleration, which is due to motion in the rotating frame.

iii) $\Omega\Omega\boldsymbol{r}^e$ is the *centrifugal* acceleration, determined by the position in the rotating frame.

iv) $\dot{\Omega}\boldsymbol{r}^e$ is sometimes referred to as *Euler*[7] acceleration or inertial acceleration of rotation. It is due to a non-constant rotation rate.

Remark 3.4 *Equation (3.13f) can be generalized to moving frames with time-variable origin. If the linear acceleration of the e-frame's origin is expressed in the i-frame with* $\ddot{\boldsymbol{b}}^i$, *the only change to be made to (3.13f) is* $R\ddot{\boldsymbol{r}}^i \to R(\ddot{\boldsymbol{r}}^i - \ddot{\boldsymbol{b}}^i)$.

[5] Élie Joseph Cartan (1869–1951), French mathematician.
[6] Gaspard Gustave de Coriolis (1792–1843).
[7] Leonhard Euler (1707–1783).

Properties of the Cartan matrix $\boldsymbol{\Omega}$. Cartan matrices are skew-symmetric, i.e. $\Omega^\mathsf{T} = -\Omega$. This can be seen in the simple 2D example above already. But it also follows from the orthogonality of rotation matrices:

$$RR^\mathsf{T} = I \implies \frac{\mathrm{d}}{\mathrm{d}t}(RR^\mathsf{T}) = \underbrace{\dot{R}R^\mathsf{T}}_{\Omega^\mathsf{T}} + \underbrace{R\dot{R}^\mathsf{T}}_{\Omega} = 0 \implies \Omega^\mathsf{T} = -\Omega\,. \tag{3.14}$$

A second interesting property is the fact that multiplication of a vector with the Cartan matrix equals the cross product of the vector with a corresponding rotation vector:

$$\Omega \boldsymbol{r} = \boldsymbol{\omega} \times \boldsymbol{r} \tag{3.15}$$

This property becomes clear from writing out the 3 Cartan matrices, corresponding to the three independent rotation matrices:

$$\left.\begin{aligned} R_1(\omega_1 t) \;&\Rightarrow\; \Omega_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\omega_1 \\ 0 & \omega_1 & 0 \end{pmatrix} \\ R_2(\omega_2 t) \;&\Rightarrow\; \Omega_2 = \begin{pmatrix} 0 & 0 & \omega_2 \\ 0 & 0 & 0 \\ -\omega_2 & 0 & 0 \end{pmatrix} \\ R_3(\omega_3 t) \;&\Rightarrow\; \Omega_3 = \begin{pmatrix} 0 & -\omega_3 & 0 \\ \omega_3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}\right\} \overset{\text{general}}{\Longrightarrow} \quad \Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \tag{3.16}$$

Indeed, when a general rotation vector $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^\mathsf{T}$ is defined, we see that:

$$\begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The skew-symmetry (3.14) of Ω is related to the fact $\boldsymbol{\omega} \times \boldsymbol{r} = -\boldsymbol{r} \times \boldsymbol{\omega}$.

Exercise 3.1 *Convince yourself that the above Cartan matrices Ω_i are correct, by doing the derivation yourself. Also verify (3.15) by writing out* LHS *and* RHS.

Using property (3.15), the velocity (3.13c) and acceleration (3.13f) may be recast into the perhaps more familiar form:

$$\dot{\boldsymbol{r}}^e = R\dot{\boldsymbol{r}}^i - \boldsymbol{\omega} \times \boldsymbol{r}^e \tag{3.17a}$$

$$\ddot{\boldsymbol{r}}^e = R\ddot{\boldsymbol{r}}^i - 2\boldsymbol{\omega} \times \dot{\boldsymbol{r}}^e - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}^e) - \dot{\boldsymbol{\omega}} \times \boldsymbol{r}^e \tag{3.17b}$$

Inertial acceleration due to Earth rotation

Neglecting precession, nutation and polar motion, the transformation from inertial to Earth-fixed frame is given by:

$$\boldsymbol{r}^e = R_3(\text{GAST})\boldsymbol{r}^i \;\overset{\text{or}}{\to}\; \boldsymbol{r}^e = R_3(\omega t)\boldsymbol{r}^i\,. \tag{3.18}$$

The latter is allowed here, since we are only interested in the acceleration effects, due to the rotation. We are not interested in the rotation of position vectors. With great precision, one can say that the Earth's rotation rate is constant: $\dot{\omega} = 0$ The corresponding Cartan matrix and its time derivative read:

$$\Omega = \begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \dot{\Omega} = 0\,.$$

The three inertial accelerations, due to the rotation of the Earth, become:

$$\text{Coriolis:} \qquad -2\Omega\dot{\boldsymbol{r}}^e = 2\omega \begin{pmatrix} \dot{y}^e \\ -\dot{x}^e \\ 0 \end{pmatrix} \tag{3.19a}$$

$$\text{centrifugal:} \quad -\Omega\Omega\boldsymbol{r}^e = \omega^2 \begin{pmatrix} x^e \\ y^e \\ 0 \end{pmatrix} \tag{3.19b}$$

$$\text{Euler:} \qquad -\dot{\Omega}\boldsymbol{r}^e = \boldsymbol{0} \tag{3.19c}$$

The Coriolis acceleration is perpendicular to both the velocity vector and the Earth's rotation axis. The centrifugal acceleration is perpendicular to the rotation axis and is parallel to the equator plane.

Exercise 3.2 *Determine the direction and the magnitude of the Coriolis acceleration if you are driving from Calgary to Banff with* 100 km/h.

Exercise 3.3 *How large is the centrifugal acceleration in Calgary? On the equator? At the North Pole? And in which direction?*

3.3 Dynamics: precession, nutation, polar motion

Instead of *linear velocity* (or *momentum*) and *forces* we will have to deal with *angular momentum* and *torques*. Starting with the basic definition of angular momentum of a point mass, we will step by step arrive at the angular momentum of solid bodies and their tensor of inertia. In the following all vectors are assumed to be given in an inertial frame, unless otherwise indicated.

Drehimpuls
Drehmoment

Angular momentum of a point mass The basic definition of angular momentum of a point mass is the cross product of position and velocity: $\boldsymbol{L} = m\boldsymbol{r} \times \boldsymbol{v}$. It is a vector quantity. Due to the definition the direction of the angular momentum is perpendicular to both $\boldsymbol{r}$ and $\boldsymbol{v}$.

In our case, the only motion $\boldsymbol{v}$ that exists is due to the rotation of the point mass. By substituting $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$ we get:

$$\begin{aligned}
\boldsymbol{L} &= m\boldsymbol{r} \times (\boldsymbol{\omega} \times \boldsymbol{r}) \qquad (3.20a)\\
&= m \begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \left[\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right] \\
&= m \begin{pmatrix} \omega_1 y^2 - \omega_2 xy - \omega_3 xz + \omega_1 z^2 \\ \omega_2 z^2 - \omega_3 yz - \omega_1 yx + \omega_2 x^2 \\ \omega_3 x^2 - \omega_1 zx - \omega_2 zy + \omega_3 y^2 \end{pmatrix} \\
&= m \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\
&= \boldsymbol{M}\boldsymbol{\omega} \, . \qquad (3.20b)
\end{aligned}$$

Massen-trägheitsmoment

The matrix $\boldsymbol{M}$ is called the *tensor of inertia*. It has units of $[\mathrm{kg\,m^2}]$. Since $\boldsymbol{M}$ is not an ordinary matrix, but a tensor, which has certain transformation properties, we will indicate it by boldface math type, just like vectors.

Compare now the angular momentum equation $\boldsymbol{L} = \boldsymbol{M}\boldsymbol{\omega}$ with the linear momentum equation $\boldsymbol{p} = m\boldsymbol{v}$, see also tbl. 3.1. It may be useful to think of m as a *mass scalar* and of $\boldsymbol{M}$ as a *mass matrix*. Since the mass m is simply a scalar, the linear momentum $\boldsymbol{p}$ will always be in the same direction as the velocity vector $\boldsymbol{v}$. The angular momentum $\boldsymbol{L}$, though, will generally be in a different direction than $\boldsymbol{\omega}$, depending on the matrix $\boldsymbol{M}$.

Exercise 3.4 *Consider yourself a point mass and compute your angular momentum, due to the Earth's rotation, in two ways:*

i) straightforward by (3.20a), and

ii) by calculating your tensor of inertia first and then applying (3.20b).

Is $\boldsymbol{L}$ parallel to $\boldsymbol{\omega}$ in this case?

Angular momentum of systems of point masses The concept of tensor of inertia is easily generalized to systems of point masses. The total tensor of inertia is just the superposition of the individual tensors. The total angular momentum reads:

$$\boldsymbol{L} = \sum_{n=1}^{N} m_n \boldsymbol{r}_n \times \boldsymbol{v}_n = \sum_{n=1}^{N} \boldsymbol{M}_n \boldsymbol{\omega} \, . \qquad (3.21)$$

Angular momentum of a solid body We will now make the transition from a discrete to a continuous mass distribution. Symbolically:

$$\lim_{N\to\infty} \sum_{n=1}^{N} m_n \cdots = \iiint_{\Omega} \ldots \, \mathrm{d}m \, .$$

Again, the angular momentum reads $\boldsymbol{L} = \boldsymbol{M}\boldsymbol{\omega}$. For a solid body, the tensor of inertia $\boldsymbol{M}$ is defined as:

$$\begin{aligned}
\boldsymbol{M} &= \iiint\limits_{\Omega} \begin{pmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{pmatrix} \mathrm{d}m \\
&= \begin{pmatrix} \iiint (y^2+z^2)\,\mathrm{d}m & -\iiint xy\,\mathrm{d}m & -\iiint xz\,\mathrm{d}m \\ & \iiint (x^2+z^2)\,\mathrm{d}m & -\iiint yz\,\mathrm{d}m \\ \text{symmetric} & & \iiint (x^2+y^2)\,\mathrm{d}m \end{pmatrix} .
\end{aligned}$$

The diagonal elements of this matrix are called *moments of inertia*. The off-diagonal terms are known as *products of inertia*.

Exercise 3.5 *Show that in vector-matrix notation the tensor of inertia* $\boldsymbol{M}$ *can be written as:* $\boldsymbol{M} = \iiint (\boldsymbol{r}^\mathsf{T}\boldsymbol{r}I - \boldsymbol{r}\boldsymbol{r}^\mathsf{T})\,\mathrm{d}m$ *.*

Torque If no external torques are applied to the rotating body, angular momentum is conserved. A change in angular momentum can only be effected by applying a torque $\boldsymbol{T}$:

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{T} = \boldsymbol{r} \times \boldsymbol{F} . \tag{3.22}$$

Equation (3.22) is the rotational equivalent of $\dot{\boldsymbol{p}} = \boldsymbol{F}$, see tbl. 3.1. Because of the cross-product, the change in the angular momentum vector is always perpendicular to both $\boldsymbol{r}$ and $\boldsymbol{F}$. Try to intuitively change the axis orientation of a spinning wheel by applying a force to the axis and the axis will probably go a different way. If no torques are applied ($\boldsymbol{T} = 0$) the angular momentum will be constant, indeed.

Table 3.1: Comparison between linear and rotational dynamics

	linear	**rotational**		
		point mass	solid body	
linear momentum	$\boldsymbol{p} = m\boldsymbol{v}$	$\boldsymbol{L} = m\boldsymbol{r} \times \boldsymbol{v}$	$\boldsymbol{L} = \boldsymbol{M}\boldsymbol{\omega}$	angular momentum
force	$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{F}$	$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{r} \times \boldsymbol{F}$	$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = \boldsymbol{T}$	torque

Three cases will be distinguished in the following:

i) $\boldsymbol{T}$ is constant $\longrightarrow$ precession,
which is a secular motion of the angular momentum vector in inertial space,

ii) $\boldsymbol{T}$ is periodic $\longrightarrow$ nutation (or forced nutation),
which is a periodic motion of $\boldsymbol{L}$ in inertial space,

iii) $\boldsymbol{T}$ is zero $\longrightarrow$ free nutation, polar motion,
which is a motion of the rotation axis in Earth-fixed space.

Precession The word precession is related to the verb to precede, indicating a steady, secular motion. In general, precession is caused by constant external torques. In the case of the Earth, precession is caused by the constant gravitational torques from Sun and Moon. The Sun's (or Moon's) gravitational pull on the nearest side of the Earth is stronger than the pull on the other side. At the same time the Earth is flattened. Therefore, if the Sun or Moon is not in the equatorial plane, a torque will be produced by the difference in gravitational pull on the equatorial bulges. Note that the Sun is only twice a year in the equatorial plane, namely during the equinoxes (beginning of Spring and Fall). The Moon goes twice a month through the equator plane.

Thus, the torque is produced because of three simultaneous facts:

i) the Earth is not a sphere, but rather an ellipsoid,

ii) the equator plane is tilted with respect to the ecliptic by $23^\circ\!.5$ (the obliquity ε) and also tilted with respect to the lunar orbit,

iii) the Earth is a spinning body.

If any of these conditions were absent, no torque would be generated by solar or lunar gravitation and precession would not take place.

As a result of the constant (or mean) part of the lunar and solar torques, the angular momentum vector will describe a conical motion around the northern ecliptical pole (NEP) with a radius of ε, see fig. 3.5 and fig. 3.6. The northern celestial pole (NCP) slowly moves over an ecliptical latitude circle. It takes the angular momentum vector 25 765 years to complete one revolution around the NEP. That corresponds to $50''\!.3$ per year.

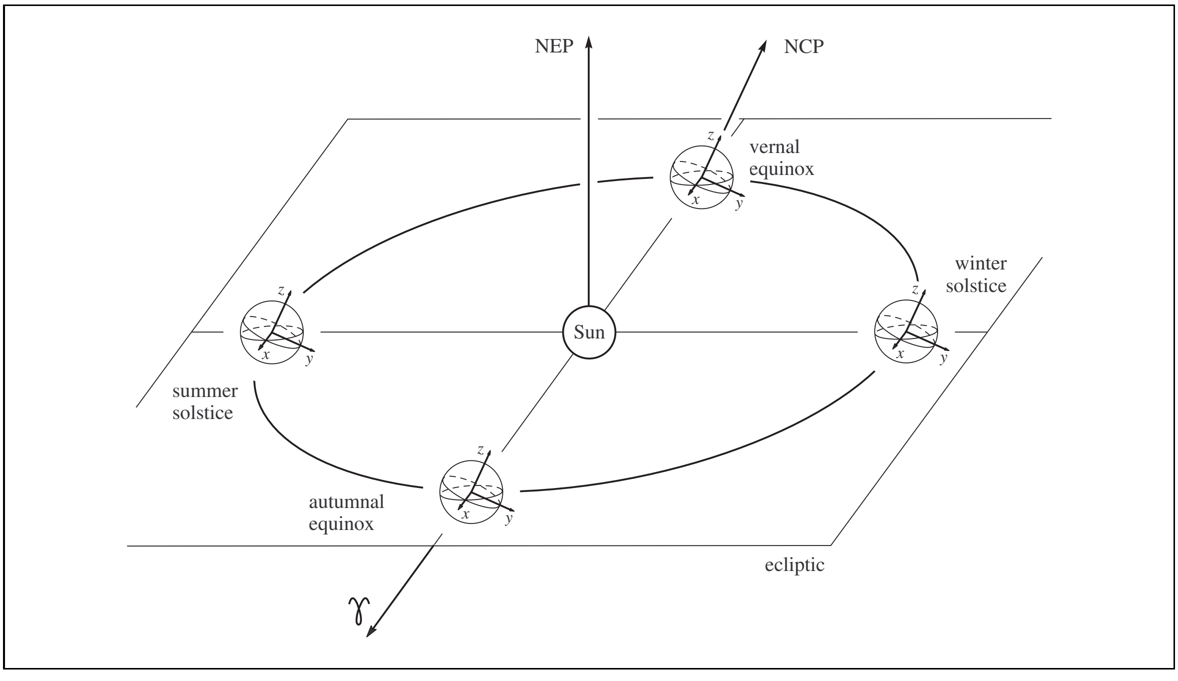


Figure 3.4: The Earth's rotation around the sun which causes the seasons and the precession effect.

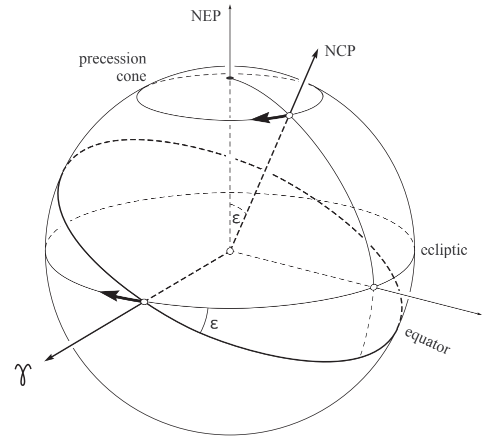


Figure 3.5: Obliquity ε with regard to the ecliptic.

Nutation The word nutation is derived from the Latin for *to nod*. Nutation is a periodic (nodding) motion of the angular momentum vector in space on top of the secular precession. There are many sources of periodic torques, each with its own frequency:

i) The orbital plane of the moon rotates once every 18.6 years under the influence of the Earth's flattening. The corresponding change in geometry causes also a change in the lunar gravitational torque of the same period. This effect is known as Bradley nutation.

ii) The sun goes through the equatorial plane twice a year, during the equinoxes. At those time the solar torque is zero. Vice versa, during the two solstices, the torque is maximum. Thus there will be a semi-annual nutation.

iii) The orbit of the Earth around the Sun is elliptical. The gravitational attraction of the Sun, and consequently the gravitational torque, will vary with an annual period.

iv) The Moon passes the equator twice per lunar revolution, which happens roughly twice per month. This gives a nutation with a fortnightly period.

Polar motion Polar motion deals with the phenomenon of polar wander, i.e. the movement of the Earth's rotation vector in the Earth-fixed system. This is called free nutation as well. Note that the rotation axis in inertial space remains constant in the considered case of $T = 0$.

To mathematically approach polar motion, we remember the definition of angular momentum and torque in inertial system:

$$\frac{\mathrm{d}\boldsymbol{L}^i}{\mathrm{d}t} = \boldsymbol{T}^i \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{L}^i}{\mathrm{d}t} = \boldsymbol{0} \quad \text{in inertial space.}$$

But as we know, angular momentum and earth's rotation vector are not aligned. That's the difference to linear dynamics in the comparison of tbl. 3.1: While $\boldsymbol{p}$ and $\boldsymbol{v}$ are pointing in the same direction, their rotational equivalents $\boldsymbol{L}$ and $\boldsymbol{\omega}$ don't. This leads to the question:

What happens in Earth-fixed space?

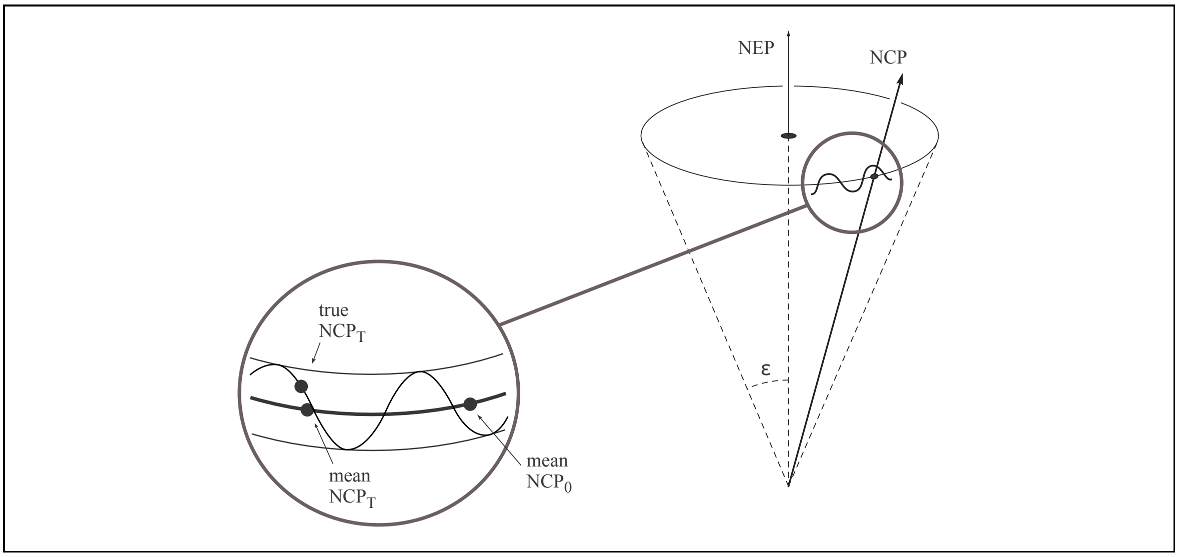


Figure 3.6: Conical motion around the NEP with radius ε.

From Equation (3.13c) we had

$$R\boldsymbol{r}^i = \dot{\boldsymbol{r}}^e + \Omega\boldsymbol{r}^e = \dot{\boldsymbol{r}}^e + \boldsymbol{\omega} \times \boldsymbol{r}^e\,, \tag{3.23}$$

and apply the same transformation to the angular momentum, which is also a kinematic quantity:

$$\begin{aligned}
R\dot{\boldsymbol{L}}^i &= \dot{\boldsymbol{L}}^e + \boldsymbol{\omega} \times \boldsymbol{L}^e = R\boldsymbol{T}^i = \boldsymbol{T}^e && (3.24a)\\
&\Downarrow\ \boldsymbol{L} = \boldsymbol{M}\boldsymbol{\omega} \\
T^e &= \dot{\boldsymbol{M}}\boldsymbol{\omega} + \boldsymbol{M}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{M}\boldsymbol{\omega} && (3.24b)\\
&\Downarrow\ \boldsymbol{T} = \boldsymbol{0} \\
\boldsymbol{0} &= \dot{\boldsymbol{M}}\boldsymbol{\omega} + \boldsymbol{M}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{M}\boldsymbol{\omega} && (3.24c)\\
&\Downarrow\ \text{assuming } \dot{\boldsymbol{M}} = \boldsymbol{0} \\
\boldsymbol{0} &= \boldsymbol{M}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{M}\boldsymbol{\omega} && (3.24d)
\end{aligned}$$

Hauptachsen-transformation

As the tensor of inertia $\boldsymbol{M}$ describes our system, the principal axes can be computed by a *body-axis transformation*. $\boldsymbol{M}$ can be diagonalized by the eigenvalue decomposition

$$\boldsymbol{M} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\mathsf{T} = \boldsymbol{Q}\begin{pmatrix} A & 0 & 0\\ 0 & B & 0\\ 0 & 0 & C\end{pmatrix}\boldsymbol{Q}^\mathsf{T}\,. \tag{3.25}$$

Figurenachsen

Defining the conventional terrestrial reference system by the *body axes* and taking an ellipsoid as the Earth's approximated shape, we can set $B = A$ and represent our ellipsoid by

$$\boldsymbol{M} = \begin{pmatrix} A & 0 & 0\\ 0 & A & 0\\ 0 & 0 & C\end{pmatrix}$$

with

$$A = 8.0131 \cdot 10^{37}\,\mathrm{kg\,m^2}$$
$$C = 8.0394 \cdot 10^{37}\,\mathrm{kg\,m^2}\,.$$

The Earth-fixed (e) system is attached to the principal axes, i.e. the Earth's body axes. Using the results from Equation (3.25), Equation (3.24d) can be written as:

$$\begin{aligned} \boldsymbol{M}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \boldsymbol{M}\boldsymbol{\omega} &= \mathbf{0} \\ \begin{pmatrix} A\dot{\omega}_1 \\ A\dot{\omega}_2 \\ C\dot{\omega}_3 \end{pmatrix} + \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} A\omega_1 \\ A\omega_2 \\ C\omega_3 \end{pmatrix} &= \mathbf{0} \end{aligned} \tag{3.26}$$

We must be aware of the two assumptions made: $\dot{\boldsymbol{M}}$ was set 0 in (3.24d) and the Earth's shape was approximated by an ellipsoid.

Under these approximations we derive the differential equations that describe polar motion in the body axes frame:

$$\begin{cases} A\dot{\omega}_1 & + (C-A)\omega_2\omega_3 & = 0 \\ A\dot{\omega}_2 & + (C-A)\omega_3\omega_1 & = 0 \\ & C\dot{\omega}_3 & = 0 \end{cases} \tag{3.27}$$

These so-called *Euler equations* are a system of coupled non-linear homogeneous ordinary differential equations. They describe the rotation of a rigid, rotationally symmetric body, which is in that case the motion of $\boldsymbol{\omega}$ in the Earth-fixed system $\boldsymbol{e}^e$. Solving non-linear differential equations is challenging. However, in this case we can solve for ω_3 first:

$$C\dot{\omega}_3 = 0 \quad \Rightarrow \quad \omega_3 = \text{constant} \tag{3.28a}$$

$$\Rightarrow \begin{cases} \dot{\omega}_1 + \frac{C-A}{A}\omega_3\omega_2 & = 0 \\ \\ \dot{\omega}_2 - \frac{C-A}{A}\omega_3\omega_1 & = 0 \end{cases} \tag{3.28b}$$

If we now set $\mu = \frac{C-A}{A}\omega_3 = 3.27 \cdot 10^{-3}\omega_3$:

$$\Rightarrow \begin{cases} \dot{\omega}_1 + \mu\omega_2 & = 0 \\ \\ \dot{\omega}_2 - \mu\omega_1 & = 0 \end{cases} \tag{3.28c}$$

Differentiating the first equation yields $\ddot{\omega}_1 + \mu\dot{\omega}_2 = 0$ and inserting $\dot{\omega}_2 = \mu\omega_1$ from above leads to:

$$\ddot{\omega}_1 + \mu^2\omega_1 = 0 \tag{3.29a}$$

$$\ddot{\omega}_2 + \mu^2\omega_2 = 0\,. \tag{3.29b}$$

The ω_2-equation arises in the same way.

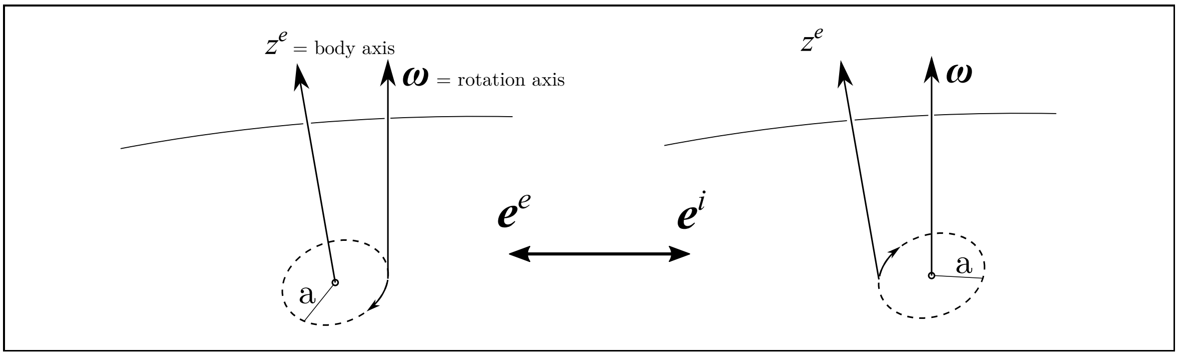


Figure 3.7: Effect of polar motion from Earth-fixed $\boldsymbol{e}^e$ to inertial system $\boldsymbol{e}^i$ where $\boldsymbol{\omega} \approx \boldsymbol{L}^i$ is fixed.

These are the equations of the harmonic oscillator with the known solution

$$\omega_1(t) = a\cos(\mu t + \varphi) \tag{3.30a}$$

$$\omega_2(t) = a\sin(\mu t + \varphi) \tag{3.30b}$$

$$\omega_3 = \text{const.}\,. \tag{3.30c}$$

This is the so-called *Euler motion* with amplitude a and phase φ (both to be determined from the initial state) and angular frequency μ. The period $T_\mu = \frac{2\pi}{\mu} = \frac{A}{C-A} T_{\omega_3} = 306\,\text{days}$ is the *Euler period*.

Interpretation of polar motion The equations above lead to two viewpoints depending on the choice of reference:

i) In the Earth-fixed system, the $\boldsymbol{\omega}$-vector describes a conical motion around the z^e-axis, which is a body axis of the Earth.

ii) In the inertial system, the body axis is rotating around $\boldsymbol{\omega}$. The Earth wobbles.

This happens when $\boldsymbol{\omega}$ and the body axis are not aligned, which is the case if our assumptions from above are not fulfilled. In reality, the Earth isn't an ellipsoid, so

$$\boldsymbol{M} \neq \begin{pmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & C \end{pmatrix}$$

and the tensor of inertia won't remain constant due to mass transports caused by hydrology, geodynamics and oceonagraphy.

$$\dot{\boldsymbol{M}} \neq 0$$

Effects in geophysics Observations in 19th century by S. C. Chandler[8] and later by the International Latitude Service validated the theory delivered by Euler. Basically, there are three effects of polar motion which can be observed (cf. fig. 3.8):

[8]Seth Carlo Chandler (1846–1913), American astronomer.

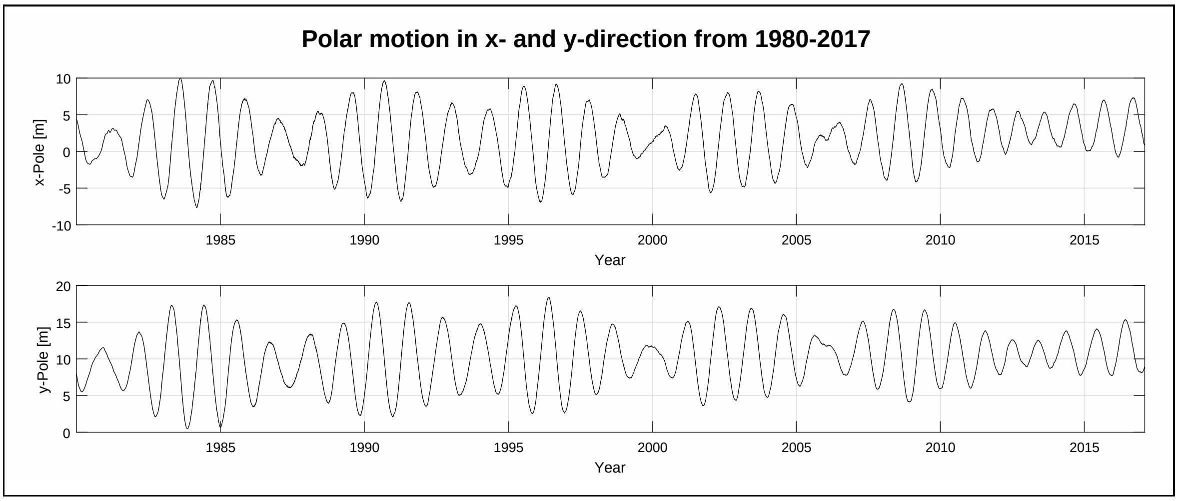


Figure 3.8: x_{p}- and y_{p}-component of the polar motion from 1980–2017. Besides the annual period a longer lasting periodic modulation and a small drift can be seen.

i) The Chandler wobble: Due to the free oscillation of the body axis, the Chandler period of $T = 435\,\mathrm{d}$ is observed instead of the 306 days predicted by Euler. That's because the Earth is a non-rigid (elastoviscous) body.

ii) Additionally, an annual period $T = 1\,\mathrm{yr}$ caused by mass transports leads to an interference of annual and Chandler period.

iii) A polar drift caused by $\dot{\boldsymbol{M}} \neq 0$ of ca. 3–4 mas/yr (10 cm/yr).

All combined, these points lead to an amplitude of $a = 0''.3$ which corresponds about 9 m on the surface of the Earth.

With a varying pole the latitude φ changes: The e-system is time-dependent due to polar motion. A correction of polar motion leads to the conventional terrestrial e_0-system.

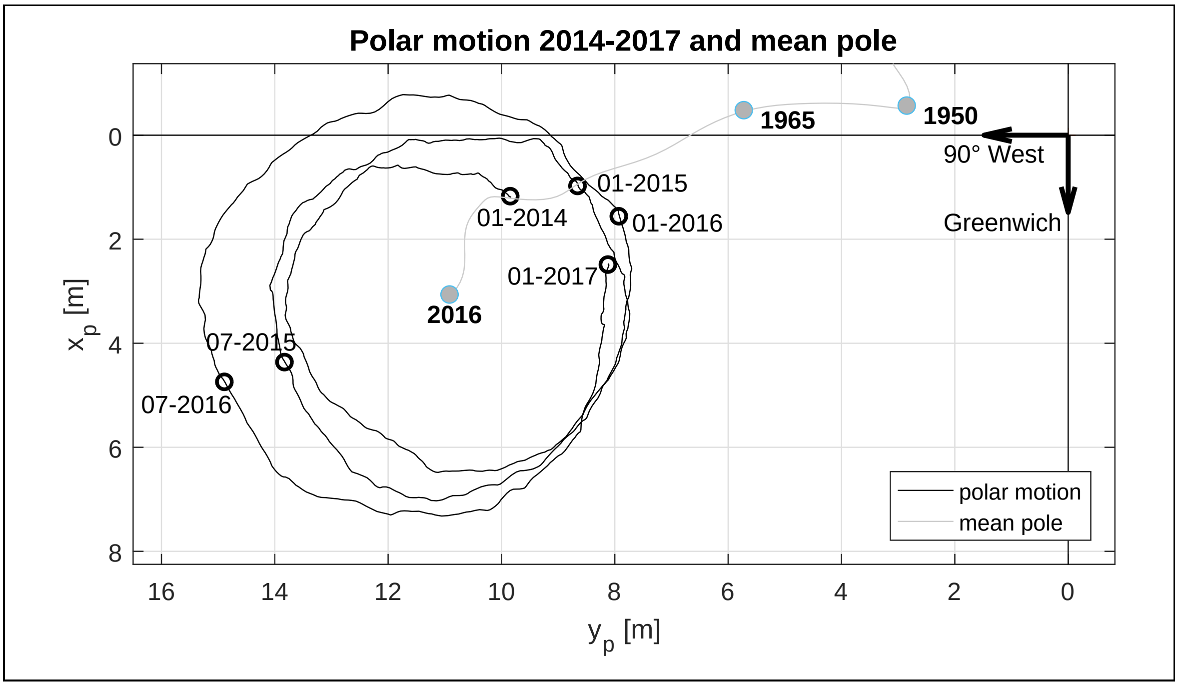


Figure 3.9: Polar motion from 2014 to January, 2017 and mean pole up to 2016.

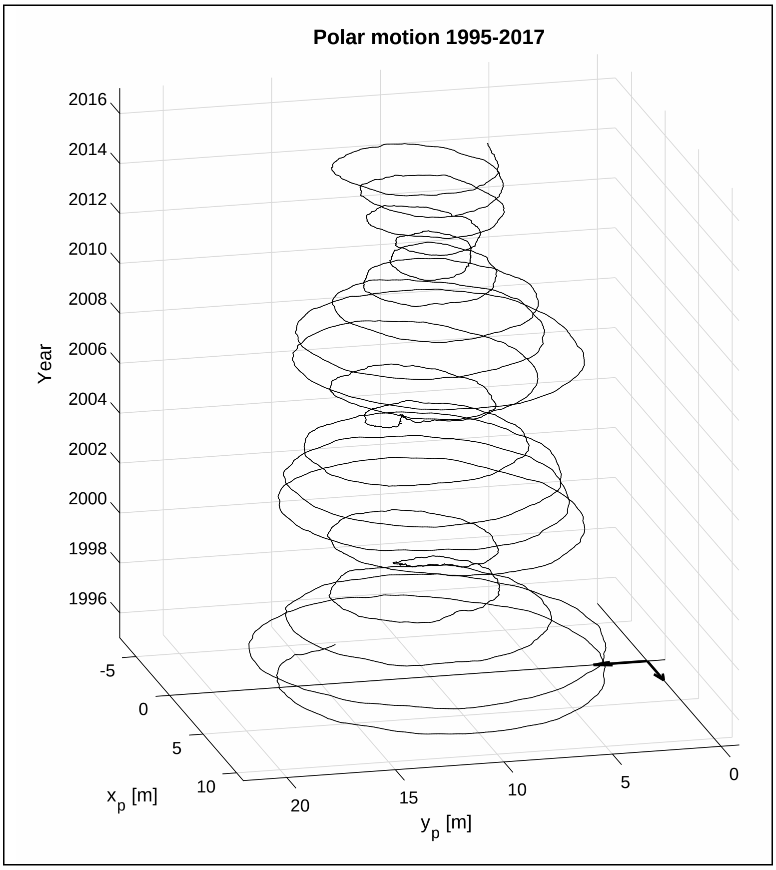


Figure 3.10: Polar motion from 1995 to January, 2017 with time in z-direction.

4 Transformations between conventional inertial and Earth-fixed reference systems $e^{i_0} \longleftrightarrow e^{e_0}$

4.1 Precession $i_0 \longleftrightarrow \bar{\imath}$

Newcomb[1] formulated the transition from the mean inertial reference system at epoch T_0 to the mean instantaneous one $i_0 \to \bar{\imath}$ as:

$$\boldsymbol{r}_{\bar{\imath}} = P\boldsymbol{r}_{i_0} = R_3(-z)R_2(\theta)R_3(-\zeta_0)\boldsymbol{r}_{i_0} \,. \tag{4.1}$$

Figure 4.2 explains which rotations need to be performed to achieve this transformation. First, a rotation around the north celestial pole at epoch T_0 (NCP$_0$) shifts the mean equinox at epoch T_0 ($\bar{\Upsilon}_0$) over the mean equator at T_0. This is $R_3(-\zeta_0)$. Next, the NCP$_0$ is shifted along the cone towards the mean pole at epoch T (NCP$_T$). This is a rotation $R_2(\theta)$, which also brings the mean equator at epoch T_0 is brought to the mean equator at epoch T. Finally, a last rotation around the new pole, $R_3(-z)$ brings the mean equinox at epoch T ($\bar{\Upsilon}_T$) back to the ecliptic. The required precession angles are given with a precision of $1''$ by:

$$\begin{aligned}
\zeta_0 &= 2306.''2181\,T + 0.''301\,88\,T^2 \\
\theta &= 2004.''3109\,T - 0.''426\,65\,T^2 \\
z &= 2306.''2181\,T + 1.''094\,68\,T^2
\end{aligned}$$

The time T is counted in *Julian centuries* (of 36 525 days) since J2000.0, i.e. January 1, 2000, 12^{h} UT1. It is calculated from calendar date and universal time (UT1) by first converting to the so-called *Julian day* number (JD), which is a continuous count of the number of days. In the following Y, M, D are the calendar year, month and day

Julianische Jahrhunderte

$$\begin{aligned}
&\text{Julian days} & \mathrm{JD} &= 367Y - \mathtt{floor}(7(Y + \mathtt{floor}((M+9)/12))/4) \\
& & &\quad + \mathtt{floor}(275M/9) + D + 1721\,014 + \mathrm{UT1}/24 - 0.5 \\
&\text{time since J2000.0 in days} & d &= \mathrm{JD} - 2451\,545.0 \\
&\text{same in Julian centuries} & T &= \frac{d}{36\,525}
\end{aligned}$$

Exercise 4.1 *Verify that the equinox moves approximately* $50''$ *per year indeed by projecting the precession angles* ζ_0, θ, z *onto the ecliptic. Use* $T = 0.01$*, i.e. one year.*

[1] Simon Newcomb (1835–1909), Canadian-American astronomer and applied mathematician.

4.2 Nutation $\bar{\imath} \longleftrightarrow i$

The following transformation describes the transition from the mean instantaneous inertial reference system to the true instantaneous one $\bar{\imath} \rightarrow i$:

$$\boldsymbol{r}_i = N\boldsymbol{r}_{\bar{\imath}} = R_1(-\varepsilon - \Delta\varepsilon)R_3(-\Delta\psi)R_1(\varepsilon)\boldsymbol{r}_{\bar{\imath}}\,. \tag{4.2}$$

Again, fig. 4.2 explains the individual rotations. First, the mean equator at epoch T is rotated into the ecliptic around $\bar{\Upsilon}_T$. This rotation, $R_1(\varepsilon)$, brings the mean north pole towards the NEP. Next, a rotation $R_3(-\Delta\psi)$ lets the mean equinox slide over the ecliptic towards the true instantaneous epoch. Finally, the rotation $R_1(-\varepsilon - \Delta\varepsilon)$ brings us back to an equatorial system, to the true instantaneous equator, to be precise. The nutation angles are known as nutation in obliquity ($\Delta\varepsilon$) and nutation in (ecliptical) longitude ($\Delta\psi$). Together with the obliquity ε itself, they are given with a precision of $1''$ by:

$$\begin{aligned}
\varepsilon &= 84\,381\overset{''}{.}448 - 46\overset{''}{.}8150\,T \\
\Delta\varepsilon &= 0\overset{\circ}{.}0026\cos(f_1) + 0\overset{\circ}{.}0002\cos(f_2) \\
\Delta\psi &= -0\overset{\circ}{.}0048\sin(f_1) - 0\overset{\circ}{.}0004\sin(f_2) \\
\text{with}& \\
f_1 &= 125\overset{\circ}{.}0 - 0\overset{\circ}{.}052\,95\,d \\
f_2 &= 200\overset{\circ}{.}9 + 1\overset{\circ}{.}971\,29\,d
\end{aligned}$$

The obliquity ε is given in seconds of arc. Converted into degrees we would have $\varepsilon \approx 23\overset{\circ}{.}5$ indeed. On top of that it changes by some $47''$ per Julian century. The nutation angles are not exact. The above formulae only contain the two main frequencies, as expressed by the time-variable angles f_1 and f_2. The coefficients to the variable d are frequencies in units of degree/day:

$$\begin{aligned}
f_1 &: \ \text{frequency} = 0.052\,95\,{}^\circ/\text{day} \quad \Rightarrow \text{period} = 18.6\,\text{years} \\
f_2 &: \ \text{frequency} = 1.971\,29\,{}^\circ/\text{day} \quad \Rightarrow \text{period} = 0.5\,\text{years}
\end{aligned}$$

The angle f_1 describes the precession of the orbital plane of the moon, which rotates once every 18.6 years. The angle f_2 describes a half-yearly motion, caused by the fact that the solar torque is zero in the two equinoxes and maximum during the two solstices. The former has the strongest effect on nutation, when we look at the amplitudes of the sines and cosines.

4.3 GAST $i \longleftrightarrow e$

For the transformation from the instantaneous true inertial system i to the instantaneous Earth-fixed sytem e we only need to bring the true equinox to the Greenwich meridian. The angle between the x-axes of both systems is the Greenwich Actual Siderial Time (GAST). Thus, the following rotation is required for the transformation $i \rightarrow e$:

$$\boldsymbol{r}_e = R_3(\text{GAST})\boldsymbol{r}_e\,. \tag{4.3}$$

The angle GAST is calculated from the Greenwich Mean Siderial Time (GMST) by applying a correction for the nutation.

$$\begin{aligned}\text{GMST} &= \text{UT1} + (24\,110.548\,41 + 8640\,184.812\,866\,T + 0.093\,104\,T^2 - 6.2\,10^{-6}\,T^3)/3600 \\ &\quad + 24n \\ \text{GAST} &= \text{GMST} + (\Delta\psi\,\cos(\varepsilon + \Delta\varepsilon))/15\end{aligned}$$

Universal time UT1 is in decimal hours and n is an arbitrary integer that makes $0 \leq \text{GMST} < 24$.

4.4 Polar motion $e \longleftrightarrow e_0$

The following transformation describes the transition from the instantaneous Earth-fixed system to the conventional terrestrial one $e \to e_0$. To correct for polar motion the Conventional International Origin (CIO) is defined as the mean pole of the years 1900–1905 measured by the International Latitude Service. A translation on the surface by x_p, y_p leads to the instantaneous pole defined by the z_e-axis. The axis through the CIO is the z_{e_0}-axis of the conventional terrestrial system e_0.

The transformation from instantaneous (true) terrestrial to the conventional terrestrial system reads

$$\boldsymbol{r}_{e_0} = R_2(-x_\text{p})R_1(-y_\text{p})\boldsymbol{r}_e\,. \tag{4.4}$$

x_p and y_p are derived from observations of the International Earth Rotation and Reference Systems Service (IERS). Written as differential rotations, (4.4) can be expressed by

$$\boldsymbol{r}_{e_0} = \begin{pmatrix} 1 & 0 & x_\text{p} \\ 0 & 1 & -y_\text{p} \\ -x_\text{p} & y_\text{p} & 1 \end{pmatrix} \boldsymbol{r}_e\,. \tag{4.5}$$

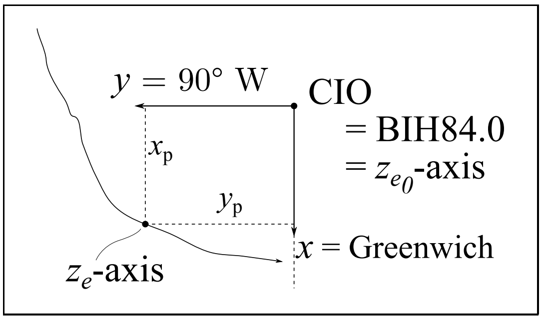


Figure 4.1: From instantaneous to conventional terrestrial system by correction of polar motion.

4.5 Conventional inertial reference system

Not only is the International Earth Rotation and Reference Systems Service (IERS) responsible for the definition and maintenance of the conventional *terrestrial* coordinate system ITRS

(International Terrestrial Reference System) and its realizations ITRF. The IERS also defines the conventional *inertial* coordinate system, called ICRS (International Celestial Reference System), and maintains the corresponding realizations ICRF.

system The ICRS constitutes a set of prescriptions, models and conventions to define at any time a triad of inertial axes.

i) origin: barycentre of the solar system ($\neq$ Sun's centre of mass),

ii) orientation: mean equator and mean equinox $\bar{\Upsilon}_0$ at epoch J2000.0,

iii) time system: barycentric dynamic time TDB,

iv) time evolution: formulae for P and N.

frame A coordinate system like the ICRS is a set of rules. It is not a collection of points and coordinates yet. It has to materialize first. The International Celestial Reference Frame (ICRF) is realized by the coordinates of over 600 that have been observed by Very Long Baseline Interferometry (VLBI). The position of the quasars, which are extragalactic radio sources, is determined by their right ascension α and declination δ.

Classically, star coordinates have been measured in the optical waveband. This has resulted in a series of fundamental catalogues, e.g. FK5. Due to atmospheric refraction, these coordinates cannot compete with VLBI-derived coordinates. However, in the early nineties, the astrometry satellite HIPPARCOS collected the coordinates of over 100 000 stars with a precision better than 1 milliarcsecond. The HIPPARCOS catalogue constitutes the primary realization of an inertial frame at optical wavelengths. It has been aligned with the ICRF.

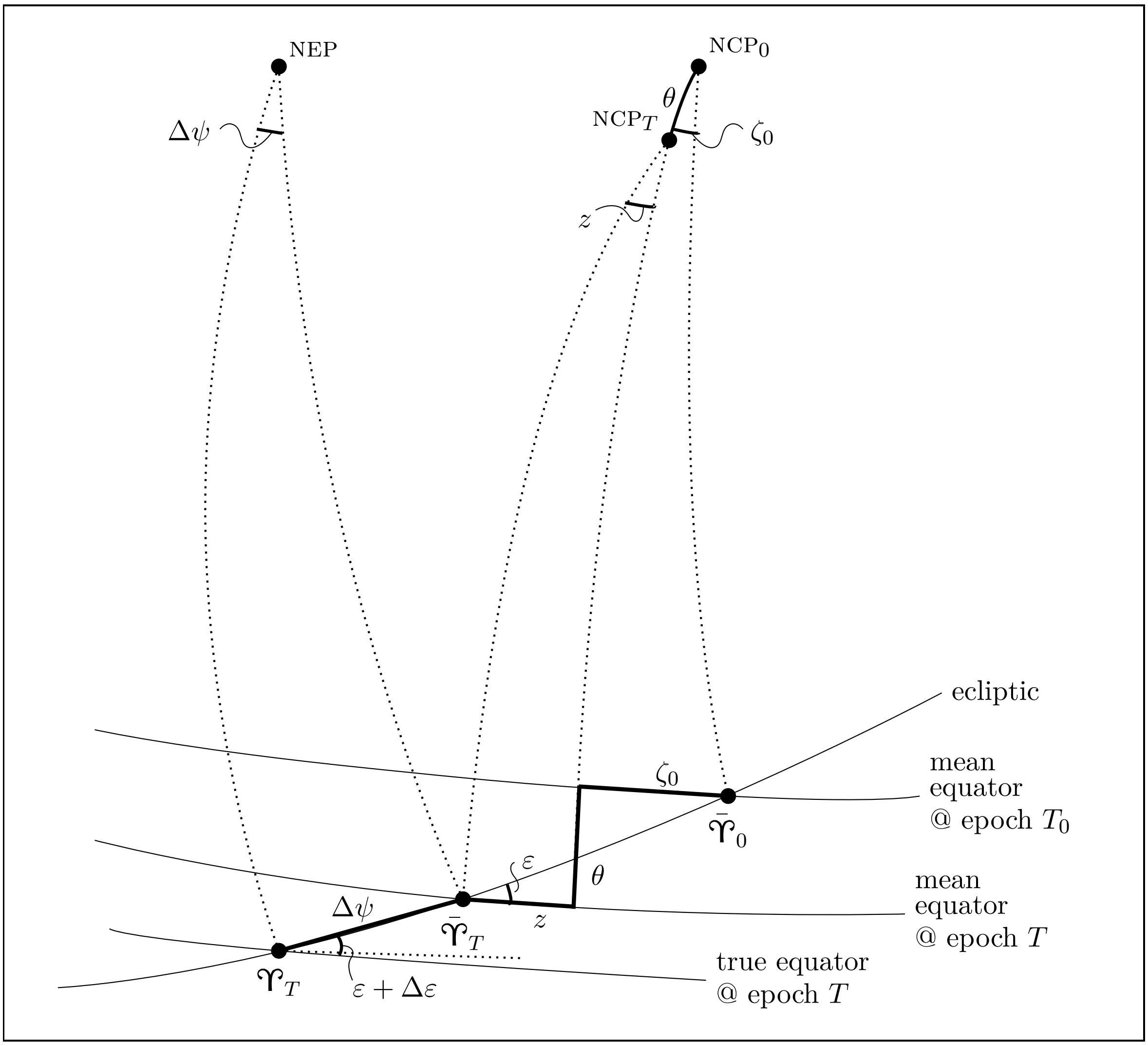


Figure 4.2: Motion of the true and mean equinox along the ecliptic under the influence of precession and nutation. This graph visualizes the rotation matrices P and N of 4.6. Note that the drawing is incorrect or misleading to the extent that *i)* The precession and nutation angles are grossly exaggerated compared to the obliquity ε, and *ii)* NCP$_0$ and NCP$_T$ should be on an ecliptical latitude circle $90° - \varepsilon$. That means that they should be on a curve *parallel* to the ecliptic, around NEP.

4.6 Summary

A. Transformation from conventional inertial (i_0) to conventional terrestrial system (e_0)

$$\begin{pmatrix} \cos\Phi\ \cos\Lambda \\ \cos\Phi\ \sin\Lambda \\ \sin\Phi \end{pmatrix} = R_2(-x_{\mathrm{P}})\ R_1(-y_{\mathrm{P}})\ R_3(\textsc{gast})\ R_1(-\varepsilon-\Delta\varepsilon)\ R_3(-\Delta\psi)\ R_1(\varepsilon)$$

$$R_3(-z)\ R_2(\theta)\ R_3(-\zeta_0)\ \begin{pmatrix} \cos\delta\ \cos\alpha \\ \cos\delta\ \sin\alpha \\ \sin\delta \end{pmatrix}$$

Required computations:

1. Julian Day (JD):

   Y: year, M: month, D: day, T: time since J2000.0 in Julian centuries

$$\begin{aligned} \textsc{jd} &= 367\,Y - \texttt{floor}(7\,(Y + \texttt{floor}((M+9)/12))/4) + \texttt{floor}(275\,M/9) \\ &\quad + D + 1721\,014 + \textsc{ut}1/24 - 0.5 \\ d &= \textsc{jd} - 2451\,545.0 \\ T &= \frac{d}{36\,525} \end{aligned}$$

2. Precession (precise to $1''$): $\quad P = R_3(-z)\ R_2(\theta)\ R_3(-\zeta_0)$

$$\begin{aligned} \zeta_0 &= 2306''.2181\ T + 0''.301\,88\ T^2 \\ \theta &= 2004''.3109\ T - 0''.426\,65\ T^2 \\ z &= 2306''.2181\ T + 1''.094\,68\ T^2 \end{aligned}$$

3. Nutation (precise to $1''$): $\quad N = R_1(-\varepsilon_0-\Delta\varepsilon)\ R_3(-\Delta\psi)\ R_1(\varepsilon_0)$

$$\begin{aligned} \varepsilon_0 &= 84\,381''.448 - 46''.8150\ T \\ \Delta\psi &= -0^\circ.0048\ \sin(f_1) - 0^\circ.0004\ \sin(f_2), \quad \text{with} \quad f_1 = 125^\circ.0 - 0^\circ.052\,95\ d \\ \Delta\varepsilon &= 0^\circ.0026\ \cos(f_1) + 0^\circ.0002\ \cos(f_2), \quad \text{with} \quad f_2 = 200^\circ.9 + 1^\circ.971\,29\ d \end{aligned}$$

4. Greenwich Mean Siderial Time (GMST) and Greenwich Apparent Siderial Time (GAST):

$$\begin{aligned} \textsc{gmst} &= (\textsc{ut}1\cdot 3600 + 24\,110.548\,41 + 8640\,184.812\,866\ T + 0.093\,104\ T^2 \\ &\quad - 6.2\ 10^{-6}\ T^3)/3600 + 24n \\ \textsc{gast} &= \textsc{gmst} + (\Delta\psi\ \cos(\varepsilon_0+\Delta\varepsilon))/15 \\ \textsc{utc} &= \textsc{cet} - 1 \\ \textsc{ut}1 &= \textsc{utc} + \Delta\textsc{ut} \end{aligned}$$

UT1 is in decimal hours, n is an arbitrary integer that makes $0 \leq \text{GMST} < 24$
CET denotes Central European Time (=MEZ).

B. Transformation from local astronomic system (g) to local geodetic system (γ)

$$\boldsymbol{r}^{\gamma} = R_3(\Delta A)\ R_2(-\xi)\ R_1(\eta)\ \boldsymbol{r}^{g} \qquad \text{with} \qquad \begin{aligned} \xi &= \Phi\ -\ \varphi \\ \eta &= (\Lambda - \lambda)\cos\varphi \\ \Delta A &= (\Lambda - \lambda)\sin\varphi \end{aligned}$$

C. Transformation from global geodetic system (ε) to local geodetic system (γ)

$$\boldsymbol{r}^{\gamma} = S_1\ R_2(\frac{\pi}{2}\ - \varphi)\ R_3(\lambda)\ \left(\boldsymbol{r}^{\varepsilon}\ -\ \boldsymbol{r}^{\varepsilon}_{0,\gamma}\right)$$

D. Transformation from conventional terrestrial system (e_0) to local astronomic system (g)

$$\boldsymbol{r}^{g} = S_1\ R_2(\frac{\pi}{2}\ - \Phi)\ R_3(\Lambda)\ \left(\boldsymbol{r}^{e_0}\ -\ \boldsymbol{r}^{e_0}_{0,g}\right)$$

E. Transformation from global geodetic system (ε) to conventional terrestrial system (e_0)

$$\boldsymbol{r}^{e_0} = \boldsymbol{r}^{e_0}_{0,\varepsilon}\ +\ \lambda\ R_1(\varepsilon_1)\ R_2(\varepsilon_2)\ R_3(\varepsilon_3)\ \boldsymbol{r}^{\varepsilon}$$

F. Transformation from instantaneous terrestrial system (e) to conventional terrestrial system (e_0)

$$\boldsymbol{r}^{e_0} = R_2(-x_{\mathrm{P}})\ R_1(-y_{\mathrm{P}})\ \boldsymbol{r}^{e}$$

5 Time systems

Time is the fourth coordinate in four-dimensional space-time. As for the coordinate systems in the previous chapters, we can discuss concepts like origin, scale, time-evolution and the distinction between a system and its realization (a frame). For these reasons alone time must be considered in a course on coordinate systems in geodesy. Moreover, the three space coordinates are strongly interwoven with time in several ways:

i) The unit of length, the metre, is defined in terms of the amount of time it takes a light wave to travel through vacuum.

ii) Most distance measurement techniques are basically timing techniques. For instance, a GPS pseudo-range is a measured time difference between satellite and receiver clocks, turned into a length measure by multiplying with the speed of light: $\rho = c\tau$. So we also have: $\mathrm{d}\rho = c\,\mathrm{d}\tau$. A ranging precision of 1 cm requires a timing precision of $33\,\mathrm{ps} = 3.3 \cdot 10^{-11}\,\mathrm{s}$.

iii) Many geodetic positioning techniques make use of satellites (GPS, SLR, DORIS) or astronomical objects (astronomical geodesy, VLBI). The rotation of the Earth and the high velocity of satellites—about $8\,\frac{\mathrm{km}}{\mathrm{s}}$ for low Earth orbiters—require precise timing.

iv) The transformation between inertial and terrestrial coordinate systems requires the angle GAST, the Greenwich Actual Sidereal Time. The Earth rotation is 360°/day or 15″/s, which amounts to about $450\,\frac{\mathrm{m}}{\mathrm{s}}$ at the equator.

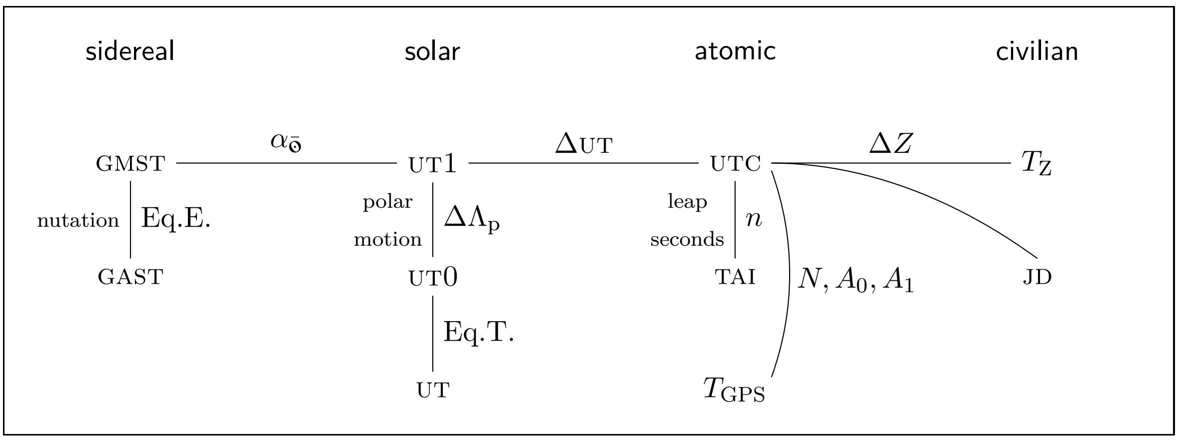


Figure 5.1: Overview: Time systems

5.1 Preliminary considerations

Time The word *time* can be understood in three senses. First, it means *epoch*, which is an instant, or a point in time. One can speak of the epoch of a GPS measurement. Second, time can be understood as an *interval*, which is just the difference between two epochs. The third sense is time *scales*, i.e. the division of an interval in time units.

Remark 5.1 *As an example, take the crossing of a star through a meridian on a certain day as epoch 1. The meridian crossing on the next day is epoch 2. The interval between these epochs is called* sidereal[1] *day. Dividing up this interval in* $24^{\mathrm{h}} = 1440^{\mathrm{m}} = 86\,400^{\mathrm{s}}$ *defines the sidereal time scale. Note that such a division in sidereal seconds is in contradiction to the definition of seconds in the SI System.*

Three classes of time systems and the transformations between them will be discussed:

i) **sidereal time** which refers to the stars,

ii) **solar or universal time** which refers to the Sun,

iii) **atomic time** which refers to atomic phenomena.

The former two categories are natural time. They describe a natural phenomenon, namely the rotation state of the Earth in space. The latter time system rather describes a physical phenomenon, namely oscillations of atoms or molecules. We will close the discussion with a few words on *calendar* dates and *Julian Day Numbers*.

There are several criteria for time systems, which may partly be contradictory.

i) The time scale should be stable, i.e. a second now should last exactly as long as it lasted yesterday. There should preferably be no drift or periodic effects in the time definition. If we only think that the Earth is slowly spinning down, this criterion is hard to meet in the long term.

ii) It should be accessible. It shouldn't be necessary, for instance, to perform complicated astronomic observations to get a time reading.

iii) It should still be available or accessible on the long term. The astronomical observation of the Babylonians can still be used. On the other hand it is doubtful whether our highly precise atomic clocks are of any use to future civilizations.

iv) For many purposes, the time system should be physically meaningful. For those purposes natural time is the preferable time system.

5.2 Sidereal time

Sternzeit

Sidereal time is the angle in the equator plane between a given meridian and the equinox ♈. This angle is conventionally expressed in units of hours, of which there are 24 in a full

[1] *Sidereal* derives from the Latin *sidus* (plural *sidera*): star.

circle: $1^{\mathrm{h}} = 15°$. According to this definition, sidereal time describes the orientation of Earth in inertial space. One sidereal day is the interval between two consecutive transits of the equinox (or of any star) through the meridian. This corresponds to a full revolution of the Earth around its axis.

If we take the true equinox Υ_T as a reference, we speak of *Actual Sidereal Time* (AST). With the mean equinox $\bar{\Upsilon}_\mathrm{T}$ we get *Mean Sidereal Time* (MST). If the particular meridian is the observer's local meridian we get *Local Sidereal Time.* On the other hand, the angle between Greenwich and the equinox is the *Greenwich Sidereal time.* All 4 potential combinations are summarized in tbl. 5.1.

	local meridian	Greenwich
true equinox	LAST	GAST
mean equinox	LMST	GMST

Table 5.1: Sidereal time.

LAST A description of sidereal time begins with the fundamental astronomical triangle on the celestial sphere, cf. fig. 5.2. This figure shows LAST as an angle between local meridian (the x-axis of the hour angle system) and true equinox (the x-axis of the instantaneous inertial system). Moreover, by including the hour circle of a given star, this figure relates LAST to the hour angle h and the right ascension α:

$$\mathrm{LAST} = \alpha + h\,. \tag{5.1}$$

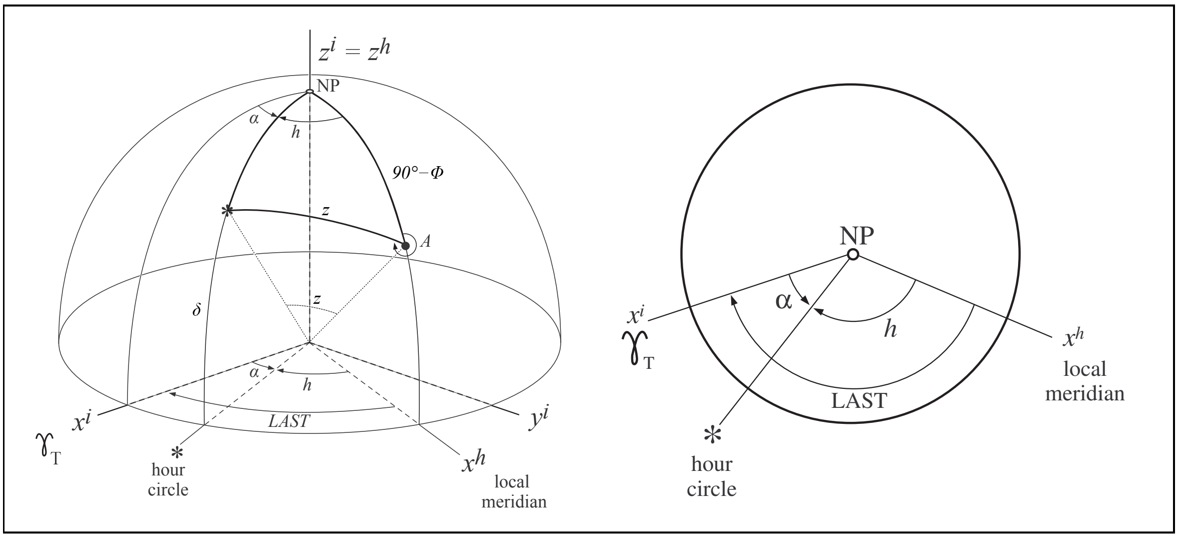


Figure 5.2: The fundamental astronomical triangle (left) and its projection on the equator plane (right).

GAST Instead of the local meridian, we can take the Greenwich meridian, see fig. 5.3. The difference between them is the astronomical longitude Λ. Thus we get the simple but fundamental relation:

$$\text{GAST} = \text{LAST} - \Lambda = \alpha + h - \Lambda\,, \tag{5.2}$$

which says that time and longitude are intimately connected. We can in principle obtain astronomical longitude from combined measurement of time (through GAST) and observation to a given star (through h), if we take the star's right ascension from a catalogue. This observation is even simpler for stars passing the local meridian, in which case we have $h = 0$. Such transits are called *upper culmination* if the star passes between zenith and north pole. A transit below the pole star is known as *lower culmination*.

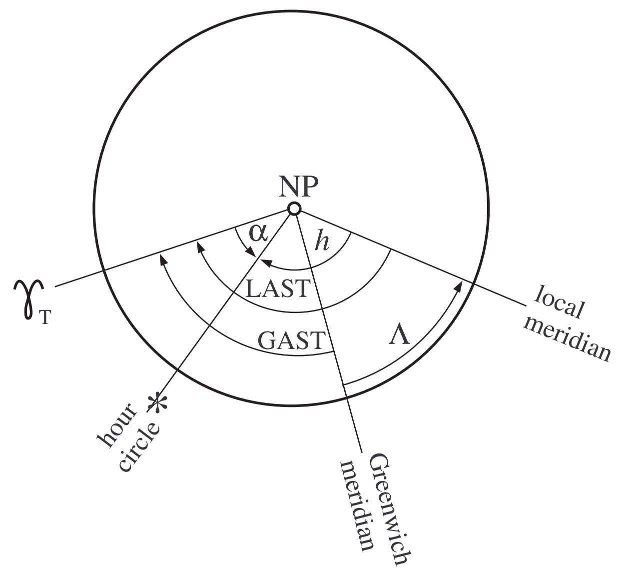


Figure 5.3: LAST vs. GAST.

LMST and GMST One of the criteria for a useful time system is the stability of the time scale. This stability is not given in case of actual sidereal time. Because of nutation, the equinox will move back and forth over the ecliptic by the angle $\Delta\psi$, the nutation in longitude. The angles LAST and GAST are relative to a time dependent reference point.

The circles drawn so far represent the equator. To correct for nutation, we have to apply the projection of $\Delta\psi$ on the equator. With the obliquity ε between ecliptic and equator, this projection becomes $\Delta\psi\, \cos\varepsilon$. This is called the *Equation of the Equinox* or Eq.E.

$$\text{Eq.E.} = \Delta\psi\, \cos\varepsilon\,, \tag{5.3a}$$

$$= \text{LAST} - \text{LMST}\,, \tag{5.3b}$$

$$= \text{GAST} - \text{GMST}\,. \tag{5.3c}$$

The variability in the Equation of the Equinox is only of the order of magnitude of roughly 1s, see fig. 5.4. For a stable definition of a sidereal time system this is relevant, though.

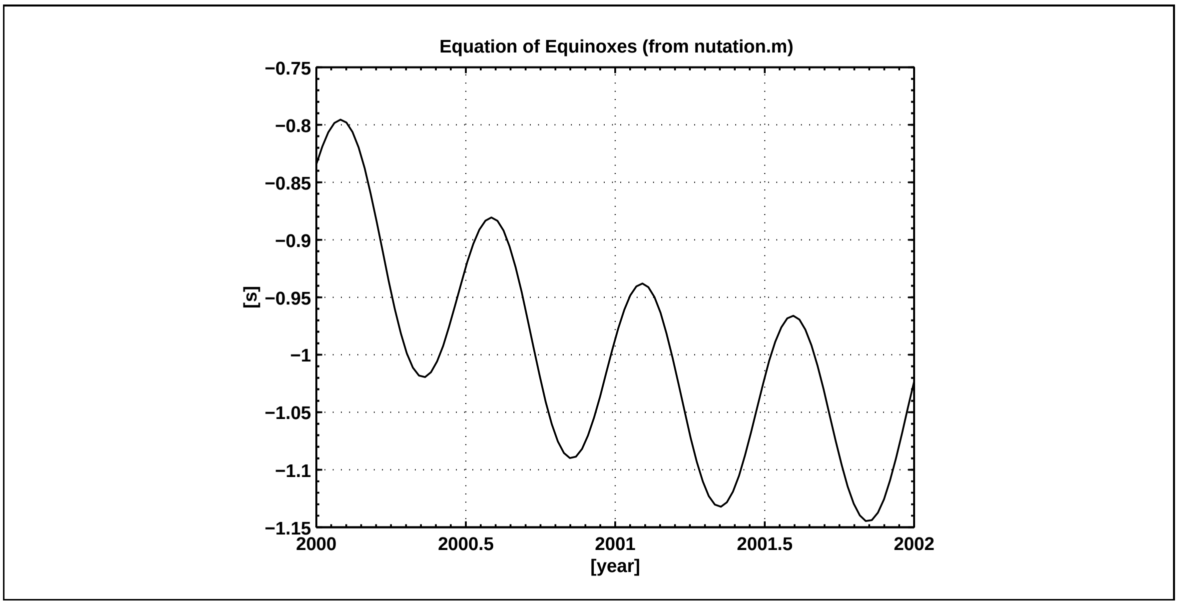


Figure 5.4: The Equation of the Equinox from a model nutation, containing only the 18.6 year and the semi-annual period.

One sidereal day is given as the time interval between two successive crossings of the mean equinox $\bar{\Upsilon}_{\mathrm{T}}$ through the local meridian. Such a crossing, during which $\textsc{lmst} = 0$ is the sidereal noon. The sidereal day is subdivided in 24 sidereal hours of 60 sidereal minutes of 60 sidereal seconds. It is important to repeat the word *sidereal* here, since these time units are slightly different from the solar equivalents in the next section.

All sidereal time angles are presented in fig. 5.5, from which we summarize the following formulas:

$$\boxed{\begin{aligned} \textsc{last} - \textsc{gast} = \textsc{lmst} - \textsc{gmst} &= \Lambda \\ \textsc{last} - \textsc{lmst} = \textsc{gast} - \textsc{gmst} &= \mathrm{Eq.E.} \\ \textsc{last} &= \alpha + h \end{aligned}} \tag{5.4}$$

Remark 5.2 *Note that the mean sidereal time refers to the mean equinox. This mean equinox is affected by precession, however. The mean equinox slides nearly uniformly over the ecliptic. For this reason the stability of the mean sidereal times as defined here is not endangered. It does mean, though, that a mean sidereal day is* 0.0084 s *shorter than a full revolution of the Earth in a conventional inertial system.*

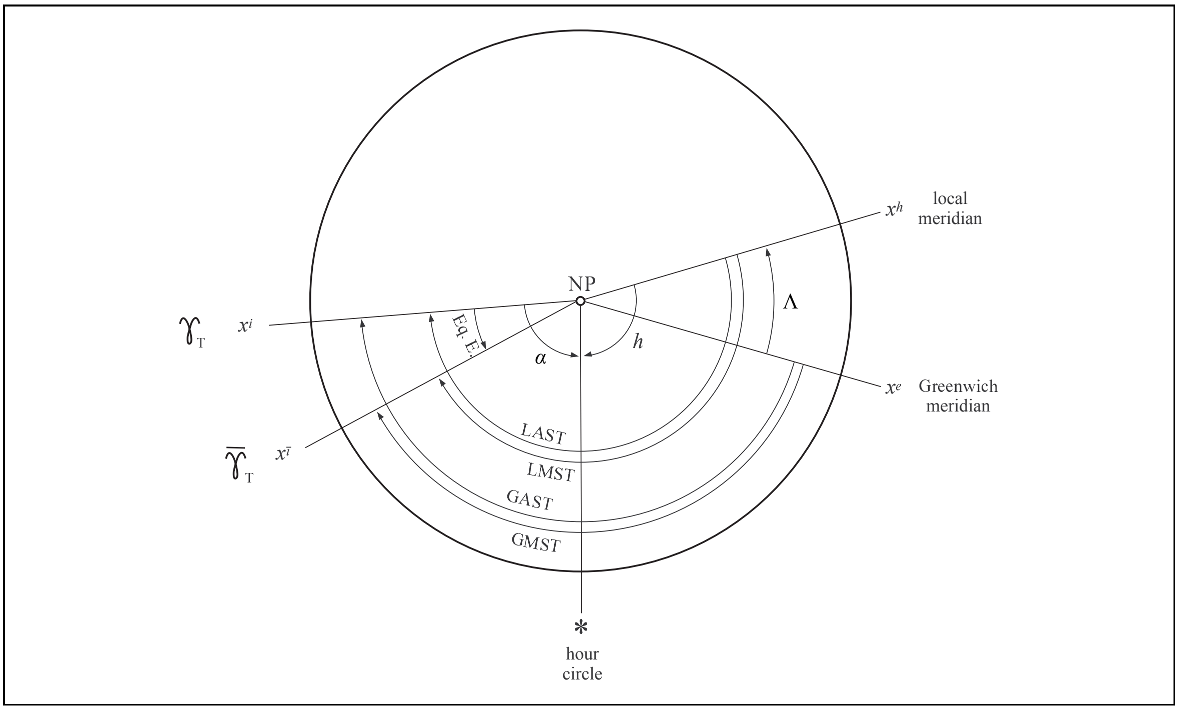


Figure 5.5: The four basic types of sidereal time.

5.3 Solar or Universal time

A civilian time system must be based on the motion of the Sun[2]. Instead of taking the equinox as a reference, we now take the Sun. Intuitively one associates noon (12^{h}) to the transition of the Sun through the local meridian. This intuitive solar time definition will require some refinement, though.

One solar day is the time span between two successive meridian transits of the Sun. This interval is divided in 24 (solar) hours of 60 minutes of 60 seconds. Opposed to a sidereal day, however, one solar day does *not* correspond to a full revolution of the Earth around its axis. Since the Earth moves around the sun once a year, one solar day is consequently slightly more then a full revolution. A year consists of 365.242 solar days. After one solar day, the Earth has travelled on average $\frac{360°}{365.242\,\mathrm{d}} = 0^\circ\!.985\,65/\mathrm{d}$ of its orbit around the Sun. This corresponds to $3^{\mathrm{m}}56^{\mathrm{s}}\!.33$ per day.

$$1 \text{ mean solar day} = 1 \text{ mean sidereal day} + 3^{\mathrm{m}}56^{\mathrm{s}}\!.33\,.$$

Consequently, a solar second is longer than a sidereal second. The scale factor between these time scales can be determined by considering that the number of sidereal days within a year is exactly one more than the amount of solar days. Therefore:

$$\text{scale factor:} \quad F = \frac{1 \text{ solar day}}{1 \text{ sidereal day}} = \frac{366.242}{365.242} = 1.002\,737\,91\,. \tag{5.5}$$

[2]We will use the symbol ☉ for the Sun.

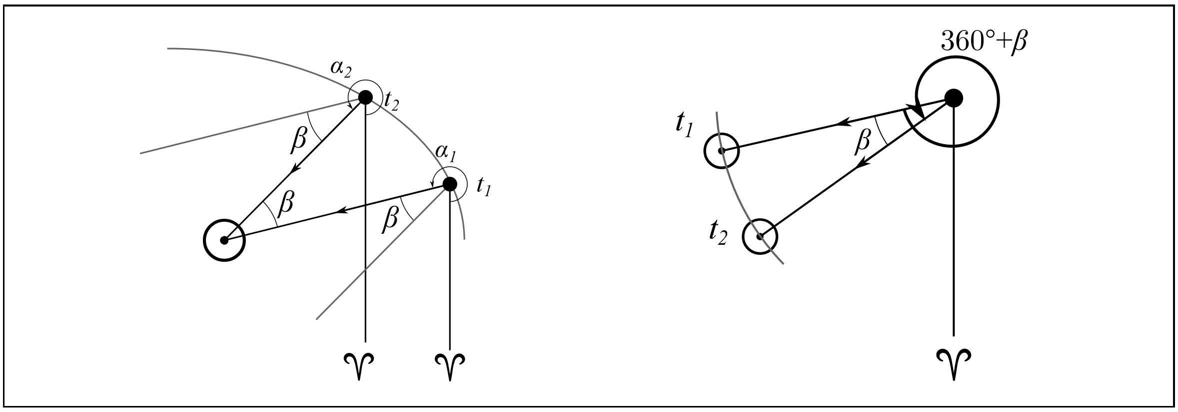


Figure 5.6: Difference (as angle β) between a solar and sidereal day relative to the sun (left) and the Earth (right).

LT, UT Figure 5.7 graphically explains the basic definition of *local solar time* (LT): it is the hour angle of the Sun $h_\odot$, i.e. the angle between local and solar meridian. However, since we associate noon with a zero hour angle, we have to add 12^{h}:

$$\mathrm{LT} = h_\odot + 12^{\mathrm{h}} \,. \tag{5.6a}$$

Similarly, by taking the Greenwich meridian, we get Greenwich solar time. This is known as true *universal time* (UT). It is called *true* because it refers to the real Sun:

$$\mathrm{UT} = h_\odot^{\mathrm{Gr}} + 12^{\mathrm{h}} \,. \tag{5.6b}$$

Local and Greenwich solar time are simply related by adding or subtracting the longitude:

$$\mathrm{LT} = \mathrm{UT} + \Lambda \,. \tag{5.6c}$$

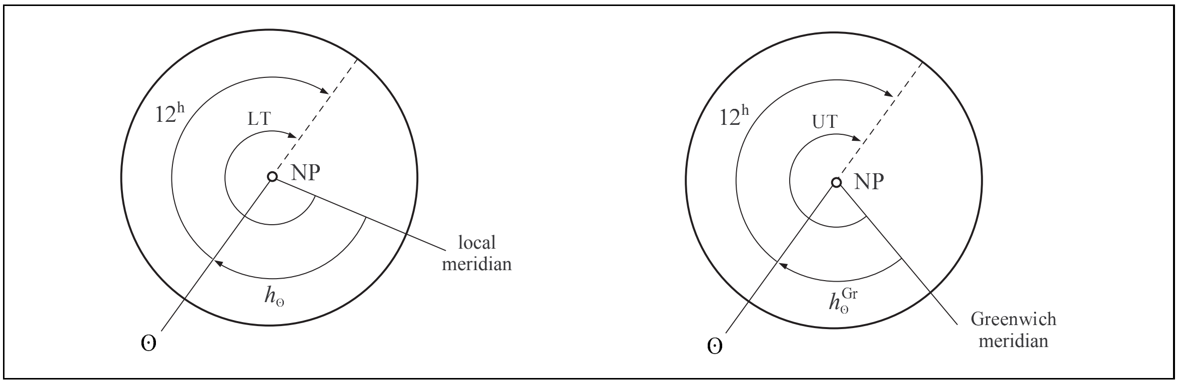


Figure 5.7: The basic local (left) and Greenwich (right) solar time.

Refinement 1: UT0 Just like the true equinox was not suitable to define a stable sidereal time system, the motion of the true Sun is not homogeneous enough to define a stable solar time system. The reason for the non-uniform motion is twofold:

Perihel

i) ellipticity of the Earth's orbit around the Sun. According to Kepler's area law, the Earth moves faster near the *perihelion* (the point closest to the Sun). We would observe the Sun lagging behind and speeding up again in a seasonal rhythm.

ii) obliquity between equator and ecliptic plane.

Therefore, a fictitious mean Sun ($\bar{\odot}$) is introduced that moves along the equator with uniform speed. The difference between the true and fictitious Sun, projected onto the equator, is called *Equation of Time*, see fig. 5.8:

$$\text{Eq.T.} = \alpha_{\bar{\odot}} - \alpha_{\odot} = h_{\odot} - h_{\bar{\odot}}\,. \tag{5.7}$$

The Equation of Time can reach values of up to ±15 minutes.

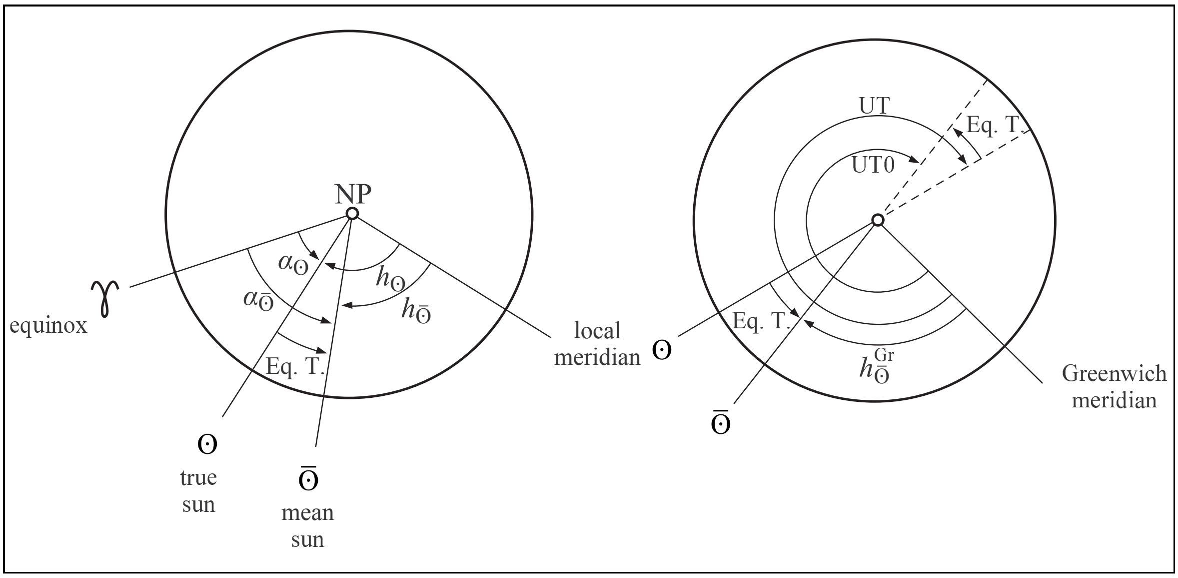


Figure 5.8: Equation of Time (left) and definition of UT0 (right).

Remark 5.3 *The equation of time is often visualized by the so-called* analemma*. This is an 8-shaped curve, which is attained by a parametric plot of $x =$ Eq.T. versus the solar declination $y = \delta_{\odot}$. The latter varies between the tropics ($\pm\varepsilon$) through the year. The analemma can also be attained photographically by opening the lens of a stationary camera every day at noon (standard time).*

Employing the mean Sun, the first refinement to UT consists of correcting for the inhomogeneous solar motion. The result is called UT0:

$$\text{UT0} = \text{UT} - \text{Eq.T.} = h^{\mathrm{Gr}}_{\bar{\odot}} + 12^{\mathrm{h}}\,. \tag{5.8}$$

Refinement 2: UT1 Because of polar motion, the instantaneous longitude of any meridian and therefore the associated solar time changes. To correct for this variability we will refer to the conventional terrestrial pole. The corresponding solar time for the conventional Greenwich meridian is called *Greenwich mean time* or UT1:

$$\begin{aligned} \mathrm{UT1} &= \mathrm{UT0} + \Delta\Lambda_\mathrm{p} \,, && (5.9a) \\ &= \mathrm{UT0} - (x_\mathrm{p} \sin\Lambda + y_\mathrm{p} \cos\Lambda) \tan\Phi \,, && (5.9b) \\ &= h_{\bar{\odot}}^{\bar{\mathrm{G}}\mathrm{r}} + 12^\mathrm{h} \,. && (5.9c) \end{aligned}$$

The result is a relatively stable time system for civilian time keeping purposes, based on a physical and observable phenomenon: the orientation of the Earth with respect to the mean Sun.

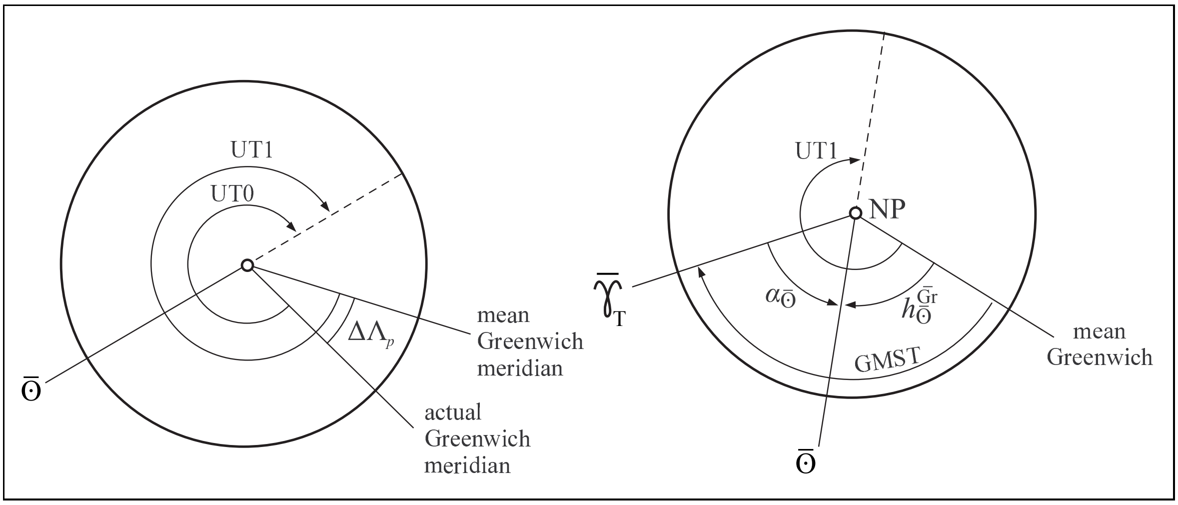


Figure 5.9: Mean Greenwich meridian and UT1 (left) and conversion between solar and sidereal time (right).

Length Of Day After these two refinement steps, UT1 will still not be stable. Due to mass redistributions on and in the Earth, the length of a mean solar day will not be constant. This fluctuation will become apparent after comparing every Lenght Of a solar Day (LOD) to 86 400 atomic seconds, cf. next section. The difference, or excess LOD is at the millisecond level, see fig. 5.10.

Remark 5.4 *A further refinement is sometimes made in order to correct for predictable periodic length-of-day variations. The resulting solar time is called* UT2.

Conversion solar $\leftrightarrow$ sidereal time The conversion between solar and sidereal time is straightforward. Figure 5.9 (right) shows the mean equinox, the mean solar meridian and the mean Greenwich meridian. From this figure we can easily verify:

$$\begin{aligned} \mathrm{GMST} &= \alpha_{\bar{\odot}} + h_{\bar{\odot}}^{\bar{\mathrm{G}}\mathrm{r}} \,, && (5.10a) \\ &= \alpha_{\bar{\odot}} + \mathrm{UT1} - 12^\mathrm{h} \,. && (5.10b) \end{aligned}$$

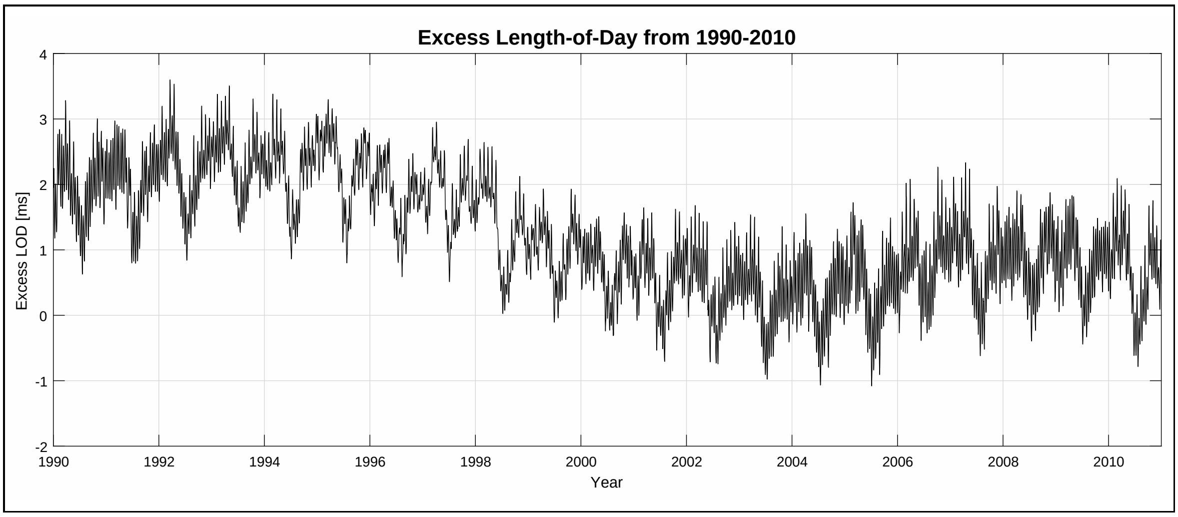


Figure 5.10: Excess Length-Of-Day.

The right ascension of the mean Sun is given by the internationally adopted formula (IAU 1976):

$$\alpha_{\bar{\odot}} = 18^{\mathrm{h}}41^{\mathrm{m}}50\overset{\mathrm{s}}{.}548\,41 + 8640\,184.812\,866\,T + \mathcal{O}(T^2)\,,$$

in which T is the time since the reference epoch J2000.0 (January 1, 2000, noon), counted in Julian centuries of 36 525 days. Thus, $T = d/36\,525$ with $d = \mathrm{JD} - 2451\,545.0$. The right ascension of the mean Sun is counted in seconds. Thus, the factor in the second term at the right hand side has units of second per Julian century. Indeed, the number 8640 184.812 866 corresponds to one full circle ($= 24^{\mathrm{h}}$) per year or $3^{\mathrm{m}}56\overset{\mathrm{s}}{.}33$ per day. With this formula for the mean solar right ascension, we get the following conversion:

$$\mathrm{GMST} = \mathrm{UT1} + 24\,110.548\,41 + 8640\,184.812\,866\,T + \mathcal{O}(T^2)\,. \tag{5.10c}$$

Again, this formula is in seconds. Moreover, an integer times 24^{h} might have to be added in order to keep $\mathrm{GMST} \in [0; 24^{\mathrm{h}}]$.

Alternatively, we could make use of the scale factor F, defined in (5.5) to perform the conversion. Consider for that purpose that if the mean Sun crosses the mean Greenwich meridian, we have $\mathrm{UT1} = 12^{\mathrm{h}} \Leftrightarrow h_{\bar{\odot}}^{\bar{\mathrm{G}}\mathrm{r}} = 0^{\mathrm{h}}$. Consequently:

$$\text{at mean noon: } \mathrm{GMST}_0 = \alpha_{\bar{\odot},0}\,.$$

Therefore, a short-term (daily) conversion formula, that only requires tabulated daily values of GMST_0, i.e. values of $\alpha_{\bar{\odot}}$ at noon, would be:

$$\begin{aligned} \mathrm{GMST} &= F\,(\mathrm{UT1} - 12^{\mathrm{h}}) + \mathrm{GMST}_0\,, \\ &= 1.002\,737\,91\,(\mathrm{UT1} - 12^{\mathrm{h}}) + \mathrm{GMST}_0\,. \end{aligned} \tag{5.11}$$

5.4 Atomic time

The second was originally defined as the fraction 1/86 400 of a mean solar day. Since the Earth's rotation shows irregularities at millisecond level in addition to a slowly decreasing spin rate, this definition became obsolete in the 20$^{\text{th}}$ century. With the advent of high precision frequency standards, based on atomic or molecular oscillations, the second was redefined in 1967:

> *The second is the duration of 9192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the* ^{133}Cs *caesium 133) atom. — 13th CGPM, 1967*

Moreover, clocks based on this atomic principle, should be resting at sea level. According to relativity theory a moving clock, or a clock that changes its potential energy level, will show a different clock rate[3].

Nowadays, Cesium clocks achieve a stability below 100 ps per day, i.e. a relative precision of 10^{-15}. More recently this stability was improved by up to two orders of magnitude by *hydrogen masers* and *optical ion traps*. This should be compared to the aforementioned astronomical definition, which may achieve a stability of 1 ms per day, i.e. a relative precision of 10^{-8}.

Two main realizations of atomic time exist.

TAI The International Atomic Time (or Temps Atomique International) is maintained by the Bureau International des Poids et Mesures (BIPM) from the readings of more than 200 atomic clocks located in metrology institutes and observatories in more than 30 countries around the world. The TAI is a weighted average of these clock readings.

$\boldsymbol{T_{\mathrm{GPS}}}$ GPS system time is given by its *Composite Clock (CC)*. The CC or *paper* clock consists of all 5 monitor stations and satellite frequency standards. Since a different set of clocks is used for this realization of atomic time, T_{GPS} will be different from TAI. Apart from a constant offset of 19 s there will be small deviations up to μs level. These differences are broadcast in the navigation message as parameters A_0 and A_1.

UTC The highly stable atomic TAI and the universal time UT1 will diverge over the years, see fig. 5.11, as the daily LOD values accumulate . The difference UT1 − TAI will increase in a non-homogeneous manner. This effect is caused by the spinning down of the Earth and by irregularities in the spin rate.

As a compromise, i.e. to keep atomic and solar time close to each other, *Universal Time Coordinated* (UTC) is introduced. The stable time scale of UTC is inherited from TAI. The trick to stay close to UT1 is to introduce *leap seconds*, such that the difference $\Delta\mathrm{UT} = \mathrm{UTC} - \mathrm{UT1} \leq 0.9$ s. The resulting time system meets all criteria, mentioned in the introduction: Schalt-sekunden

i) stability: stable definition of the second,

ii) accessibility: through atomic clocks and radio broadcast,

[3] Atomic clocks on-board GPS satellites are corrected for these relativistic effects.

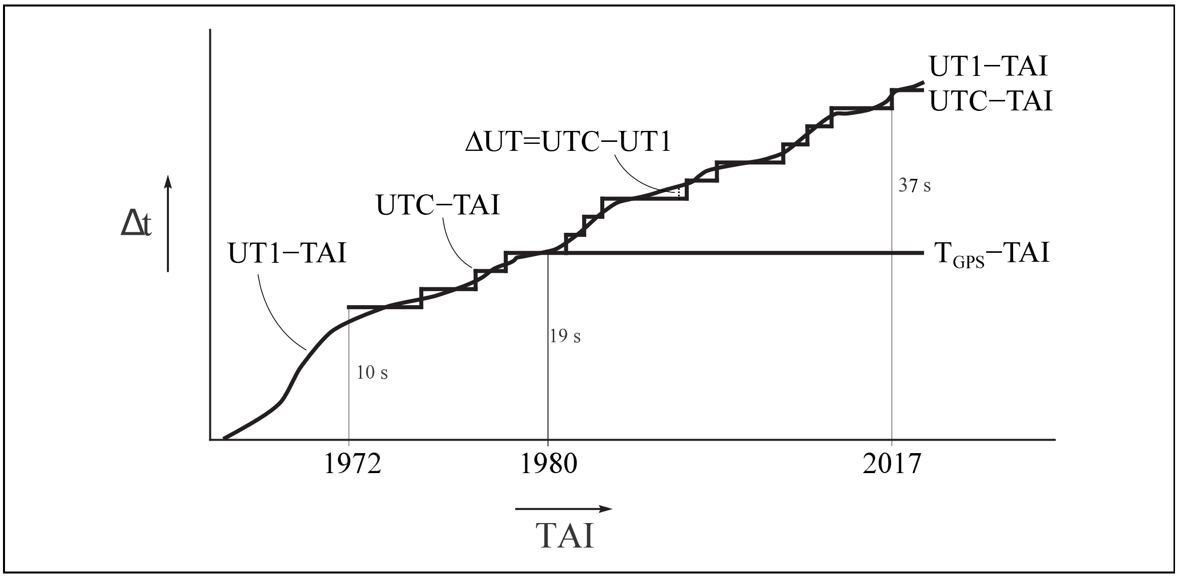


Figure 5.11: Atomic time scales and UT1.

iii) meaningfulness: very close to mean solar time, i.e. the rotation phase of the Earth.

The corresponding formulas:

$$\Delta\text{UT} = \text{UTC} - \text{UT1} \leq 0.9\,\text{s}\,, \tag{5.12a}$$

$$\text{leap seconds:} \qquad \text{UTC} - \text{TAI} = n \in \mathbb{N}\,. \tag{5.12b}$$

Leap seconds are issued by the International Earth Rotation Service (IERS), either at July 1 or January 1. The IERS makes a decision whenever ΔUT comes close to the condition (5.12a). This system started in 1972 with $n = 10$s. The last leap second was issued January 1, 2017, bringing the total to $n = 37\,$s.

At the start of GPS system time at midnight January 6, 1980, the difference between T_{GPS} and TAI was 19 s. Leap seconds are neither applied to T_{GPS} nor to TAI. As a consequence, UT1 and UTC are drifting away from GPS time as well. Thus, T_{GPS} is now 13 s apart from UTC. In general:

$$T_{\text{GPS}} - \text{UTC} = N - f(A_0, A_1), \qquad \text{with } N \in \mathbb{N} \text{ and } f(A_0, A_1) < 1\,\mu\text{s}\,. \tag{5.13}$$

5.5 Calendar time

The UTC is the basis of standard time[4], in which the Earth is roughly divided in 24 meridional zones extending 15° in longitude. Each individual zone has one single civilian time system, which differs an integer amount of hours from UTC. This guarantees that civilian time is never off by more than half an hour compared to true local time. In reality the situation is

[4]One of the most prominent *fathers* of standard time was the Canadian engineer Sir Sandford Fleming (1827–1915), surveyor for the Canadian Pacific Railway.

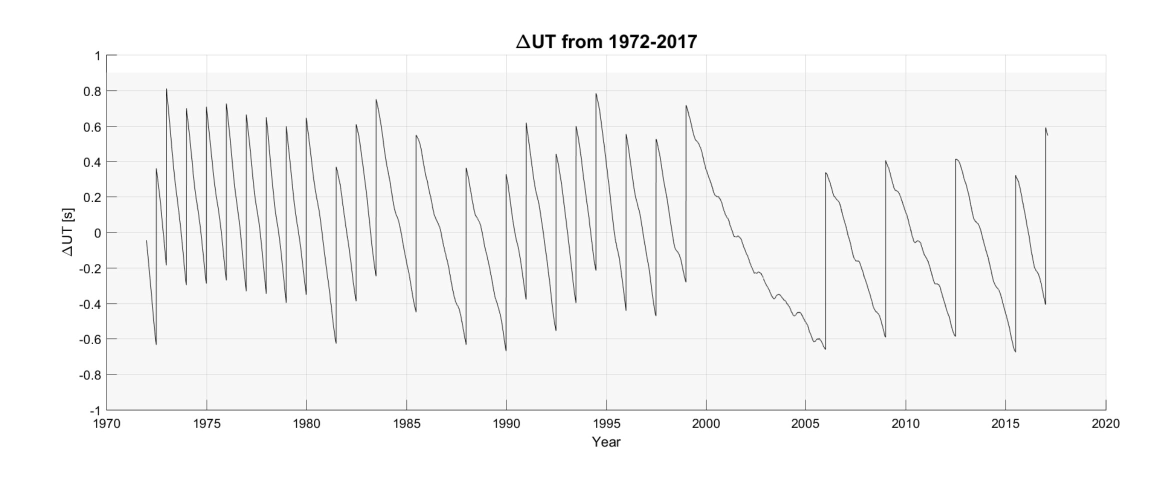


Figure 5.12: ΔUT since 1972.

slightly more complex, depending on local geography and state boundaries. Some zones even have a non-integer difference ΔZ to UTC.

$$\begin{aligned} &\text{standard time:} & T_\mathrm{Z} &= \mathrm{UTC} + \Delta Z\,. \qquad (5.14)\\ &\text{e.g.} \quad \text{Mountain Standard Time} &&= \mathrm{UTC} - 7^\mathrm{h}\,,\\ &\text{Central European Time} &&= \mathrm{UTC} + 1^\mathrm{h}\,. \end{aligned}$$

Julian Days Civilian time is counted in the *Gregorian Calendar* in days (D), months (M) and years (Y). Since the length of months is variable and since some years (leap years) have an additional day, it is difficult to calculate time intervals in terms of days. For geodetic, astronomic and chronological purposes a chronological counting of days would be more practical. This is exactly what Julian Day Numbers are.

Remark 5.5 *The Gregorian calendar reform was issued by Pope Gregory in 1582. It was motivated by the need for a correct determination of the date of Easter. The preceding Julian calendar—named after Julius Caesar—made use of Julian centuries of 36 525 days. Since the tropical year is slightly shorter than 365.25 days, the Julian calendar is one day out of sync after about 131 years. To make up for this error 10 days were skipped in October 1582 and the rule for leap years was changed. In the Julian calendar* every *fourth year was a leap year. In the Gregorian calendar years that are divisible by 100 are no leap years anymore. As an exception, years divisible by 400 are leap years.*

The Julian Day system starts counting at noon on January 1, -4712[5]. A new Julian Day, i.e. a new integer Julian Day Number, starts at noon. This makes sense for astronomical

[5] When counting days and years chronologically, the year preceding 1 AD, i.e. 1 BC, becomes the year 0. In general the year Y BC becomes $-(Y-1)$. Thus, the starting year -4712 would be 4713 BC.

purposes (hour angle equals zero). Many algorithms to compute Julian Day numbers from calendar dates exist. For instance:

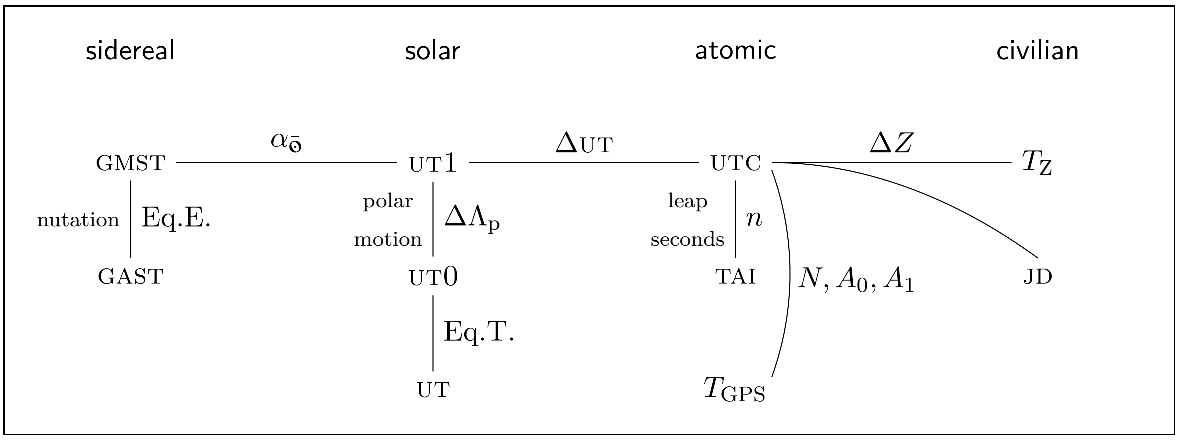
$$
\begin{aligned}
\textsc{jd} &= 367\,Y - \texttt{floor}(7(Y + \texttt{floor}((M+9)/12))/4) \\
&\quad + \texttt{floor}(275\,M/9) + D + 1721\,014 + \textsc{ut}1/24 - 0.5\,, \qquad (5.15a)\\
\textsc{mjd} &= \textsc{jd} - 2400\,000.5\,. \qquad (5.15b)
\end{aligned}
$$

The latter is the *Modified Julian Day*, which starts at midnight November 16–17, 1858. It guarantees that at most 5 digits are required in the period from 1859 to about 2130. The 0.5 at the end means that a new MJD starts at midnight.

Exercise 5.1 *Sputnik 1 was launched October 4, 1957,* $19^{\mathrm{h}}44$ *Greenwich time. What was the corresponding Julian Day?*

Exercise 5.2 *How many days have passed between your birthday and January 1, 2000?*

5.6 Overview and summary of formulas



local time — Greenwich time For local sidereal vs. Greenwich sidereal time we have:

$$\textsc{last} - \textsc{gast} = \textsc{lmst} - \textsc{gmst} = \Lambda$$

For local solar time (true or mean) vs. $\textsc{ut}i$, similar equations hold.

inertial — Earth-fixed

$$\textsc{last} = \alpha + h$$

actual time — mean time Due to nutation (sidereal) or to the combined effect of eccentricity of Earth's orbit and obliquity of ecliptic (solar):

$$\begin{aligned} \text{sidereal:} \quad & \textsc{last} - \textsc{lmst} = \textsc{gast} - \textsc{gmst} = \text{Eq.E.} = \Delta\psi \cos\varepsilon \\ \text{solar: } \textsc{ut} - \textsc{ut}0 &= \alpha_{\bar{\odot}} - \alpha_{\odot} = h_{\bar{\odot}}^{\mathrm{Gr}} - h_{\odot}^{\mathrm{Gr}} = \text{Eq.T.} \end{aligned}$$

sidereal — solar

$$\begin{aligned} \textsc{gmst} &= \alpha_{\bar{\odot}} + h_{\bar{\odot}}^{\bar{\mathrm{Gr}}} = \alpha_{\bar{\odot}} + \textsc{ut}1 - 12^{\mathrm{h}} \\ &= \textsc{ut}1 + 6^{\mathrm{h}}41^{\mathrm{m}}50\overset{\mathrm{s}}{.}548\,41 + 8640\,184\overset{\mathrm{s}}{.}812\,866\ T + \mathcal{O}(T^2) \end{aligned}$$

UT1 — UT0 Due to polar motion:

$$\textsc{ut}1 - \textsc{ut}0 = \Delta\Lambda_{\mathrm{p}} = -(x_{\mathrm{p}} \sin\Lambda + y_{\mathrm{p}} \cos\Lambda) \tan\Phi$$

solar — atomic Due to (remaining) non-uniform Earth rotation:

$$\textsc{utc} - \textsc{ut}1 = \Delta\textsc{ut} \overset{\text{def.}}{\leq} 0\overset{\mathrm{s}}{.}9$$

TAI — UTC Leap seconds due to Earth's spin-down and definition of $\Delta\textsc{ut}$:

$$\textsc{tai} - \textsc{utc} = n[\mathrm{s}] \in \mathbb{N}$$

Bibliography

Bruns, H. (1878), *Die Figur der Erde*, Publ. Königl. Preuß. Geod. Inst., Berlin.

Heiskanen, W. A. and Moritz, H. (1967), *Physical Geodesy*, W. H. Freeman and Co., San Francisco.

Torge, W. (1991), *Geodesy*, Walter de Gruyter, New York.

Vanicek, P. and Krakiwsky, E. J. (1982), *Geodesy: The Concepts*, North Holland, Amsterdam.

6 The Greek alphabet

α	A	alpha
β	B	beta
γ	Γ	gamma
δ	Δ	delta
ε, ε	E	epsilon
ζ	Z	zeta
η	H	eta
θ, ϑ	Θ	theta
ι	I	iota
κ	K	kappa
λ	Λ	lambda
μ	M	mu
ν	N	nu
ξ	Ξ	ksi
o	O	o$micron$
π, ϖ	Π	pi
ρ, ϱ	P	rho
σ, ς	Σ	sigma
τ	T	tau
υ	Υ	upsilon
φ, φ	Φ	phi
χ	X	chi
ψ	Ψ	psi
ω	Ω	o$mega$