1

1

Geodetic Coordinate Systems

Wolfgang Keller

February 16, 2001

[1]Geodetic Institute
University Stuttgart

Contents

Chapter 1

Introduction

One of the fundamental issues of Geodesy is the transition of geometrical figures into our environment. Basically, all geometrical figures are determined by lengths of straight line segments and angles, which on the first sight makes the introduction of coordinates obsolete.
But as in analytical geometry, the introduction of coordinates simplifies most of the geodetic operations. Hence, from a practical point of view coordinates are a very useful facility making geodesists life easier.

Of course in a given situation, there are always several possibilities to introduce coordinates . Which coordinate system finally is chosen is a question of practical usefulness. The consequence of the ambiguity of possible coordinate systems is the necessity of coordinate transformations.

In some respect the role of coordinates is comparable to the role of a currency: Coins do net represent a value themselves. They are a convention, which makes the exchange of goods and services easier.
As in the financial world also for coordinate systems proper conventions about their definition and transformation parameters are absolutely vital.

This lecture, gives some insight about

- the mathematical background of coordinate transformations,
- international conventions about coordinate systems and
- the way of materialization of coordinate systems.

Chapter 2

Matrices

2.1 Basic concepts

Matrices are a rectangular schemes of real numbers. They are used to describe mappings from one vector to another. Since coordinates will be defined using vectors, matrices will be an important tool for the transformation of coordinates.

Definition 1 *A rectangular scheme A of real numbers a_{ij} with m rows and n columns*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{2.1}$$

is called a (m, n) matrix.

There are a number of special types of matrices:

- zero matrix $O : a_{ij} = 0, \quad i = 1, \dots, m, j = 1, \dots, n$
- identity matrix $I : a_{ij} = 0, i \neq j, a_{ii} = 1$
- lower triangular matrix $L : a_{ij} = 0, \quad j > i, m = n$
- upper triangular matrix $U : a_{ij} = 0, \quad i > j, m = n$

The matrix algebra comprises of three operations

1. addition of matrices,
2. multiplication of matrices with a real number
3. multiplication of two matrices.

These operations are defined in the following way:

Definition 2 *A (m,n) matrix C is the sum of the two (m,n) matrices A, B, if*

$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1,\dots,m, j = 1,\dots,n \tag{2.2}$$

holds.

Definition 3 *A (m,n) matrix B is the product of the real number α and the (m,n) matrix A, if*

$$b_{ij} = \alpha \cdot a_{ij}, \quad i = 1,\dots,m, j = 1,\dots,n \tag{2.3}$$

holds.

Definition 4 *A (m,p) matrix is called the product of the (m,n) matrix A and the (n,p) matrix B , if*

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}, \quad i = 1,\dots,m, j = 1,\dots,p \tag{2.4}$$

holds.

In contrast to the addition of matrices, where both operants must have the same dimension two matrices can only be multiplied if the number of columns of the first factor equals the number of rows of the second factor.
The matrix algebra is governed by the following rules:

Theorem 1

$$
\begin{aligned}
A + B &= B + A \quad \textit{commutativity} && (2.5)\\
(A + B) + C &= A + (B + C) \quad \textit{associativity} && (2.6)\\
(A + O) &= A \quad \textit{existence of neutral element} && (2.7)\\
A + (-1) \cdot A &= O \quad \textit{existence of inverse element} && (2.8)\\
1 \cdot A &= A \quad \textit{neutral element of scalar multiplication} && (2.9)\\
\alpha(\beta A) &= (\alpha\beta)A \quad \textit{associativity 1} && (2.10)\\
(\alpha A)B &= \alpha(AB) \quad \textit{associativity 2} && (2.11)\\
A(BC) &= (AB)C \quad \textit{associativity 3} && (2.12)\\
(\alpha + \beta)A &= \alpha A + \beta A \quad \textit{distribuitivity}1 && (2.13)\\
\alpha(A + B) &= \alpha A + \alpha B \quad \textit{distributivity 2} && (2.14)\\
(A + B)c &= AC + BC \quad \textit{distributivity 3} && (2.15)\\
A(B + C) &= AB + AC \quad \textit{distributivity 4} && (2.16)
\end{aligned}
$$

Important: The multiplication of matrices is **not** commutative, i.e in general

$$AB \neq BA \tag{2.17}$$

Definition 5 *A (n, m) matrix B is called the transposed of the (m, n) matrix A*

$$B = A^\top \tag{2.18}$$

if

$$b_{ij} = a_{ji}, \quad i = 1, \ldots, n, j = 1, \ldots, m \tag{2.19}$$

holds.

For the transposition the following rules are valid:

Theorem 2

$$(A^\top)^\top = A \tag{2.20}$$
$$(A + B)^\top = A^\top + B^\top \tag{2.21}$$
$$(\alpha A)^\top = \alpha \cdot A^\top \tag{2.22}$$
$$(A \cdot B)^\top = B^\top A^\top \tag{2.23}$$

Of a special importance are matrices, which do not change when transposed.

Definition 6 *A matrix Q is called symmetric if*

$$Q = Q^\top \tag{2.24}$$

holds, it is called skew-symmetric, if

$$Q = -Q^\top \tag{2.25}$$

holds.

Theorem 3 *For an arbitrary (n, n) matrix Q there exist two $(n.n)$ matrices S, T with*

$$(i) \qquad T = T^\top \tag{2.26}$$
$$(ii) \qquad S = -S^\top \tag{2.27}$$
$$(iii) \qquad Q = S + T \tag{2.28}$$

As previously mentioned, vector transformation can be described by matrices. The concatenation of two transformations than translates into the multiplication of the corresponding matrices. The concatenation of a transformation of a transformation and its inverse transformation is the identical transformation. Hence the product of the matrix describing the transformation and the matrix describing the inverse transformation has to be the identity matrix. This gives rise to the definition of the concept of an inverse matrix.

Definition 7 *A (n, n) matrix Q^{-1} is called the inverse of the (n, n) matrix Q, if*

$$Q \cdot Q^{-1} = Q^{-1} \cdot Q = I \tag{2.29}$$

holds.

The computation of the inverse is a cumbersome process. Therefore, matrices where the inversion coincides with the simple transposition are of a particular importance.

Definition 8 *A (n, n) matrix Q is called orthogonal, if*

$$Q \cdot Q^\top = Q^\top \cdot Q = I \tag{2.30}$$

holds.

As a simple conclusion for orthogonal matrices holds

$$Q^{-1} = Q^\top \tag{2.31}$$

It is not easy, to decide whether or not a matrix is orthogonal or whether or not a matrix has an inverse. Therfore, it is useful do have an indicator for these properties. Such an indicator is the derteminant of a matrix

Definition 9 *Let be*

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix} \tag{2.32}$$

a permutation of the natural numbers $(1, 2, \dots, n)$ and $\sigma(\pi)$ the number of inversions in this permutation.
The real number

$$detA := \sum_\pi (-1)^{\sigma(\pi)} a_{1,i_1} \cdot a_{2,i_2} \cdot \dots \cdot a_{n,i_n} \tag{2.33}$$

is called the determinant of the (n, n) matrix A.

For the computation of the determinant the following rules hold:

Theorem 4

(i) *The change of two rows or two columns changes the sign* (2.34)
of the determinant (2.35)

(ii) $detQ = detQ^\top$ (2.36)

(iii) $det(\alpha Q) = \alpha^n detQ$ (2.37)

(iv) $\det(A \cdot B) = detA \cdot detB$ (2.38)

(v) $\det Q^{-1} = \frac{1}{detQ}$ (2.39)

(vi) Q *orthogonal* $\Leftrightarrow detQ = \pm 1$ (2.40)

(2.41)

2.2 Linear systems of equations

Many problems in Geodesy, engineering and natural sciences lead to the question, how to find a vector x, which for a given matrix A and a given other vector b fulfills

$$Ax = b \tag{2.42}$$

The basic algorithms for the solution of such problems are the Gaussian algorithm and its variant for symmetric positive definite matrices–the Cholesky algorithm. The essence of both algorithms is the decomposition of the matrix A into the product of two triangular matrices. For equations like (2.42) with a triangular matrix A the solution is very simple.

Theorem 5 *Let A be a lower triangular matrix. Then the solution of (2.42) is given by*

$$x_i = (b_i - \sum_{j=1}^{i-1} a_{ij} x_j)/a_{ii}, \quad i = 1, \ldots, n \tag{2.43}$$

Theorem 6 *Let A be an upper triangular matrix. Then the solution of (2.42) is given by*

$$x_i = (b_i - \sum_{j=i+1}^{n} a_{ij} x_j)/a_{ii}, \quad i = n, n-1, \ldots, 1 \tag{2.44}$$

Hence, the only question , which still has to be solved is how to decompose a given matrix into two triangular matrices. This can be done by the algorithm of Gauß - Banachiewic.

Theorem 7 *Let be a a strongly regular matrix. Then there exist*

- *a uniquely determined upper triangular matrix U and*
- *a uniquely determined lower triangular matrix L having $l_{ii} = 1, i = 1, \ldots, n$*

with

$$A = L \cdot U \tag{2.45}$$

The matrices U, L can be computed according to:

$$\left.\begin{array}{ll} u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} & j = i, \ldots, n \\ l_{ji} = (a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki})/u_{ii} & j = i+1, \ldots, n \end{array}\right\} i = 1, \ldots, n \tag{2.46}$$

The original problem (2.42) is now equivalent to

$$L \cdot Ux = b \tag{2.47}$$

With the substitution $y = Ux$ (2.42) decomposes in to two systems with triangular matrix

$$Ly = b, \quad Ux = y \tag{2.48}$$

Having available the algorithm for the triangular decomposition of a regular matrix, the solution of (2.42) consists of three steps

1. triangular decomposition $A = L \cdot U$
2. forward elimination $Ly = b$
3. backward substitution $Ux = y$.

The main numerical effort nis spent for the decomposition of the matrix A. In the case of a symmetric matrix there should be a variant of the decomposition algorithm with a strongly reduced numerical effort. This variant is the Cholesky algorithm.

Definition 10 *A (n, n) matrix A is called positive definite, if*

$$x^\top Ax > 0, \quad \forall x \neq 0 \tag{2.49}$$

holds.

Theorem 8 *Let A be a symmetric positive definite matrix. Then there exist a uniquely determined lower triangular matrix L, with*

$$A = L \cdot L^\top \tag{2.50}$$

The matrix L can be computed according to:

$$\left.\begin{array}{l} l_{ii} = \sqrt{a_{ij} - \sum_{k=1}^{i-1} l_{ik}^2} \\ \\ l_{ji} = (a_{ji} - \sum_{k=1}^{i-1} l_{ik} l_{jk})/l_{ii} \quad j = i+1, \ldots, n \end{array}\right\} i = 1, \ldots, n \tag{2.51}$$

2.3 Linear adjustment

So far we have considered only linear systems of equations where the number of equations equals the number of unknowns. Typical for Geodesy is that for the purpose of error control the number of equations exceeds the number of unknowns. Therefore, it will not be possible to find a solution vector , which fulfills all equations. It will only be possible, to find a solution vector , which optimally adjusts Ax to the inhomogeneity b.
The discipline of Geodesy, which deals with such overdetermined systems of linear equations is called linear adjustment theory.

For an overdetermined system of linear equations there is no hope that the vector of residuals

$$r := Ax - b \tag{2.52}$$

will be identical to zero. It can only be tried to make this residual r as small as possible . This leads to the linear adjustment problem

$$\min_{x \in \mathbb{R}^n} \{r^\top r \mid r = Ax - b\} \tag{2.53}$$

The solution of this adjustment problem is given by

Theorem 9 *The vector*

$$\hat{x} := (A^\top A)^{-1} A^\top b \tag{2.54}$$

is the solution of the linear adjustment problem (2.53).

It has to be noticed that for the special case of a uniquely determined system the adjustment solution coincides with the traditional solution

$$\hat{x} = (A^\top A)^{-1} A^\top b = A^{-1}(A^\top)^{-1} A^\top b = A^{-1} b$$

Chapter 3

Vector-spaces

The concept of a vector is familiar to everybody: a directed line segment. Also the vector operations are very well known: addition of vectors by concatenating them. On the other hand a clear mathematical definition of a vector is by no means a trivial task. Therfore for the definition of vectors an inverse method is selected: The central set of rules for the vector operations is chosen as criterion. All objects with operations fulfilling these rules are called vectors.

Definition 11 *A set $V(\mathbb{R})$ is called real vector space, if there are two operations*

$$+ : V(\mathbb{R}) \times V(\mathbb{R}) \to V(\mathbb{R}), \quad \textit{vector addition} \tag{3.1}$$

$$\cdot : \mathbb{R} \times V(\mathbb{R}) \to V(\mathbb{R}), \quad \textit{scalar multiplication} \tag{3.2}$$

defined, which fulfill the following conditions:

$$A1 \quad : \quad u + v = v + u \tag{3.3}$$

$$A2 \quad : \quad u + (v + w) = (u + v) + w \tag{3.4}$$

$$A3 \quad : \quad \textit{There is an element } \mathcal{O} \in V(\mathbb{R}) \textit{with } u + \mathcal{O} = u \tag{3.5}$$

$$A4 \quad : \quad \textit{For any } u \in V(\mathbb{R}) \textit{ there is exactly one element } (-u) \in V(\mathbb{R}) \tag{3.6}$$

$$\textit{with } u + (-u) = \mathcal{O} \tag{3.7}$$

$$M1 \quad : \quad 1 \cdot u = u \tag{3.8}$$

$$M2 \quad : \quad \alpha(\beta u) = (\alpha\beta)u \tag{3.9}$$

$$M3 \quad : \quad (\alpha + \beta)u = \alpha u + \beta u \tag{3.10}$$

$$M4 \quad : \quad \alpha(u + v) = \alpha u + \alpha v \tag{3.11}$$

Besides the directed line segments, which obviously fulfill these conditions there is a very large variety of mathematical objects which also can be considered as vectors:

1. The set of matrices with identical dimensions.
2. The set of n-tuples $(u_1, \ldots, u_n)$ of real numbers.
3. The set of all polynomials of fixed degree:

$$P_n(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \quad (3.12)$$

One of the essential technique in the vector algebra is the representation of all vectors of the vector space by a number of prototype vectors.

Definition 12 *The vectors $u_1, \ldots, u_m \in V(\mathbb{R})$ are called linear dependent, if there are real numbers $a_1, \ldots, a_m$ which are not all identical to zero and which fulfill*

$$\mathcal{O} = a_1 u_1 + a_2 u_2 + \ldots + a_m u_m \quad (3.13)$$

If there are no such numbers, the vectors are called linearly independent.

The motivation for this definition comes from the fact that for linearly dependent vectors it is possible to express one vector by the rest of the set. For independent vectors this is impossible.

Definition 13 *The maximal number of linearly independent vectors of a vector space $V(\mathbb{R})$ is called its dimension.*

Vectors can be very complicated objects and vector operations can be even more complicated. It would be desirable to map vector operations to operations on real numbers. This can be achieved by the introduction of so-called vector bases.

Definition 14 *A set of n linearly independent vectors of a n-dimensional vector space $V(\mathbb{R})$ is called a base of this vector space.*

Theorem 10 *Let $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ be a base of the n-dimensional vector space $V(\mathbb{R})$ and let be $\mathbf{x} \in V(\mathbb{R})$. There are uniquely determined real numbers $x_1, \ldots, x_n$ with*

$$\mathbf{x} = x_1 \mathbf{g}_1 + \ldots + x_n \mathbf{g}_n \quad (3.14)$$

Definition 15 *The uniquely determined real numbers $x_1, \ldots, x_n$ are called the coordinates of $\mathbf{x}$ with respect to the base $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$.*

Due to the one-to-one mapping of vectors to their coordinates vector operations can be mapped to operations between real numbers.

Theorem 11 *Let be $V(\mathbb{R})$ a n-dimensional vector space and $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ a base of $V(\mathbb{R})$. Two vectors $\mathbf{x}, \mathbf{y} \in V(\mathbb{R})$ then have the following representation with respect to the base*

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{g}_1, \quad \mathbf{y} = \sum_{i=1}^{n} y_i \mathbf{y}_i \tag{3.15}$$

Then the following relations hold:

$$(\mathbf{x} + \mathbf{y}) = \sum_{i=1}^{n} (x_i + y_i) \mathbf{g}_i, \quad \alpha \mathbf{x} = \sum_{i=1}^{n} (\alpha x_i) \mathbf{g}_i \tag{3.16}$$

The interpretation of this theorem is that vectors are added by adding their coordinates and that vectors are multiplied by real numbers by multiplying their coordinates by these real numbers.

Of course there is alway more than one base in a vector space and the choice of a particular base is dictated by the application one has in mind. This ambiguity of possible bases often requires a change of the basis. This means: given the coordinates of a vector with respect to one base how can the coordinates of this vector with respect to another base be computed?

Theorem 12 *Let $\{\mathbf{g}_1, \ldots, \mathbf{g}_n\}$ and $\{\bar{\mathbf{g}}_1, \ldots, \bar{\mathbf{g}}_n\}$ be two different bases of a n-dimensional vector space $V(\mathbb{R})$. Then there are uniquely determined real numbers $T_{ji}, \bar{T}_{ij}, i, j = 1, \ldots, n$ with*

$$\mathbf{g}_i = \sum_{j=1}^{n} T_{ji} \bar{\mathbf{g}}_j, \quad i = 1, \ldots, n \tag{3.17}$$

$$\bar{\mathbf{g}}_j = \sum_{i=1}^{n} \bar{T}_{ij} \mathbf{g}_i, \quad j = 1, \ldots, n \tag{3.18}$$

Theorem 13 *Let $x_1, \ldots, x_n$ and $\bar{x}_1, \ldots, \bar{x}_n$ be the coordinates of a vector $\mathbf{x}$ with respect to the base $\{\mathbf{g}, \ldots, \mathbf{g}_n\}$ and to the base $\{\bar{\mathbf{g}}_1, \ldots, \bar{\mathbf{g}}_n\}$ respectively. Then it holds*

$$\bar{x}_j = \sum_{i=1}^{n} T_{ji} x_i, \quad j = 1, \ldots, n \tag{3.19}$$

$$x_i = \sum_{j=1}^{n} \bar{T}_{ij} \bar{x}_j, \quad i = 1, \ldots, n \tag{3.20}$$

Chapter 4

Cartesian Coordinate Systems

Coordinates are a useful and necessary prerequisite to perform geometrical constructions in an analytical way. A coordinate system is nothing heavenly given but it man-made. Therefore, there is not only one coordinate system for the analytical description of a geometrical situation but in many cases a great variety of possible coordinate system.
Depending on the choice of a particular coordinate system the geometrical problem can be solved easily or can be extremely complicated. It is therefore necessary to formulate a given problem in different coordinate systems and transform the given information between those coordinate systems.

4.1 Definition of Cartesian Coordinate Systems

Definition 16 *A bijective mapping of the Euclidean space E^3 to the real numbers $\mathbb{R}^3$ is called a coordinate system in E^3.*
The real numbers, a point is mapped to are called the coordinates of this point.

The simplest coordinate system in E^3 is a Cartesian coordinate system. This can be done by the connection of Geometry with vector algebra.

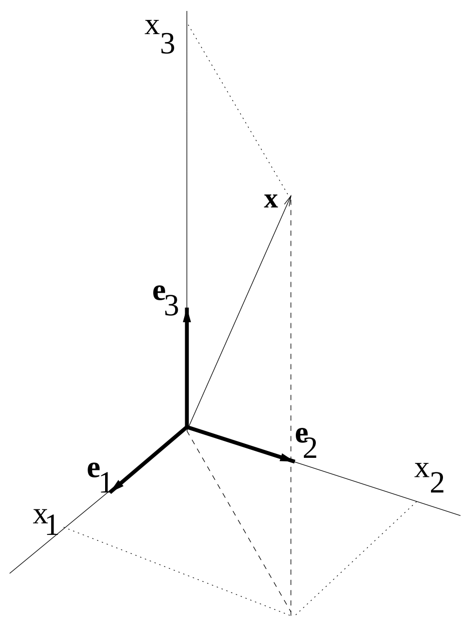
Definition 17 *A Cartesian coordinate system $\{\mathcal{O}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of the Euclidean space E^3 consists of*

- *an arbitrary but fixed point $\mathcal{O}$-the origin of the coordinate system*
- *three mutually orthogonal vectors $\mathbf{e}_i$ having unit length*

If $\mathbf{x}$ is the uniquely defined vector connecting a given point $P \in E^3$ with the origin $\mathcal{O}$, the real numbers

$$x_i := \mathbf{x}^\top \mathbf{e}_i \tag{4.1}$$

are the Cartesian coordinates of P.



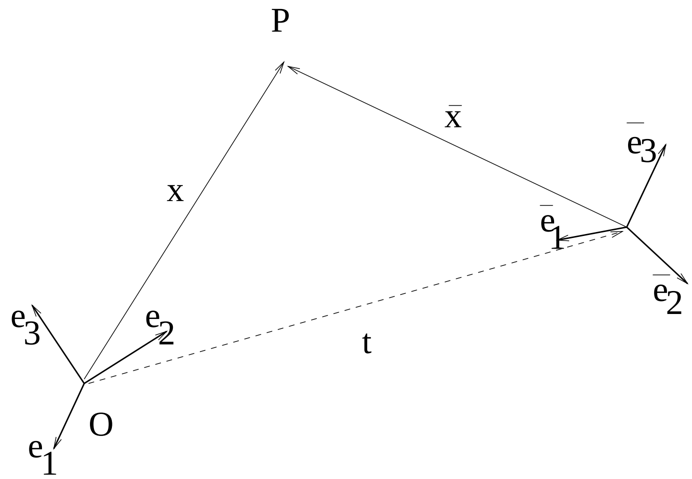
4.2 Transformation Between Cartesian Coordinate Systems

Since different location of the origins and different orientation of the unit vectors are possible different cartesian coordinate systems can be used. One point can have different coordinates with respect to different coordinate systems. Therefore it is necessary to transform the coordinates of one point from one coordinate system to another.

Theorem 14 *Let (x_1, x_2, x_3) be the cartesian coordinates of a point with respect to the coordinate system $\{\mathcal{O}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ the coordinates of the same point with respect to a different coordinate system $\{\bar{\mathcal{O}}, \bar{\mathbf{e}}_,\bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3\}$.*
Then, there exist a uniquely defined vector $\mathbf{t}$ and a uniquely defined matrix $\mathbf{T}$ that

$$\bar{\mathbf{x}} = \mathbf{t} + \mathbf{T} \cdot \mathbf{x} \tag{4.2}$$

holds.



Proof: Let $\mathbf{t}$ be the vector, which connects the two origins $\mathcal{O}$ and $\bar{\mathcal{O}}$. With respect to the base $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3\}$ this vector has the unique representation

$$\mathbf{t} = t_1\bar{\mathbf{e}}_1 + t_2\bar{\mathbf{e}}_2 + t_3\bar{\mathbf{e}}_3$$

.
On the other hand, each of the base vectors $\mathbf{e}_i$ has a unique representation with respect to the base $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3\}$

$$\mathbf{e}_i = \sum_{k=1}^{3} t_{ki}\bar{\mathbf{e}}_k$$

This leads to the following representation of the point P

$$\begin{aligned} P &= \mathcal{O} + \sum_{i=1}^{3} x_i \mathbf{e}_i \\ &= \mathcal{O} + \sum_{i=1}^{3} x_i \sum_{k=1}^{3} t_{ki}\bar{\mathbf{e}}_k \\ &= \bar{\mathcal{O}} + \sum_{k=1}^{3} t_k\bar{\mathbf{e}}_k + \sum_{k=1}^{3}(\sum_{i=1}^{3} t_{ki}x_i)\bar{\mathbf{e}}_k \\ &= \bar{\mathcal{O}} + \sum_{k=1}^{3}(t_k + \sum_{i=1}^{3} t_{ki}x_i)\bar{\mathbf{e}}_k \end{aligned}$$

On the other hand , the same point P has another representation

$$P = \bar{\mathcal{O}} + \sum_{k=1}^{3} \bar{x}_k\bar{\mathbf{e}}_k$$

Since the representation with respect to a base is unique, the relation

$$\bar{x}_k = t_k + \sum_{i=1}^{3} t_{ki} x_i$$

or in matrix notation

$$\bar{\mathbf{x}} = \mathbf{t} + \mathbf{T} \cdot \mathbf{x} \quad , \quad \mathbf{T} = (t_{ki})$$

Since both bases are orthogonal,but not necessarily orthonormal the matrix $\mathbf{T}$ has some special properties:

Theorem 15 *For the transformation matrix* $\mathbf{T}$ *holds*

$$\mathbf{T}^\top \mathbf{T} = \mathbf{T}\mathbf{T}^\top = m\mathbf{I} \tag{4.3}$$

Proof: Both, the base $\{\mathbf{e}_i\}$ and $\{\bar{\mathbf{e}}_j\}$ are orthogonal.

$$m\delta_{ij} = \mathbf{e}_i^\top \mathbf{e}_j \quad , \quad \bar{m}\delta_{ij} = \bar{\mathbf{e}}_i^\top \bar{\mathbf{e}}_j$$

Consequently,

$$\begin{aligned} m\delta_{ij} &= (\sum_{k=1}^{3} t_{ki}\bar{\mathbf{e}}_k)^\top (\sum_{l=1}^{3} t_{lj}\bar{\mathbf{e}}_l) \\ &= \bar{m} \sum_{k=1}^{3} t_{ki} t_{kj} \end{aligned}$$

in matrix notation this yields

$$\frac{\bar{m}}{m}\mathbf{I} = \mathbf{T}^\top \mathbf{T}$$

This means, that for cartesian coordinate systems the transformation matrix $\mathbf{T}$ is a multiple of an orthogonal matrix. Every orthogonal matrix can be represented as the product of three rotation matrices. Either in the Cartan representation

$$\mathbf{T} = m\mathbf{R}_3(\gamma)\mathbf{R}_2(\beta)\mathbf{R}_1(\alpha) \tag{4.4}$$

or in the Eulerian representation

$$\mathbf{T} = m\mathbf{R}_3(\psi)\mathbf{R}_1(\vartheta)\mathbf{R}_3(\varphi) \tag{4.5}$$

Hence, the transformation between two coordinate systems is completely known, if 7 parameters

- three translation parameters,
- three rotation angles,

- one scale ratio

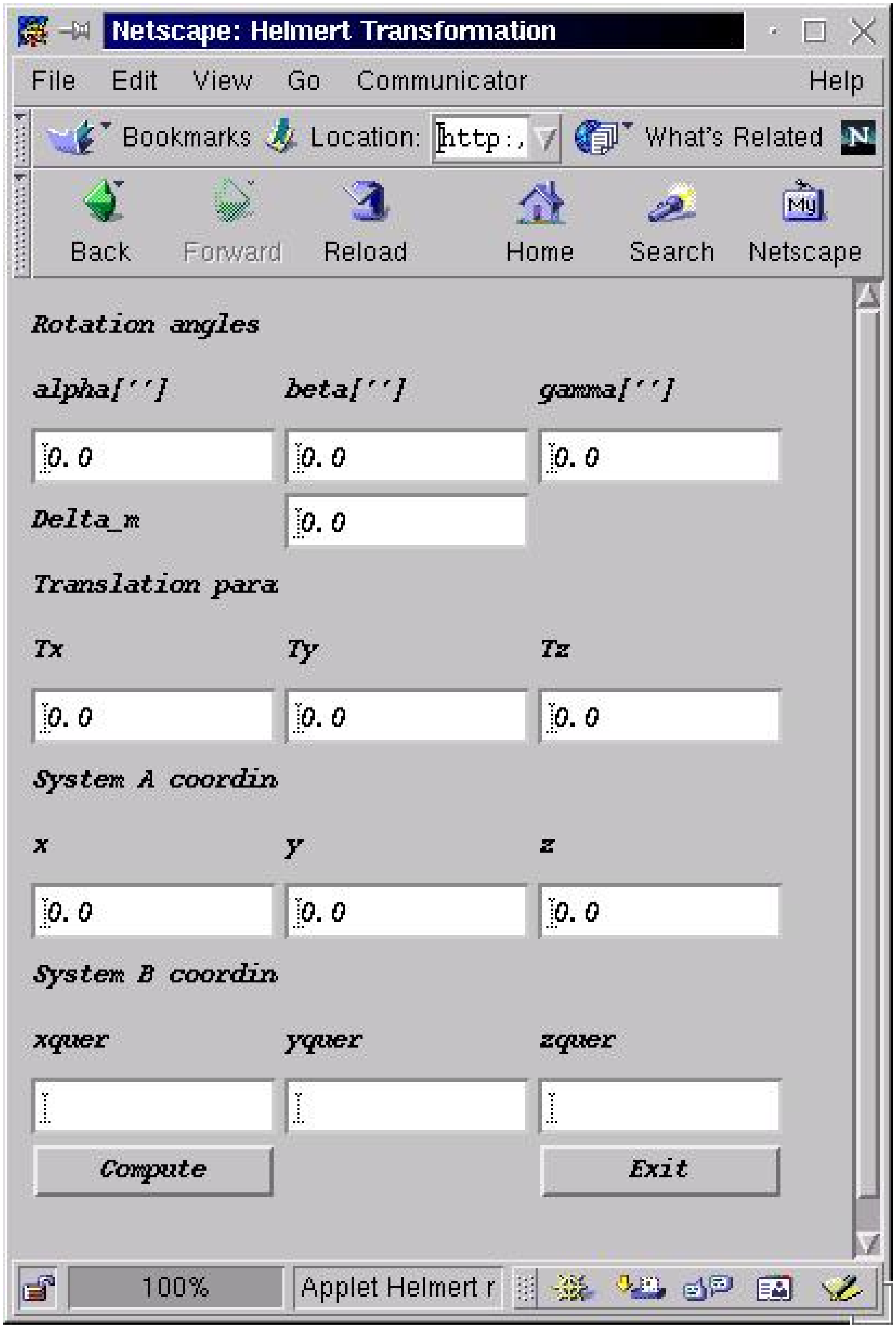
are known.
In most cases the rotation angles between two coordinate systems are small and for this reason the transformation matrix can be approximated either by

$$\mathbf{T} = m \begin{bmatrix} 1 & \delta\gamma & -\delta\beta \\ -\delta\gamma & 1 & \delta\alpha \\ \delta\beta & -\delta\alpha & 1 \end{bmatrix} \tag{4.6}$$

or

$$\mathbf{T} = m \begin{bmatrix} 1 & (\delta\varphi + \delta\psi) & 0 \\ -(\delta\varphi + \delta\psi) & 1 & \delta\vartheta \\ 0 & -\delta\vartheta & 1 \end{bmatrix} \tag{4.7}$$

For given coordinates in a system A and for given transformation parameters between the system A and a new system B click here to start an applet.


Netscape: Helmert Transformation
File Edit View Go Communicator Help
Bookmarks Location: http:/ What's Related
Back Forward Reload Home Search Netscape
Rotation angles
alpha[''] beta[''] gamma['']
0.0 0.0 0.0
Delta_m 0.0
Translation para
Tx Ty Tz
0.0 0.0 0.0
System A coordin
x y z
0.0 0.0 0.0
System B coordin
xquer yquer zquer
Compute
Exit
100%
Applet Helmert r

In most cases the transformation parameters between two cartesian systems are not previously known. The can be determined by the comparison of the coordinates of identical points in both coordinate systems.

4.3 Determination of the transformation Parameters

Let us assume that for at least three points their coordinates (x_i, y_i, z_i) with respect to the cartesian coordinate system $\{\mathcal{O}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ as well as their coordinates $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ with respect to the coordinate system $\{\bar{\mathcal{O}}, \bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3\}$ are known. Unknown are

- the translation parameters (t_1, t_2, t_3) between the origins of the two systems,
- the rotation angle $\delta\alpha, \delta\beta, \delta\gamma$ between the axis of the two systems and
- the difference δm between 1.0 and the scale ratio of the two systems.

The unknown transformation parameters and the known coordinates are in the following relationship

$$\begin{bmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{z}_i \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} + (1 + \delta m) \cdot \begin{bmatrix} 1 & \delta\gamma & -\delta\beta \\ -\delta\gamma & 1 & \delta\alpha \\ \delta\beta & -\delta\alpha & 1 \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \tag{4.8}$$

This a partly linear and partly nonlinear system of equations for the unknown transformation parameters. Nonlinear because the the products $\delta m\delta\alpha, \delta m\delta\beta, \delta m\delta\gamma$ between the unknown parameters occur in these equations. Since both the parameters δm and $\delta\alpha, \delta\beta, \delta\gamma$ are small their products can be neglected. The simplified equations are

$$\begin{bmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{z}_i \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} + \begin{bmatrix} 1+\delta m & \delta\gamma & -\delta\beta \\ -\delta\gamma & 1+\delta m & \delta\alpha \\ \delta\beta & -\delta\alpha & 1+\delta m \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \tag{4.9}$$

The equations (4.9) are now linear with respect to the unknown transformation parameters. A re-arrangement leads to the usual matrix form of an overdetermined linear system of equations.

$$\underbrace{\begin{bmatrix} \bar{x}_1 - x_1 \\ \bar{y}_1 - y_1 \\ \bar{z}_1 - z_1 \\ \bar{x}_2 - x_2 \\ \bar{y}_2 - y_2 \\ \bar{z}_2 - z_2 \\ \vdots \end{bmatrix}}_{\mathbf{b}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & -z_1 & y_1 & x_1 \\ 0 & 1 & 0 & z_1 & 0 & -x_1 & y_1 \\ 0 & 0 & 1 & -y_1 & x_1 & 0 & z_1 \\ 1 & 0 & 0 & 0 & -z_2 & y_2 & x_2 \\ 0 & 1 & 0 & z_2 & 0 & -x_2 & y_2 \\ 0 & 0 & 1 & -y_2 & x_2 & 0 & z_2 \\ & & & \vdots & & & \end{bmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ \delta\alpha \\ \delta\beta \\ \delta\gamma \\ \delta m \end{bmatrix}}_{\mathbf{x}} \tag{4.10}$$

The unknown transformation parameter $\mathbf{x}$ are the usual least-squares solution of the overdetermined system of equations

$$\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b} \tag{4.11}$$

Click here to start an applet for 3D Helmert transformation

Besides the general three-dimensional case there is a much simpler solution for the determination of the transformation parameters in the two-dimensional special case. In two dimensions the relations between the coordinates of a point with respect to two different systems are

$$\begin{bmatrix} \bar{x}_i \\ \bar{y}_i \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} (m\cos\alpha) & (m\sin\alpha) \\ -(m\sin\alpha) & (m\cos\alpha) \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \end{bmatrix} \tag{4.12}$$

In contrast to the three-dimensional case there is only one rotation angle α and this angle is not restricted in size. The nonlinearity of the problem can be eliminated by introducing

$$o := -m\sin\alpha \quad , \quad a := m\cos\alpha \tag{4.13}$$

This leads to a linear problem for o, a

$$\begin{bmatrix} \bar{x}_i \\ \bar{y}_i \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} a & -o \\ o & a \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \end{bmatrix} \tag{4.14}$$

From the solution a, o of this linear system the original unknowns can be recovered by

$$m = \sqrt{a^2 + o^2} \quad , \quad -\alpha = \arctan(\frac{o}{a}) \tag{4.15}$$

For the determination of the four unknown parameters a, o, α_0, β_0 four equations, i.e. two identical points are necessary.

As a first step for the determination of the transformation parameters the translation parameters are eliminated from the transformation equations

$$\begin{bmatrix} \bar{x}_2 - \bar{x}_1 \\ \bar{y}_2 - \bar{y}_1 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 & -(y_2 - y_1) \\ y_2 - y_1 & x_2 - x_1 \end{bmatrix} \cdot \begin{bmatrix} a \\ o \end{bmatrix} \tag{4.16}$$

The equations (4.16) have a unique solution if the determinate does not vanish, i.e. if

$$0 \neq (x_2 - x_1)^2 + (y_2 - y_1)^2 =: s_{12}^2 \tag{4.17}$$

,which is automatically fulfilled as long as the two identical points do not coincide.

Schreibers rule gives an explicit expression for the solution

$$\begin{bmatrix} a \\ o \end{bmatrix} = \frac{1}{s_{12}^2} \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ -(y_2 - y_1) & x_2 - x_1 \end{bmatrix} \cdot \begin{bmatrix} \bar{x}_2 - \bar{x}_1 \\ \bar{y}_2 - \bar{y}_1 \end{bmatrix} \tag{4.18}$$

Inserting the solution a, o into (4.12) and solving for α_0, β_0 gives the remaining parameters:

$$\begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} \bar{x}_i \\ \bar{y}_i \end{bmatrix} - \begin{bmatrix} x_i & -y_i \\ y_i & x_i \end{bmatrix} \cdot \begin{bmatrix} a \\ o \end{bmatrix} \tag{4.19}$$

Of course, this algorithm is applicable only, if the number of identical points equals two. The disadvantage is that every error contained in the coordinates of the identical points propagates uncontrolled into the derived transformation parameters. The usual measure for error control in Geodesy is overdetermination. This means that more than two identical points are used for the determination of the four transformation parameters. In this way, errors that are contained in the coordinates of one of the identical point can be identified by intolerable residuals for this point.

Let us assume that $(x_i, y_i) \quad , \quad i = 1, \ldots, n$ are coordinates of n points with respect to a cartesian System A and $(\bar{x}_i, \bar{y}_i) \quad , \quad i = 1, \ldots, n$ are the coordinates of the same points with respect to another cartesian system B. If the transformation parameters α_0, β_0, a, o from System A to system B were already known, the transformed coordinates would be

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix} + \begin{bmatrix} x_i & -y_i \\ y_i & x_i \end{bmatrix} \cdot \begin{bmatrix} a \\ o \end{bmatrix} \quad , \quad i = 1, \ldots, n \tag{4.20}$$

The residuals

$$\xi_i = \bar{x}_- x_i' \quad , \quad \eta_i = \bar{y}_i - y_i' \quad , \quad i = 1, \ldots, n \tag{4.21}$$

are the differences between the coordinates of a point in system B and and the transformed coordinates of the same point. If the transformation parameters were perfect, all residuals should of course vanish. Since there are always some small errors in the coordinates in practice they will not.
Therefore, a reasonable strategy for the choice of the transformation parameters is to chose them in such a way that the square sum of the residuals gets minimal.

$$\Phi := \sum_{i=1}^{n} (\xi_i^2 + \eta_i^2) \rightarrow \min \tag{4.22}$$

The necessary conditions for an extremal point are

$$0 = \frac{\partial \Phi}{\partial \alpha_0} = 2 \sum_{i=1}^{n} \xi_i \frac{\partial \xi_i}{\partial \alpha_0} + \eta_i \frac{\partial \eta_i}{\partial} \alpha_0 \tag{4.23}$$

$$= -2 \sum_{i=1}^{n} \bar{x}_i - \alpha_0 - x_i \cdot a + y_i \cdot o \tag{4.24}$$

$$0 = \frac{\partial \Phi}{\partial \beta_0} = 2 \sum_{i=1}^{n} \xi_i \frac{\partial \xi_i}{\partial \beta_0} + \eta_i \frac{\partial \eta_i}{\partial} \beta_0 \tag{4.25}$$

$$
\begin{aligned}
&= -2\sum_{i=1}^{n} \bar{y}_i - \beta_0 - y_i \cdot a - x_i \cdot o && (4.26)\\
0 = \frac{\partial \Phi}{\partial a} &= 2\sum_{i=1}^{n} \xi_i \frac{\partial \xi_i}{\partial a} + \eta_i \frac{\partial \eta_i}{\partial a} && (4.27)\\
&= 2\sum_{i=1}^{n} (\bar{x}_i - \alpha_0 - x_i \cdot a + y_i \cdot o)(-x_i) && (4.28)\\
&\quad + (\bar{y}_i - \beta_0 - y_i \cdot a - x_i \cdot o)(-y_i) && (4.29)\\
0 = \frac{\partial \Phi}{\partial o} &= 2\sum_{i=1}^{n} \xi_i \frac{\partial \xi_i}{\partial o} + \eta_i \frac{\partial \eta_i}{\partial o} && (4.30)\\
&= 2\sum_{i=1}^{n} (\bar{x}_i - \alpha_0 - x_i \cdot a + y_i \cdot o)(y_i) && (4.31)\\
&\quad + (\bar{y}_i - \beta_0 - y_i \cdot a - x_i \cdot o)(-x_i) && (4.32)
\end{aligned}
$$

A summation over all n points gives the linear system of equations

$$
\begin{bmatrix} n & 0 & \sum x_i & -\sum y_i \\ 0 & n & \sum y_i & \sum x_i \\ \sum x_i & \sum y_i & \sum(x_i^2 + y_i^2) & 0 \\ -\sum y_i & \sum x_i & 0 & \sum(x_i^2 + y_i^2) \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \beta_0 \\ a \\ o \end{bmatrix} = \begin{bmatrix} \sum \bar{x}_i \\ \sum \bar{y}_i \\ \sum(x_i \bar{x}_i +_i \bar{y}_i) \\ \sum(x_i \bar{y}_i - y_i \bar{x}_i) \end{bmatrix} \tag{4.33}
$$

The equations (4.33) can be simplified, if the origins of the systems A and B are shifted into the mean values of the given coordinates: Let be

$$
x_s := \frac{1}{n}\sum_{i=1}^{n} x_i, \quad y_s := \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \bar{x}_s := \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i, \quad \bar{y}_s := \frac{1}{n}\sum_{i=1}^{n} \bar{y}_i \tag{4.34}
$$

The differences between the individual coordinates and the coordinate mean values are

$$
u_i := x_i - x_s, \quad v_i := y_i - y_s, \quad \bar{u}_i := \bar{x}_i - \bar{x}_s, \quad \bar{v}_i := \bar{y}_i - \bar{y}_s \tag{4.35}
$$

With respect to these variables the equations (4.33) obtain the form

$$
\begin{bmatrix} n & 0 & nx_s & -ny_s \\ 0 & n & ny_s & nx_s \\ nx_s & ny_s & n(x_s^2 + y_s^2) + \sum(u_i^2 + v_i^2) & 0 \\ -ny_s & nx_s & 0 & n(x_s^2 + y_s^2) + \sum(u_i^2 + v_i^2) \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \beta_0 \\ a \\ o \end{bmatrix} \tag{4.36}
$$

$$
= \begin{bmatrix} n\bar{x}_s \\ n\bar{y}_s \\ n(x_s \bar{x}_s + y_s \bar{y}_s) + \sum(u_i \bar{u}_i + v_i \bar{v}_i) \\ n(x_s \bar{y}_s - y_s \bar{x}_s) \sum(u_i \bar{v}_i - v_i \bar{u}_i) \end{bmatrix}
$$

Multiplication of the first equation with x_s and of the second equation with y_s and subtraction from the third and fourth equation, respectively , yields

$$\left[\begin{array}{cc} \sum(u_1^2+v_i^2) & 0 \\ 0 & \sum(u_i^2+v_i^2) \end{array}\right] \cdot \left[\begin{array}{c} a \\ o \end{array}\right] == \left[\begin{array}{c} \sum(u_i\bar{u}_i+v_i\bar{v}_i) \\ \sum(u_i\bar{v}_i-v_i\bar{u}_i) \end{array}\right] \tag{4.37}$$

The solution of (4.37) can be given explicitely

$$a = \frac{\sum u_i\bar{u}_i+v_i\bar{v}_i}{\sum(u_i^2+v_i^2)}, \quad o = \frac{\sum u_i\bar{v}_i-v_i\bar{u}_i}{\sum(u_i^2+v_i^2)} \tag{4.38}$$

Inserting this in (4.36) the remaining translation parameters can be obtained

$$\alpha_0 = \bar{x}_s - ax_s + oy_s, \quad \beta_0 = \bar{y}_s - ay_s - ox_s \tag{4.39}$$

Please note, that in contrast to the three-dimensional case the two-dimensional Helmert transformation does not imply any restriction to the rotation angle.

For a JAVA Applet, performing a two-dimensional Helmert transformation click here

Chapter 5

Curvilinear Coordinate Systems

5.1 Coordinate lines

Coordinates of a point in the thee-dimensional Euclidean space are defined as the triple (x_1, x_2, x_3) of real number, which are in a unique way assigned to this point. There are of course different ways of such an assignment an each assignment is called a coordinate system.

If two of the three coordinates are held fixed and the remaining coordinate runs through the real numbers the corresponding points, addressed by the varying coordinates, describes a curve in the Euclidean space: A so called coordinate line. Depending of the curvature of the coordinate lines the coordinate system is either Cartesian or curvilinear.

Definition 18 *Let $K : E^3 \to \mathbb{R}^3$ a coordinate system with the coordinates (x_1, x_2, x_3). The set of points*

$$L1 := \{(x_1, x_2, x_3) | x_2 = const, x_3 = const\} \tag{5.1}$$
$$L2 := \{(x_1, x_2, x_3) | x_1 = const, x_3 = const\} \tag{5.2}$$
$$L3 := \{(x_1, x_2, x_3) | x_1 = const, x_2 = const\} \tag{5.3}$$

are called the x_1, x_2 and the x_3 coordinate lines in the coordinate system K.

Definition 19 *A coordinate system with at least one coordinate not being a straight line is called a curvilinear coordinate system.*

The simplest way to define a curvilinear coordinate system is, to relate it to an existing Cartesian coordinate system. Let (x_1, x_2, x_3) be Cartesian coordinates in E^3. A bijective mapping $U : \mathbb{R} \to \mathbb{R}^3$ generates new, in general curvilinear coordinates by

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = U(x_1, x_2, x_3) \tag{5.4}$$

In the following some curvilinear coordinate systems, typical for Geodesy, will be described.

5.2 Spherical Coordinates

Spherical coordinates are usually denoted by r, ϑ, λ and related to Cartesian coordinates by

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2}, \quad \vartheta = \arccos(x_3/r), \quad \lambda = arctan(x_2/x_1) \tag{5.5}$$

The inverse transformation from spherical to Cartesian coordinates is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = r \begin{bmatrix} \sin\vartheta \cdot \cos\lambda \\ \sin\vartheta \cdot \sin\lambda \\ \cos\vartheta \end{bmatrix} \tag{5.6}$$

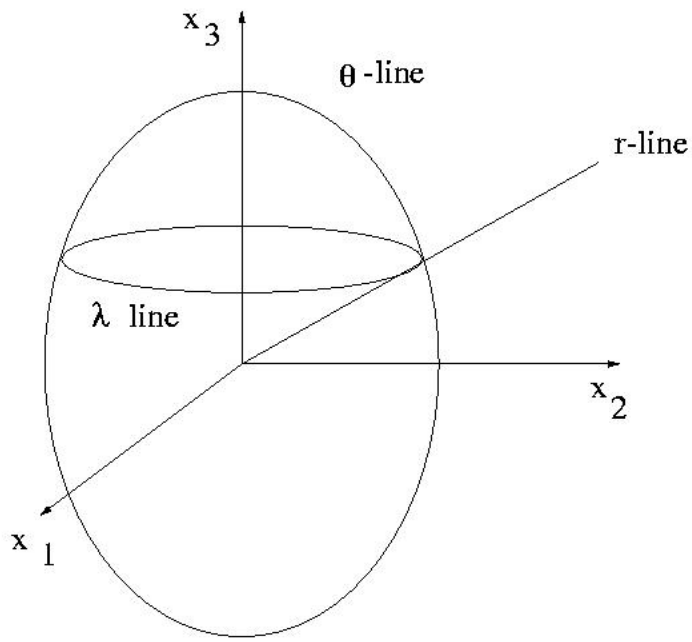
The coordinate lines are

- $r - line$: straight line through the origin of the Cartesian coordinate system.

- $\lambda - line$ circles parallel to the $x_1 - x_2$ plane.

- $\vartheta - line$ circles having the x_3 axis as diameters.

The coordinate lines for spherical coordinates are displayed in the following figure

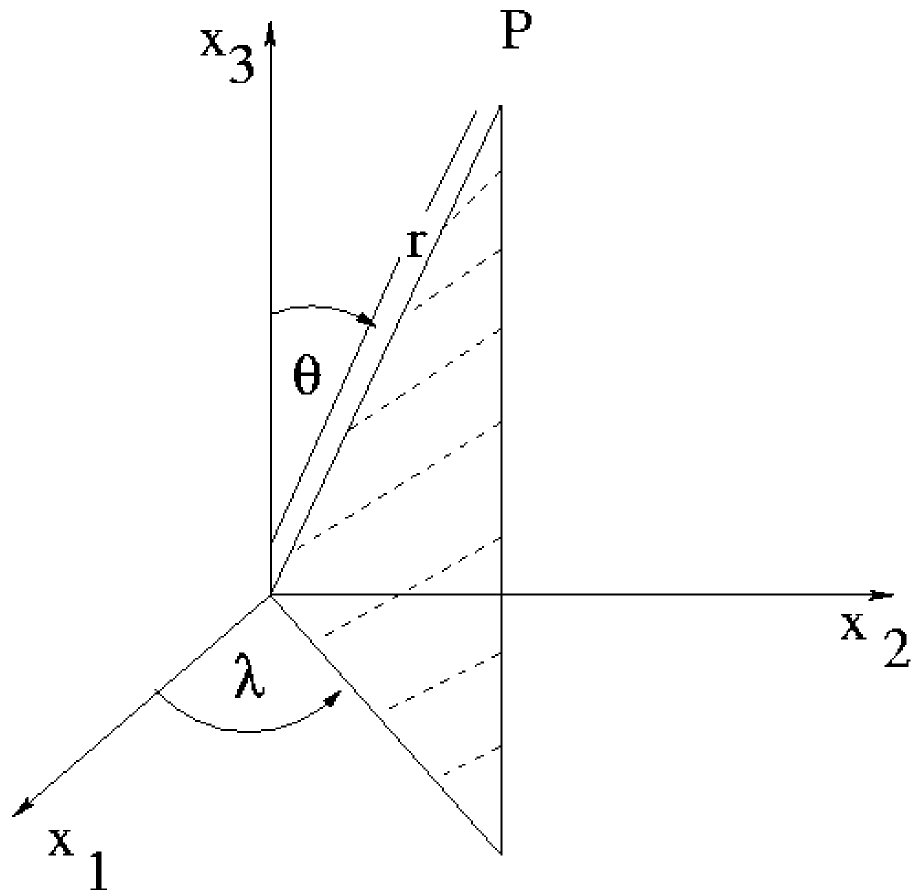


The interpretation of spherical coordinates is quit simple:

- r is the distance from the origin of the Cartesian coordinate system.

- ϑ is the angle between the x_3 axis of the Cartesian coordinate system and the straight line connecting the point with the origin of the coordinate system.

- λ is the angle between the $x_1 - x_3$ plane and the plane containing the x_3 axis and the point.

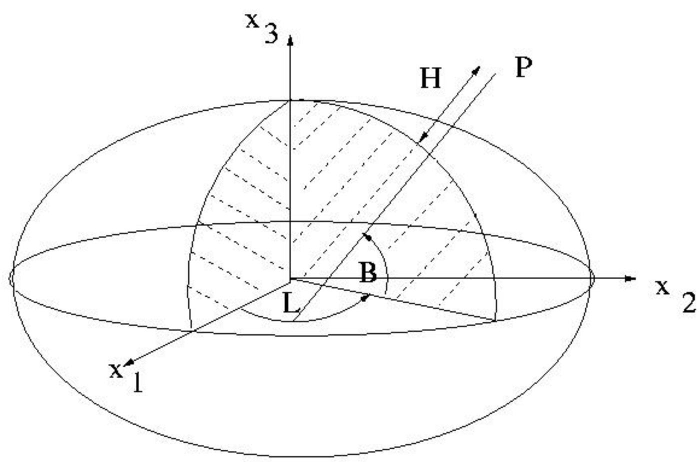


If the Earth were spherical the coordinates at its surface would only change slightly in the r component. The changes in the r- component were exclusively due to the changes in the height. In reality the Earth is flattened at its poles. This means that the r component changes by about 20 km due to the flattening. This exceeds the r-changes due to the hight changes considerably. Therefore, a coordinate system with the third coordinate not changing along the surface of the flattened Earth would be very useful. Such a coordinate system is the ellipsoidal coordinate system.

5.3 Ellipsoidal Coordinates

Ellipsoidal coordinates are based upon an ellipsoid of revolution, centered at the origin of a Cartesian coordinate system and having its rotation axis to coincide with the x_3 axis of the Cartesian system.

The definition of ellipsoidal coordinates can be explained best geometrically.



Definition 20 *The ellipsoidal height H is the distance to the surface of the ellipsoid.*
The ellipsoidal latitude B is the angle between the straight line connecting the point with the origin and the $x_1 - x_2$ plane.
The ellipsoidal longitude L is the angle between the $x_1 - x_3$ plane and the plane containing the x_3 axis and the point.

The computation of Cartesian coordinates from given ellipsoidal coordinates is straightforward

$$N = \frac{a}{\sqrt{1 - e^2 \sin^2(B)}} \tag{5.7}$$

$$x_1 = (N + H)\cos(B)\cos(L) \tag{5.8}$$

$$x_2 = (N + H)\cos(B)\sin(L) \tag{5.9}$$

$$x_3 = (N(1 - e^2) + H)\sin(B) \tag{5.10}$$

The connection between ellipsoidal and Cartesian coordinates is more complicated than in the spherical case. This back-transformation is given by the following set of equations

$$b = a \cdot \sqrt{1 - e^2} \tag{5.11}$$

$$e_1 = 1 - \frac{b^2}{a^2} \tag{5.12}$$

$$e_2 = \frac{a^2}{b^2} - 1 \tag{5.13}$$

$$e_3 = \sqrt{x_1^2 + x_2^2} \tag{5.14}$$

$$F = 54 \cdot b^2 \cdot x_3^2 \tag{5.15}$$

$$G = r^2 + (1 - e_1) \cdot x_3^2 - e_1 \cdot e_3 \tag{5.16}$$

$$c = e_1^2 \cdot F \frac{r^2}{G^3} \tag{5.17}$$

$$s_1 = \sqrt{c^2 + 2c} \tag{5.18}$$

$$s_2 = 1 + c + s_1 \tag{5.19}$$

$$s = \sqrt[3]{s2} \tag{5.20}$$

$$p_1 = (s + \frac{1}{s} + 1)^2 \tag{5.21}$$

$$p_2 = 3p_1 \cdot G^2 \tag{5.22}$$

$$p = \frac{F}{p_2} \tag{5.23}$$

$$Q = \sqrt{1 + 2e_1^2 p} \tag{5.24}$$

$$\rho_1 = \frac{1}{2} a^2 (1 + \frac{1}{Q}) \tag{5.25}$$

$$\rho_2 = \frac{1}{2} p r^2 \tag{5.26}$$

$$\rho_3 = \frac{p(1 - e_1) x_3^2}{Q(1 + Q)} \tag{5.27}$$

$$\rho_4 = \sqrt{\rho_1 - \rho_2 - \rho_3} \tag{5.28}$$

$$\rho_5 = \frac{P e_1 r}{1 + Q} \tag{5.29}$$

$$\rho = \rho 4 - \rho 5 \tag{5.30}$$

$$U = \sqrt{(r - e_1 \rho)^2 + (1 - e_1) x_3^2} \tag{5.31}$$

$$z_0 = \frac{b^2 x_3}{aV} \tag{5.32}$$

$$h_1 = \frac{b^2}{aV} \tag{5.33}$$

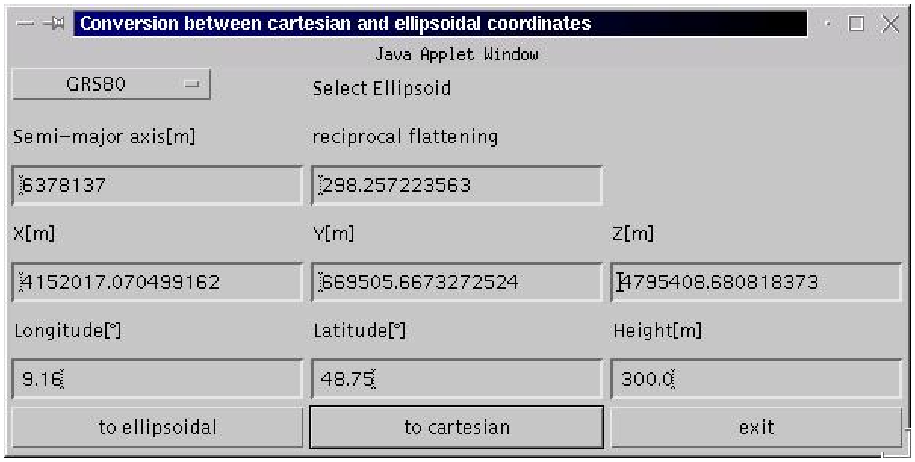
$$bt = \sqrt{r^2 + (x_3 + e_2 z_0)^2} \tag{5.34}$$

$$H = U(1 - h_1) \tag{5.35}$$

$$B = \arccos(\frac{r}{bt}) \tag{5.36}$$

$$L = \arccos(\frac{x_1}{r}) \tag{5.37}$$

The JAVA Applet providing a tool for the conversion between Cartesian and ellipsoidal coordinates



can be invoked by clicking here.

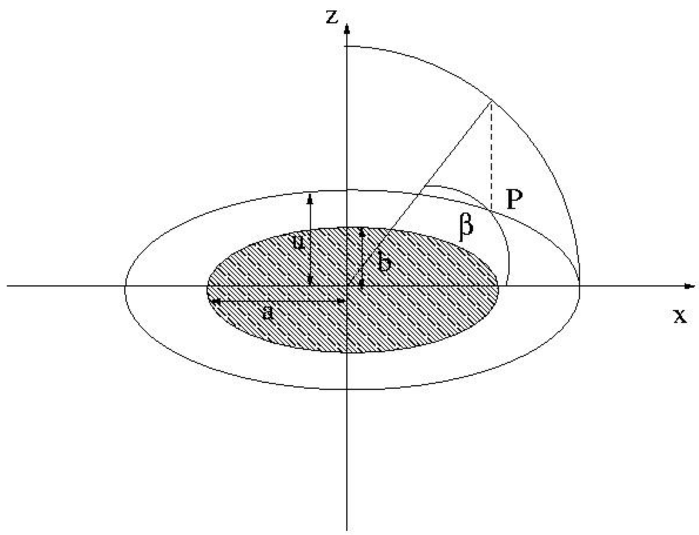
The ellipsoid, underlying an ellipsoidal coordinate system does not only serve as a model for the figure of the Earth but also as a model of the Earth's gravity field. Therefore, a gravity field model has to be assigned to a given ellipsoid in such a way that a certain equipotential surface of this model coincides with the surface of the ellipsoid. After specifying the value for GM , the product of gavitational constant and mass of the Earth and fixing the value for the rotation rate ω of the Earth this so-called normal potential is given by the *Somigliani-Pizetti* formula:

$$U = \frac{GM}{\epsilon} arctan(\frac{\epsilon}{u}) + \frac{\omega^2}{2} a^2 \frac{q}{q_0} (\sin^2 \beta - \frac{1}{3}) \tag{5.38}$$

with

$$q = \frac{1}{2}((1 + 3\frac{u^2}{\epsilon^2}) arctan(\frac{u}{\epsilon}) - 3\frac{epsilon}{u}), \quad q_0 = q_{u=b} \tag{5.39}$$

The meaning of the coordinates u, ϵ, β can be read from the following figure.



The norm of the gradient of U is called normal gravity γ. It si given by the normal gravity formula

$$\gamma = \frac{a\gamma_e \cos^2\varphi + b\gamma_p \sin^2\varphi}{\sqrt{a^2\cos^2\varphi + b^2\sin^2\varphi}} \tag{5.40}$$

with γ_e, γ_p being the normal gravity at the equator and at the pole.

Chapter 6

Map Coordinates

For historical but also for practical reasons the description of a position of a point can be separated into the description of its horizontal and its vertical position. Historically, this distinction was motivated by different measurement technologies in the horizontal and in the vertical domain. Practically, this separation is useful since along the Earth's surface the variation of horizontal coordinates is many times larger than the variation of the vertical coordinates.
Coordinate systems, which naturally separate the horizontal and the vertical description of a position from its horizontal description are the spherical and the ellipsoidal system of coordinates. There λ, θ or L, B describe the horizontal position and r or H describe the vertical position, respectively.
Unfortunately, vertical coordinate surfaces are not planes but curved surfaces. Therefore, the simple formulas of planar Euclidean geometry are not longer applicable for geodetic computations with horizontal coordinates. This would have been replaced by the much more complicated relation of differential geometry.
One compromise is to map the horizontal coordinate surfaces into the Euclidean plane and to perform the computations with the plane coordinates of the mapped points. The Euclidean coordinates of the mapped surface are called map coordinates.
Any mapping of a surface with non vanishing Gaussian curvature into the plane is connected with the distortion of geometric quantities as distances, angles, directions. This distortion is accounted for in two ways:

- The horizontal coordinate surfaces are mapped only piece-wise, in order to keep the unavoidable distortions for every piece.
- The geometric quantities derived from map-coordinates have to undergo certain corrections before they can enter geodetic computations.

In the pre-GPS time the differences in the vertical coordinates could not be measured directly. Therefore, the vertical coordinate was replaced by a closely related quantity which could be observed by geodetic measurements. Since this could be done in several different ways several height systems are currently still in use.

6.1 Conformal mappings

Definition 21 *A injective mapping $M : U \subset S \to \mathbb{R}^2$ of a part U of a vertical coordinate surface S into the pairs of real numbers is called a map-projection. The numbers, which a point P is mapped onto are called map- coordinates of P*

Let P and $P+dP$ be two points in $U \subset S$ having the coordinates $u_i, i = 1,2$ and $u_i + du_i, i = 1,2$ respectively. Let be (x,y) and $(x+dy, y+dy)$ the coordinates of of their images under the map-projection M.

The the square of the infinitesimal distance of P and $P+dP$ is

$$dS^2 = \sum_{i,j=1}^{2} g_{ij} du_i du_j \tag{6.1}$$

with g_{ij} being the metric tensor of the horizontal coordinate surface S. The square of the infinitesimal distance of the images of these points is

$$ds^2 = dx^2 + dy^2 = \sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) du_i du_j \tag{6.2}$$

Definition 22 *The quantity*

$$m := \frac{ds}{dS} = \sqrt{\frac{\sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) du_i du_j}{\sum_{i,j=1}^{2} g_{i,j} du_i du_j}} \tag{6.3}$$

is the length-distortion of the map-projection M.

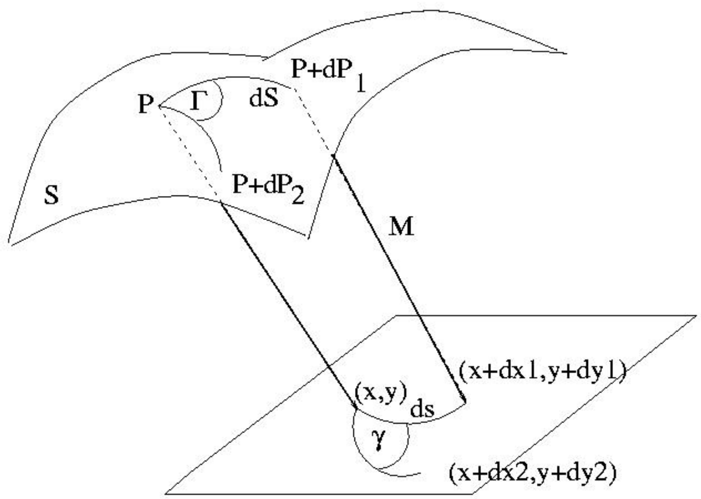
Let now be $P, P+dP^1, P+dp^2$ be three infinitesimal close points in the horizontal coordinate plane S. Let $u_i, u_i + du_i^1, u_i + du_i^2$ be the coordinates of these points and let $(x,y), (x+dx^1, y+dy^1), (x+dx^2, y+dy^2)$ be the coordinates of their images. The angle between these thee points is

$$\cos\Gamma = \frac{\sum_{i,j=1}^{2} g_{ij} du_i^1 du_j^2}{\sqrt{\sum_{i,j=1}^{2} g_{ij} du_i^1 du_j^1}\sqrt{\sum_{i,j=1}^{2} g_{ij} du_i^2 du_j^2}} \tag{6.4}$$

and the angle between their images is

$$\cos\gamma = \frac{dx^1 dx^2 + dy^1 dy^2}{\sqrt{(dx^1)^2 + (dy^1)^2}\sqrt{(dx^2)^2 + (dy^2)^2}} \tag{6.5}$$

$$= \frac{\sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) du_i^1 du_j^2}{\sqrt{\sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) du_i^1 du_j^1}\sqrt{\sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) du_i^2 du_j^2}} \tag{6.6}$$



Definition 23 *The ratio*

$$\delta\gamma := \frac{\gamma}{\Gamma} \tag{6.7}$$

is called the angle distortion of the map projection M.

Definition 24 *A map projection is called*

- *length-preserving, im $m = 1$ and*
- *angle-preserving,or conformal if $\delta\gamma = 1$*

Theorem 16 *A map-projection M is*

- *length-preserving, if*

$$g_{ij} = \sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) \tag{6.8}$$

- *angle-preserving, if*

$$g_{ij} = \lambda(u_1, u_2) \cdot \sum_{i,j=1}^{2} \left(\frac{\partial x}{\partial u_i}\frac{\partial x}{\partial u_j} + \frac{\partial y}{\partial u_i}\frac{\partial y}{\partial u_j} \right) \tag{6.9}$$

Every length-preserving map-projection is also angle-preserving.

Neither for a horizontal coordinate surface of the spherical coordinate system nor for the horizontal coordinate surface of an ellipsoidal coordinate system a length-preserving map projection is possible. Therefore, frequently used map projections are only angle preserving. In order to keep the unavoidable length-distortion small the coordinate surface is not mapped as a whole but piecewise.
Now some of the most important map projections will be discussed.

6.2 Gauß-Krüger projection

This map projection was first used by the mathematician Gauß and the geodesist Krüger. It is one of the most frequently map projections in Europe. The horizontal coordinate surface for the Gauß-Krüger map projection is an ellipsoid of a certain size and flattening. This surface is divided in several zones included between Meridians. Each of the zones is mapped conformally into the map-plane.
This means a Gauß-Krüger projection is defined by

1. The ellipsoid a, f which defines the ellipsoidal coordinate system.
2. The width of the zones.
3. the central meridian L_0 of each zone
4. false easting E_0

The central meridian is mapped length-preserving to the x-axis of the map-coordinate system and the intersection of the central meridian and the equator is mapped to the origin of the map-coordinate system. In order to avoid negative coordinate values in the map-plane the coordinates (x, y) are replaced by the values N, E (northing, and easting) in the following way:

$$N = x, \quad E = y + E_0 \tag{6.10}$$

The transformation from ellipsoidal coordinates L, B into Gauß-Krüger coordinates N, E can be achieved by the following set of formulae:

Definition of reference point on the central meridian

$$(L_0, B_0) \quad B_0 \approx B \Rightarrow \Delta B = B - B_0, \quad \Delta L = L - L_0 \tag{6.11}$$

Meridional arc length

$$(e')^2 = \frac{f(2-f)}{(1-f)^2} \tag{6.12}$$

$$\begin{aligned}
a_0 &= 1 - ((e')^2 + 3\frac{(e')^4}{16} + 5\frac{(e')^6}{64} + 175\frac{(e')^8}{4096})/4 \\
a_2 &= 3((e')^2 + \frac{(e')^4}{4} + 15\frac{(e')^6}{128} - 455\frac{(e')^8}{4096})/8 \\
a_4 &= 15((e')^4 + 3\frac{(e')^6}{4} - 77\frac{(e')^8}{128})/256 \\
a_6 &= 35((e')^6 - 41\frac{(e')^8}{32})/3072 \\
a_8 &= -315\frac{(e')^8}{131072} \\
G &= a \cdot (a_0 B_0 - a_2 \sin(2B_0) + a_4 \sin(4B_0) - a_6 \sin(6B_0))
\end{aligned}$$

auxiliary quantities:

$$\eta^2 = (e')^2 \cos^2(B_0), \quad t = \tan(B), \quad N = \frac{a}{\sqrt{1-(e')^2 \sin^2(B_0)}} \tag{6.13}$$

power series coefficients:

$$\begin{aligned}
a_{10} &= N(1-\eta^2+\eta^4-\eta^6)\\
a_{01} &= N\cos(B_0)\\
a_{20} &= 3Nt(\eta^2-2\eta^4)/2\\
a_{11} &= N\cos(B_0)t(-1+\eta^2-\eta^4)\\
a_{02} &= N\cos^2(B_0)t/2\\
a_{30} &= N\eta^2(1-t^2-2\eta^2+7\eta^2t^2)/2\\
a_{21} &= N\cos(B_0)(-1+\eta^2-3\eta^2t^2-\eta^4+6\eta^4t^2)/2\\
a_{12} &= N\cos^2(B_0)(1+t^2+\eta^2t^2-\eta^4t^2)/2\\
a_{03} &= N\cos^3(B_0)(1-t^2+\eta^2)/6\\
a_{40} &= Nt(-\eta^2)/2\\
a_{31} &= N\cos(B_0)t(1-10\eta^2+3\eta^2t^2)/6\\
a_{22} &= N\cos^2(B_0)t(-4+3\eta^2t^2)/4\\
a_{04} &= N\cos^4(B_0)t(5-t^2+9\eta^2)/24\\
a_{13} &= N\cos^3(B_0)t(-5+t^2-4\eta^2-\eta^2t^2)/6
\end{aligned} \tag{6.14}$$

Gauß- Krüger coordinates

$$\begin{aligned}
x &= G+a_{10}\Delta B\\
&+ a_{20}\Delta B^2+a_{02}\Delta L^2\\
&+ a_{30}\Delta B^3+a_{12}\Delta B\Delta L^2\\
&+ a_{40}\Delta B^4+a_{22}\Delta B^2\Delta l^2+a_{04}\Delta L^4\\
y &= a_{01}\Delta L\\
&+ a_{11}\Delta B\Delta L\\
&+ a_{21}\Delta B^2\Delta L+a_{03}\Delta L^3\\
&+ a_{31}\Delta B^3\Delta L+a_{13}\Delta B\Delta L^3
\end{aligned} \tag{6.15}$$

For the inverse transformation the following set of equations is available:
Definition of a reference point on the central meridian

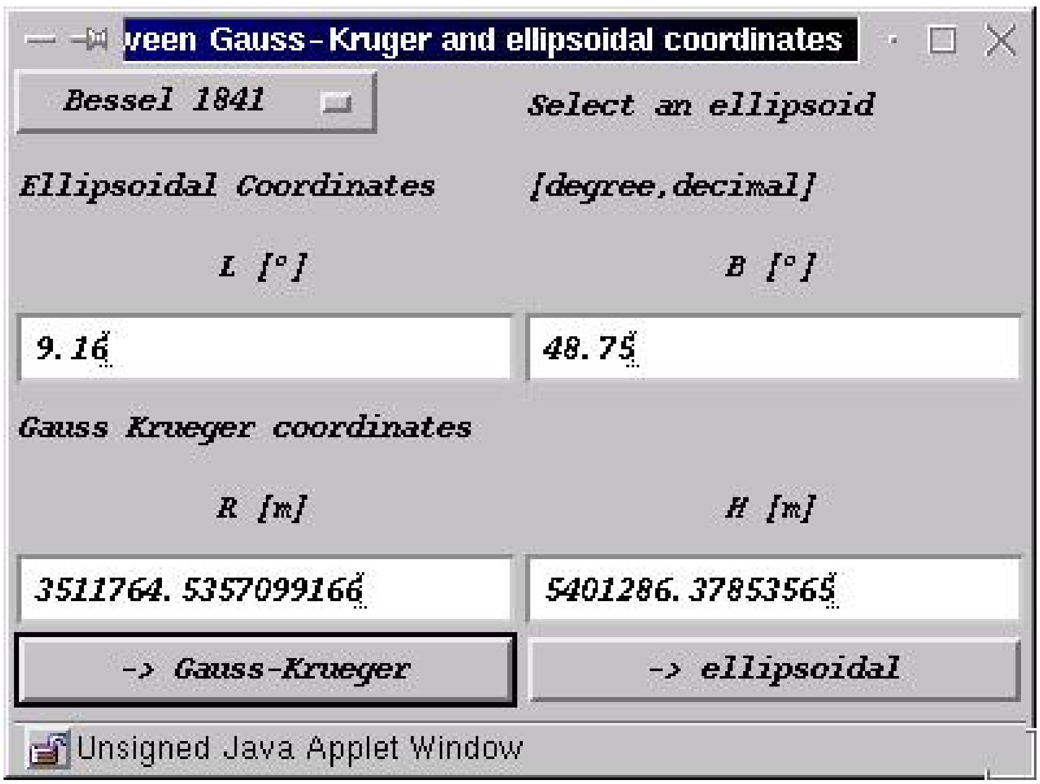
$$(L_0,B_0), B_0\approx B \Rightarrow x_0=G, \Delta x = x-x_0 \tag{6.16}$$

power series coefficients

$$\begin{aligned}
b_{10} &= (1+\eta^2)/N\\
b_{01} &= \frac{1}{N\cos(B_0)}\\
b_{11} &= \frac{t}{N^2\cos(B_0)}\\
b_{20} &= \frac{3\eta^2t}{2N^2}(-1-\eta^2)\\
b_{02} &= \frac{t}{2N^2}(-1-\eta^2)\\
b_{30} &= \frac{\eta^2}{2N^3}(-1+t^2-2\eta^2+6\eta^2t^2)\\
b_{21} &= \frac{1}{2N^3\cos(B_0)}(1+2t^2+\eta^2)\\
b_{03} &= -\frac{1}{3}b_{21}\\
b_{12} &= \frac{1}{2N^3}(-1t^2-2\eta^2+2\eta^2t^2-\eta^4+3\eta^4t^2)\\
b_{40} &= \frac{\eta^2t}{2n^4}\\
b_{31} &= \frac{t}{6N^4\cos(B_0)}(5+6t^2+\eta^2)\\
b_{22} &= \frac{t}{4N^4}(-2-2t^2+9\eta^2+\eta^2t^2)\\
b_{13} &= -b_{31}\\
b_{04} &= \frac{t}{24N^4}(5+3t^2+6\eta^2-6\eta^2t^2)
\end{aligned} \tag{6.17}$$

ellipsoidal coordinates

$$
\begin{aligned}
B &= B_0 + \Delta x \\
&+ b_{20}\Delta x^2 + b_{02}y^2 \\
&+ b_{30}\Delta x^3 + b_{12}\Delta x y^2 \\
&+ b_{40}\Delta x^4 + b_{22}\Delta x^2 y^2 + b_{04}y^4 \\
L &= b_{01}y \\
&+ b_{11}\Delta x y \\
&+ b_{21}\Delta x^2 y + b_{03}y^3 \\
&+ b_{31}\Delta x^3 y + b_{13}\Delta x y^3
\end{aligned}
\tag{6.18}
$$

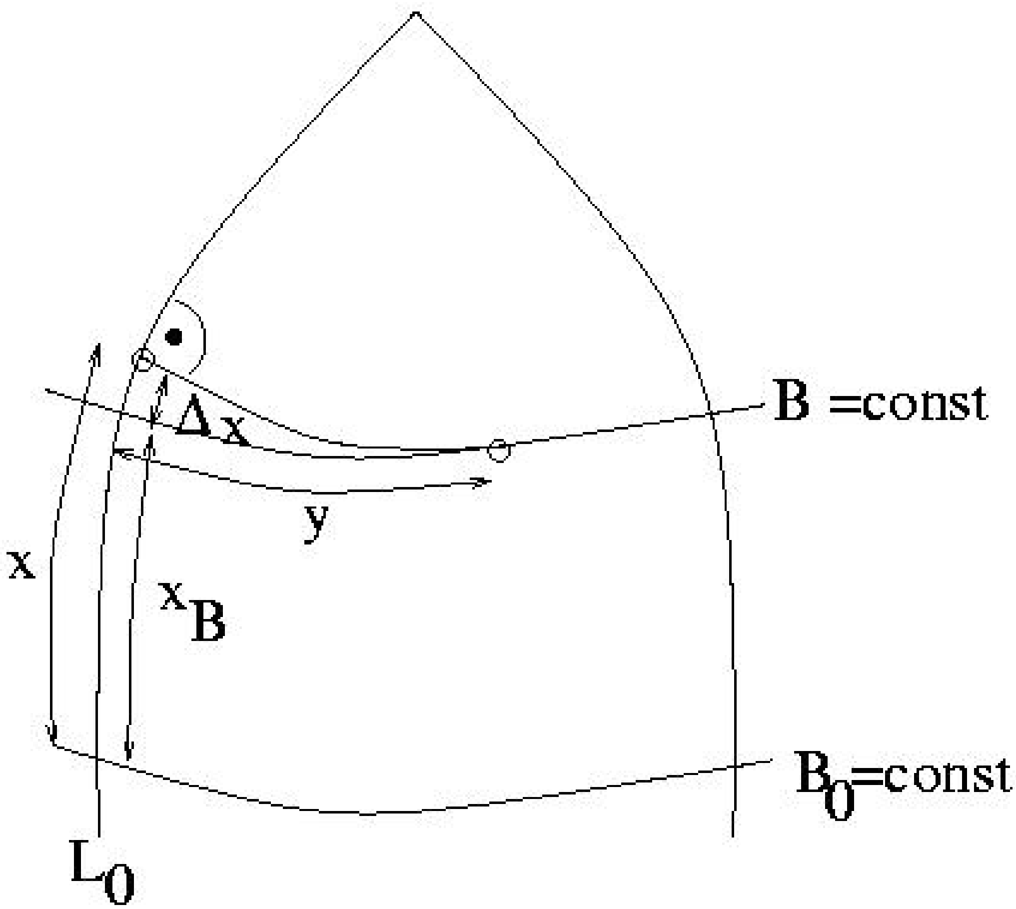


For a JAVA applet converting ellipsoidal coordinates in Gauß-Krüger coordinates click here.

6.3 Soldner coordinates

Soldner coordinates are map coordinates, which were frequently used in the second halve of the 19th century. They do not map a whole meridional strip but only a certain patch of the ellipsoid surface.
For the definition of a particular Soldner map projection a central meridian with the latitude L_0 is chosen, being approximatively in the middle of the area to be mapped

into the map plane. On this central meridian a fundamental point $P_0 = (B_0, L_0)$ is selected, which is approximatively in the center of the area to be mapped.



Definition 25 *Let g be the geodesic through the point $P = (L, B)$, which intersects the central meridian $L = L_0$ orthogonally.*
The length of the geodesic arc between P and the central meridian is the Soldner coordinate y,
The length of the meridional arc between B_0 and the intersection of the geodesic and the central meridian is the Soldner coordinate x of P.

The computation of Soldner coordinates from given ellipsoidal coordinates (L, B) is given by the following set of formulae:

$$\begin{array}{rcl} l & = & L - L_0 \\ x & = & G(B) - G(B_0) + \frac{N}{2}\sin B\cos Bl^2 + \frac{N}{24}\sin B\cos^2(5 - t^2 + 5\eta^2)l^4 \\ y & = & N\cos Bl + \frac{N}{6}\sin^2 B\cos Bl^3 - \frac{N}{120}\sin^2 B\cos^3 B(8 - t^2)l^5 \end{array} \quad (6.19)$$

The inverse transformation from give Soldner coordinates to ellipsoidal coordinates can be accomplished by the set of formulas:
First determine the latitude B_f as the solution of

$$G(B_0) + x = G(B_f) \quad (6.20)$$

The determine $B_f - B$ and $l = L - L_0$ as

$$\begin{array}{rcl} B_f - b & = & \frac{V_f^2}{2N_f^2}t_f y^2 - \frac{V_f^2}{24N_f^4}t_f(1 + 3t_f^2 + \eta_f^2 - 9\eta_f^2 t_f^2)y^4 \\ l & = & \frac{1}{N_f\cos B_f}y - \frac{t_f^2}{3N_f^3\cos B_f}y^3 + \frac{t_f^2}{15N_f\cos B_f}[1 + 3t_f^2)y^5 \end{array} \quad (6.21)$$

In these formula the quantity V is defined by

$$V = \sqrt{1 + (e')^2\cos^2 B} \quad (6.22)$$

The remaining quantities G, N, t, η are defined exactly in the same way as in the section *Gauß - Krüger coordinates*.

6.4 UTM coordinates

Definition 26 *The UTM projection is a Gauß - Krüger projection of that part of the ellipsoid, which is between* $80°$ *S and* $80°$ *N latitude. The meridional zones have a width of* $6°$ *with their central meridians at* $3°, 9°, \ldots, 177°$ *longitude East and West of Greenwich. As length-distortion along the central meridian the value* $m = 0.99996$ *was assigned.*

Chapter 7

Reference systems and Reference Frames

Physical quantities which are observed during a geodetic measurement are

- travel times of electromagnetic waves,
- interference patterns between a received signal and a reference signal,
- interference patterns between a signal received at two different locations, ...

These observed quantities are converted to geometrical quantities like distances and angles. From distances and angles observed between the points of a network coordinates of these points are to be derived. This is only possible,if

- a coordinate system is defined, where the coordinates refer to and
- physical quantities like the vacuum velocity of light are adopted, which allow the conversion of the observed physically quantity into geometric quantity.

Definition 27 *A coordinate system together with a set of parameters, which completely describe the physical model of observations which are to be related to this coordinate system are called a* reference system.

Definition 28 *Let be N a network of points with given coordinates with respect to a coordinate system. A minimal set of parameters, which uniquely define the position and orientation of N in space and the physical model of the observations is called* a datum *of this reference system.*

Obviously, a reference system can have several equivalent datum parameter sets. For instance the orientation and position datum parameters can be given as

1. the position of the origin and the orientation of the axes or
2. as adopted coordinates for a set of points or
3. as transformation parameters which relate this reference system to another reference system.

The definition of a coordinate system can be made in two different ways:

1. **explicit:**
   by describing the the location of the origin and the direction of the axes in relation to some material points.
2. **conventional:**
   by assigning coordinates to a selected number of material points

Example:
Let P, Q, R be three points in the *Euclidean* plane. Then a *Cartesian* co-ordinate system can be defined explicitely by

- letting the origin coincide with the point Q,
- letting the x-axis coincide with the straight line $\vec{QR}$,
- letting the y axis being perpendicular to the x axis and
- defining a length-unit l

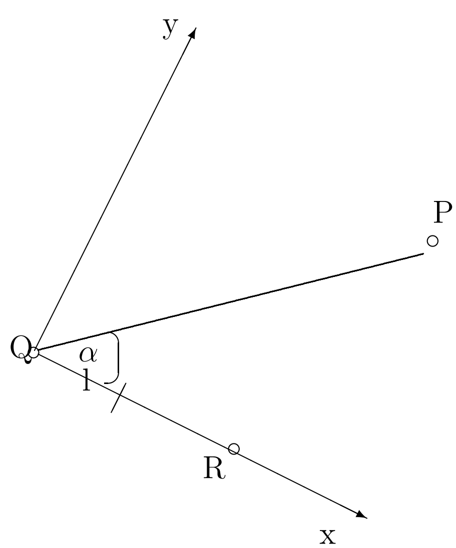
With respect to this co-ordinate system the points P, Q, R have the coordinates

$$P = \frac{1}{l}(|\vec{QP}|\cos\alpha, |\vec{QP}\sin\alpha) \quad (7.1)$$

$$Q = (0,0) \quad (7.2)$$

$$R = (0, \frac{|\vec{QR}|}{l}) \quad (7.3)$$

, where α is the angle between the straight lines $\vec{QR}$ and $\vec{QP}$.



The conventional way to define a coordinate system is to assign coordinate-values to a fixed number of points.

$$P = (x_P, y_P) \quad , \quad Q = (x_Q, y_Q) \quad , \quad R = (x_R, y_R)$$

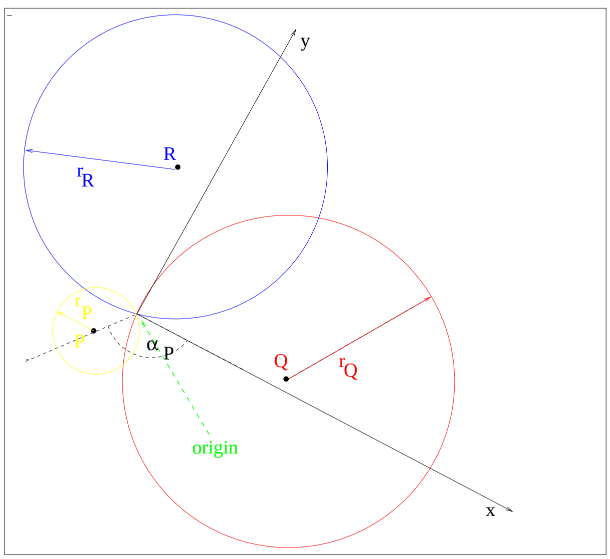
Hence the distances of the origin of the coordinate system to the three points are

$$r_P = \sqrt{x_P^2 + y_P^2} \quad , \quad r_Q = \sqrt{x_Q^2 + y_Q^2} \quad , \quad r_R = \sqrt{x_R^2 + y_R^2}$$

Consequently, the origin is at the intersection of three circles with the radii r_P, r_Q, r_R and the centers at P, Q, R.
From the assigned coordinates the angles $\alpha_P, \alpha_Q, \alpha_R$ between the lines connecting the origin with these points and the x-axis can be computed:

$$\alpha_P = \arctan(\frac{y_P}{x_P}), \alpha_Q = \arctan(\frac{y_Q}{x_Q}), \alpha_R = \arctan(\frac{y_R}{x_R})$$

The x-axis can be chosen so that it includes the angle α_P with the line connecting$P$ with the origin.



Click here to see an animation

Independent of the way of its definition a coordinate system is always a mathematical fiction. Therefore, it is impossible to have access to the coordinate system by geodetic measurements. A materialization of the coordinate system is needed. Such an realization of a coordinate system by material points with given coordinates with respect to the coordinate system under consideration is called *reference frame* .

Definition 29 *A set of material points with given coordinates with respect to a particular coordinate system is called* a reference frame *of this coordinate system.*

Unfortunately, the terminology is not clearly used in geodetic literature. Frequently, the concepts of reference systems and reference frames are not clearly distinguished. Additionally, every reference frame defines, by conventional definition also a coordinate system, which approximates the underlying coordinate system and forms together with the geodetic datum parameters a new reference system. For this reasons, the points of a reference frame are sometimes also called a reference system.

Chapter 8

Time Systems

Many geodetic observation techniques measure travel times of electromagnetic waves. Therefore a precise definition of time is fundamental to geodetic observations. Presently, two time systems are in use

- atomic time,
- dynamical time

Before atomic time was available civilian time systems were based on the Earth's rotation and were called *universal* or *sideral* time.

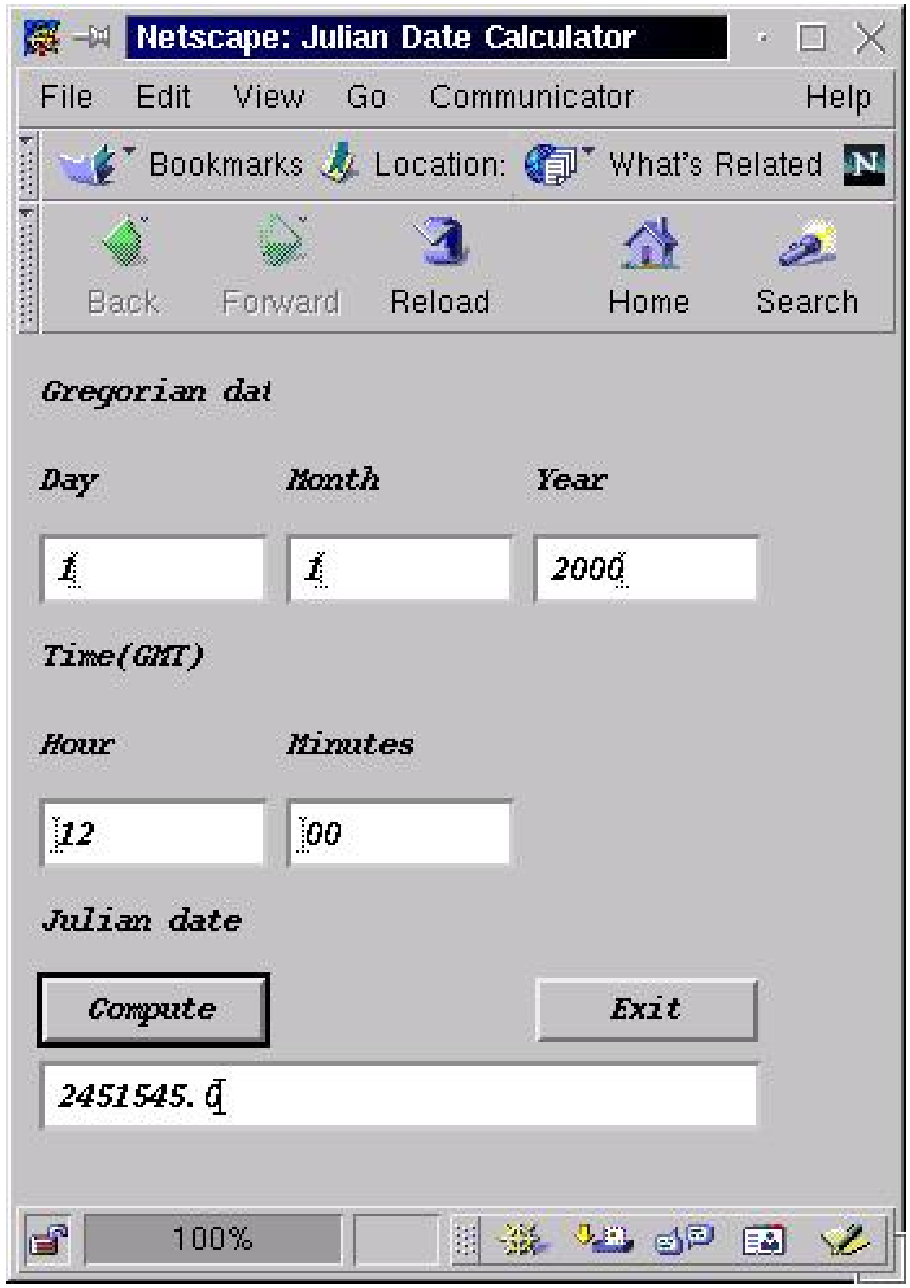
8.1 Atomic Time

The fundamental atomic time scale *Temps Atomique International - TAI* is based on atomic clocks, operated by various national agencies and kept by the *International Earth Rotation Service - IERS* and the *Bureau International de Poids et Mesures -BIPM).*

Definition 30 *TAI is a uniform time scale coinciding with Universal Time (UT) at midnight January 1, 1958.*
The fundamental interval unit of TAI is one SI second.
The SI day is defined as 86400 SI seconds and the Julian century is defined as 36525 SI days.

Definition 31 *The Julian Date is the number of days and the fraction of a day elapsed since* $12, h$ *UT on January 1, 4713 BCE.*

For conversion from Gregorian to Julian date click here



Since the origin of Julian Date is much to far in the past, UT has another fundamental epoch to refer time differences to:

Definition 32 *The Julian Date of the standard epoch of UT is called J2000.0. It is defined as the Julian Date at 1 January 2000 12:00 GMT.*

Hence,

$$J2000.0 = JD2451545.0 \tag{8.1}$$

All time variables, denoted by T are measured in Julian centuries relative to the epoch J2000.0.

TAI is a continuous time scale, it does not maintain synchronization with the solar day (UT) .Since the rotation rate of the Earth is slowing down the TAI will get more and more ahead UT. This problem is solved by the definition of the *Universal Coordinated Time* UTC.

Definition 33 *UTC runs at the same rate as TAI but is incremented periodically by leap seconds.*

Leap seconds are introduced by the IERS if necessary. The introduction of leap seconds makes sure that the difference between UTC and UT (more precisely: between UTC and UT1) is not larger than 0.9 s. The difference DUT1 := UT1-UTC is brodcasted by the IERS.

A third atomic time is the GPS time.

Definition 34 *The GPS time (GPST) runs at the rate of the atomic clock of the GPS Master Control station in Colorado Springs. GPST and UTC coincided at* o^h *January 6 1980.*

Since GPST is not incremented there is a 19 seconds offset between TAI and GPST

$$GPST = 19s + TAI \tag{8.2}$$

8.2 Dynamical Time

Definition 35 *Dynamical Time is the independent variable in the equations of motion of bodies under gravitational forces according to the theory of General Relativity.*

Since the best approximation of an inertial system is centered at the barycentre of the solar system, the dynamical time measured in this system is called Barycentric Dynamical Time (Temps Dynamique Barycentrique - TDB).
An Earth based clock will show periodic variations of about 1.6 milliseconds with respect to TDB due to the motion of the Earth in the gravitational field of the sun.

Definition 36 *Terrestrial Dynamical Time TDT (Temps Dynamique Terrestre) is the independent variable in the equation of motion of a body in the Earth's gravitational field.*

The relations between TDT and TDB are given by

$$TDB = TDT + 0^s.001658 \sin(g + 0.0167 \sin g) \tag{8.3}$$

where

$$g = (357^\circ.528 + 35999^\circ.050T)\frac{\pi}{180} \tag{8.4}$$

T is the time in Julian centuries TDT.

8.3 Sideral and Universal Time

Prior to the operationality of atomic clocks, the Earth's diurnal rotation was used to measure time. Two different time systems were connected to the rotation Earth:

- sideral time
- universal time

These two times are still used as an angle measure for the transformation between celestial and terrestial systems.

Definition 37 *The angle between the observers local meridian and and the true vernal equinox corrected for precession and nutation is called* apparent sideral time *(AST).*
If this angle is referred to the Greenwich mean astronomical meridian, it is called Greenwich apparent sideral time *(GAST)*

besides the times AST and GAST, which refer to the *true* vernal equinox there are corresponding times MST and GMST, which refer to the *mean* vernal equinox.(corrected only for precession)

Definition 38 *The difference between GAST and GMST is called the equation of Equinox EqE*

$$Eq.E := GAST - GMST = AST - MST = \Delta\psi \cos(\varepsilon + \Delta\varepsilon) \tag{8.5}$$

where the nutations in longitude and obliquity $\Delta\psi, \Delta\varepsilon$ are given by (9.8) and (9.9).

According to Kepler's second law the Earth doesn't revolve the sun at a constant angular velocity. For this reason a fictious sun was invented which moves with constant velocity.

Definition 39 *The hour angle of the fictious sun is called Universal Time UT. The Time UT1 is the Universal Time corrected for polar motion.*

The relationship between sideral and universal time is given in terms of the IAU(1967) system of constants by

$$\begin{aligned} GMST &= UT1 + 6^h 41^m 50''.548481 + 8640184''.812866 T_u \\ &\quad + 0''.093104 T_u^2 - 6''.2 \cdot 10^{-6} T_u^3 \end{aligned} \tag{8.6}$$

With T_u being the Julian date since J2000.0 in Julian centuries

$$T_u = \frac{JulianUT1date - 2451545.0}{36525} \tag{8.7}$$

Chapter 9

Geodynamics

9.1 Earth rotation

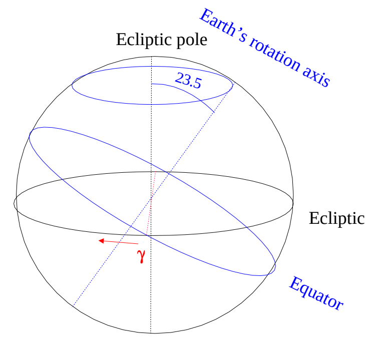
The rotation axis of the Earth is not fixed in inertial space neither it is with respect to the Earth's body. The gravitative forces of the Sun and Moon acting on the equatorial bulge of the Earth are changing the orientation of the rotation axis in inertial space. These changes are called *precession* and *nutation* and can be predicted with a very high accuracy.
Additionally, there is a small movement of the Earth's rotation axis with respect to its crust, which is called *polar motion*. Both nutation and polar motion are the Earth's response to external forces. Nutation is primarily the forced response of the Earth and can be predicted by geophysical and orbital models. The polar motion represents the forced and the free response of the Earth to external forces in almost equal parts. Again the forced part can be predicted but the free part can only be determined by Space Geodesy methods.

9.1.1 Motion in Celestial System

Moon and Sun and the planets exert gravitational forces on the equatorial bulge. Since the rotating Earth behaves like a gyro, it reacts to this forces by a clockwise movement of its rotation axis. This movement consist of two constituents

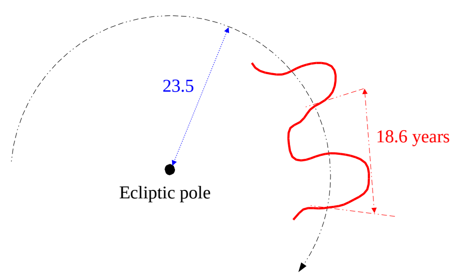
- precession and
- nutation

Theorem 17 Luni-solar precession *is the the circular motion of the celestial pole with a period of* $25,800$ *years and an amplitude equal to the obliquity of the ecliptic of* $23°.5$. *The precession causes a westerly movement of the equinox of about* $50''.3$ *per year.*
Planetary precession consist of a $0°.5$ *per year rotation of the ecliptic resulting in a easterly motion of the equinox by about* $12''.5$ *per century and an decrease of the obliquity of the ecliptic by about* $47''$ *per century.*



Definition 40 *The combined effect of luni-solar and planetary precession is called* general precession *or simply* precession

Definition 41 *The short periodic motion of the pole superimposed on the precession with oscillations of* 1 *day to* 18.6 *years (the main period) and a maximum amplitude of* $9''.2$ *is called* nutation.



Precession transformation

The transformation of stellar coordinates from the mean equator and equinox at epoch t_i to the mean equator and equinox at another epoch t_j is performed by the means of the following rotation matrix

$$\mathbf{P} = \mathbf{R}_3(-z_A)\mathbf{R_2}(\theta_A)\mathbf{R}_3(-\zeta_A) \tag{9.1}$$

The precession angles , defined by the 1976 IAU conventions, are given by

$$\begin{aligned} \zeta_A \quad = \quad & (2306''.2181 + 1''.39656T_u - 0''.000139T_u^2)t \\ & +(0''.30188 - 0''.000344T_u)t^2 + 0''.017998t^3 \end{aligned} \tag{9.2}$$

$$\begin{aligned} z_A &= (2306''.2181 + 1''.39656 T_u - 0''.000139 T_u^2)t \\ &\quad +(1''.09468 - 0''.000066 T_u)t^2 + 0''.018203 t^3 \end{aligned} \tag{9.3}$$

$$\begin{aligned} \theta_A &= (2004''.3109 - 0''.85330 T_u - 0''.000217 T_u^2)t \\ &\quad -(0''.42665 - 0''.000217 T_u)t^2 - 0''.041833 t^3 \end{aligned} \tag{9.4}$$

where

$$T_u := (JD - 2451545.0)/36525 \tag{9.5}$$

and t is the interval between t_j and t_i in Julian centuries.

Nutation transformation

The transformation of stellar coordinates from the mean to the true equator and equinox at a epoch is given by

$$\mathbf{N} = \mathbf{R}_1(-\varepsilon - \Delta\varepsilon)\mathbf{R}_3(-\Delta\psi)\mathbf{R}_1(\varepsilon) \tag{9.6}$$

The nutation time series according to the 1980 IAU conventions are

$$\begin{aligned} \varepsilon &= (84381''.448 - 46''.8150 T_u + 0''.00059 T_u^2 + 0''.001813 T_u^3) \\ &\quad +(-46''.8150 - 0''.00177 T_u + 0''.005439 T_u^2)t \\ &\quad +(-0''.00059 + 0''.005439 T_u)t^2 + 0''.00181 t^3 \end{aligned} \tag{9.7}$$

The nutation parameters $\Delta\psi$ and $\Delta\varepsilon$ can be represented by series expansions

$$\Delta\psi = \sum_{j=1}^{N} \left[(A_{0j} + A_{1j}T) \sin \left(\sum_{i=1}^{5} k_{ji}\alpha_i(T) \right)) \right] \tag{9.8}$$

$$\Delta\varepsilon = \sum_{j=1}^{N} \left[(B_{0j} + B_{1j}T) \sin \left(\sum_{i=1}^{5} k_{ji}\alpha_i(T) \right)) \right] \tag{9.9}$$

The α coefficients are arguments of the motion of Sun and Moon:

1. mean anomaly of the Moon

$$\alpha_1 = 485866''.733 + (1325^r + 715922''.633)T + 31,''.310T^2 + 0''.064T^3 \tag{9.10}$$

2. mean anomaly of the Sun

$$\alpha_2 = 1287009''.804 + (99^r + 1292581''.224)T - 0''.577T^2 - 0''.012T^3 \tag{9.11}$$

3. mean argument of latitude of the Moon

$$\alpha_3 = 335778''.877 + (1342^r + 2995263''.137)T - 13''.257T^2 + 0''.011T^3 \tag{9.12}$$

4. mean elongation of the Moon from the Sun

$$\alpha_4 = 10072261''.307 + (1236^r + 1105601''.328)T - 6''.891T^2 + 0''.019T^3 \quad (9.13)$$

5. mean longitude of the ascending lunar node

$$\alpha_5 = 450160''.280 - (5^r + 482890''.539)T + 7''.455T^2 + 0''.008T^3 \quad (9.14)$$

Here 1^r means one revolution , i.e. $1^r = 360° = 1296000''$ The coefficients A_{ij}, B_{ij}, k_{ij} are given by the standard 1980 IAU series and can be found in [?].

9.1.2 Motion in the Terrestrial System

Besides the movement of the Earth's rotation axis in space there is an additional variation of the rotation axis relative to the Earth's crust. This motion is primarily due to the elastic properties of the Earth and due to the exchange of angular momentums between the solid Earth, the oceans and the atmosphere.

Definition 42 *Polar motion is the rotation of the true celestial pole as defined by the precession and nutation models with respect do the z-axis of a conventionally chosen terrestrial reference system.*

Polar motion consists of a free and a forced oscillation. The free oscillation is counterclockwise with a period of 430 days (*Chandler period*) and an amplitude of $3 - 6m$.
The forced component again consists of two parts. One part is exited by the tidal forces and therefore has a diurnal period, with an amplitude of one order of magnitude smaller then the free oscillation. The second part has an annual period since it is exited by the annual changes in the atmosphere. Its amplitude is about as large as the free oscillation.

Polar motion cannot be predicted by models, it has to be observed by space techniques. The accuracy of those observation has achieved a very high level, accounting for 0.2 - 0.5 miliarcseconds which is equivalent to 6 - 15 mm at the Earth's surface. Polar motion values can be downloaded from the *International Earth Rotation Service* (IERS) as tables of daily values of pole coordinates.

Earth Orientation Transformation

The transformation from the celestial to the terrestrial system includes the Earth's rotation and the polar motion. Consequently it can by represented as the product of three rotation matrices:

$$\mathbf{S} = \mathbf{R}_2(-x_p)\mathbf{R}_1(-y_p)\mathbf{R}_3(GAST) \quad (9.15)$$

The Earth rotation is the rotation around the instantaneous rotation axis with the rotation angle being the difference between the true vernal equinox of the date to the meridian of the 1903.0 Greenwich zero longitude.

$$\mathbf{R}_3(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \theta = GAST \tag{9.16}$$

where GAST is given by

$$GAST = GMST_0 + \frac{d(GMST)}{dt}(UTC - (UTC - UT1) + Eq.E \tag{9.17}$$

with $Eq.E$ being again the equation of equinox. Here the difference $UTC - UT1$ has to be interpolated from IERS tables.

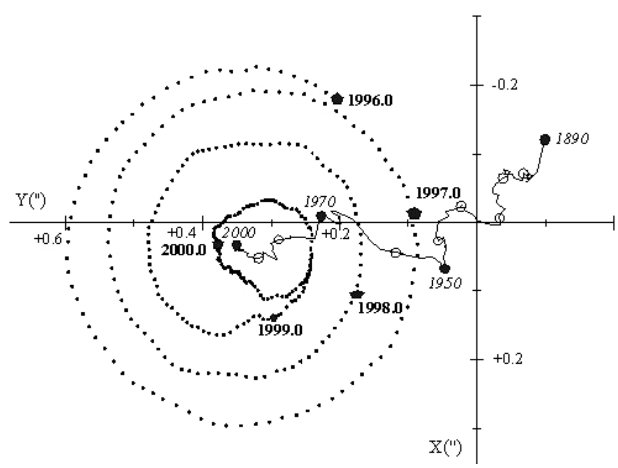
The polar motion rotation is the transformation between the instantaneous pole to the the pole given by nutation and precession theories.

Definition 43 *Polar motion rotation is defined by the left-handed pair of angles* (x_p, y_p) *. The first angle is the angle between the mean direction of the pole during the period 1900.0 - 1906.0 (the IERS Reference Pole (IRF)) and the true rotation axis. It is defined positive in the direction of the x-axis of the terrestrial system. the second angle is positive in the direction of the* $270°$ *meridian.*

Since both angles are small the rotation can be approximated by

$$\mathbf{R}_2(-x_p)\mathbf{R}_1(-y_p) = \begin{bmatrix} 1 & 0 & x_p \\ 0 & 1 & -y_p \\ -x_p & y_p & 1 \end{bmatrix} \tag{9.18}$$

where the angles are interpolated from the IERS tables.



9.2 Earth Deformation

9.2.1 Rotation versus Deformation

The location of a point at the Earth's surface in inertial space changes due to two reasons

1. the rotation of the Earth's body,

2. the deformation of the Earth.

Since, with the help of Space Geodesy only the position of a point or the change of the position of a point in inertial space can be measured, a additional criterion is needed to distinguish the deformation from rotation.
Let $\mathbf{v}_I$ be the velocity of a point at the Earth's surface in inertial space. Then it can be decomposed into

$$\mathbf{v}_i = \mathbf{v}_T + \omega \times \mathbf{r} \tag{9.19}$$

The first term is the movement of the point with respect to an Earth-fixed system, i.e. the deformation and the second term is the movement of the point due to Earth's rotation. The vector ω is the rotation vector of the Earth. From measurements only $\mathbf{v}_i$ is accessible, and this vector has to be separated somehow into deformation and rotation. This separation is ambiguous, but at least there is a restriction which all possible deformations have to fulfill.

Theorem 18 *The velocity vector* $\mathbf{v}_T$ *of the Earth's deformation has to fulfill the following condition*

$$\int \rho(\mathbf{r} \times \mathbf{v}_T)dV = 0 \tag{9.20}$$

Proof: Since the deformation rate is small compared to the rotation rate, the following condition can be used to define a mean rotation axis ω

$$T := \int \rho(\mathbf{v}_T^\top \mathbf{v}_T)dV \to_{\omega}^{min}$$

Inserting (9.19) one obtains

$$\begin{aligned} T &= \int \rho(\mathbf{v}_i - \omega \times \mathbf{r})^\top (\mathbf{v}_i - \omega \times \mathbf{r})dV \\ &= \int \rho(\mathbf{v}_i^\top \mathbf{v}_i - 2\mathbf{v}_i^\top (\omega \times \mathbf{r}) + (\omega \times \mathbf{r})^\top (\omega \times \mathbf{r})dV \end{aligned}$$

The necessary condition for an extremum is the vanishing of the gradient of T with respect to ω.

$$0 = \frac{\partial T}{\partial \omega} = -2 \int \rho(\mathbf{v}_i \times \mathbf{r})dV + 2 \int \rho(\mathbf{r} \times (\omega \times \mathbf{r}))dV$$

or equivalently

$$\int \rho(\mathbf{v}_i \times \mathbf{r})dV = \int \rho(\mathbf{r} \times (\omega \times \mathbf{r}))dV \tag{9.21}$$

On the other hand the total angular momentum is

$$\begin{aligned} H &= \int \rho(\mathbf{v}_i \times \mathbf{r})dV \\ &= \int \rho(\omega \times \mathbf{r} + \mathbf{v}_T) \times \mathbf{r}dV \\ &= \int \rho(\mathbf{r} \times (\omega \times \mathbf{r}))dV + \int \rho(\mathbf{r} \times \mathbf{v}_T)dV \end{aligned}$$

Comparing this with(9.21) yields

$$\int \rho(\mathbf{r} \times \mathbf{v}_T)dV = 0$$

9.2.2 Global Plate Motion

The NNR-NUVEL1 plate tectonic model by [**?**] describes the angular velocities of the 14 major tectonic plates defined by the constraint (9.21).
The velocity a station i in a plate j is given on a spherical Earth as a function of φ, λ, R by

$$\mathbf{v}_{ij} = \Omega_j \times \mathbf{r}_i = R\omega_j \begin{bmatrix} \cos\varphi_j \sin\varphi_i \sin\lambda_j - \sin\varphi_j \cos\varphi_j \sin\lambda_i \\ \sin\varphi_j \cos\varphi_i \cos\lambda_i - \cos\varphi_j \sin\varphi_i \cos\lambda_j \\ \cos\varphi_j \cos\varphi_i \sin(\lambda_i - \lambda_j) \end{bmatrix} \tag{9.22}$$

Consequently, the station coordinate corrections for global plate motion are given by

$$\mathbf{r}_{ij}(t) = \mathbf{r}_{ij}(t_0) + \mathbf{v}_{ij} \cdot (t - t_0) \tag{9.23}$$

9.2.3 Tidal Effects

The gravitational attractions of Moon and Sun cause tidal deformations of the Earth which result in periodical changes of station coordinates. Therefore, a tidal model needs tio be included in the definition of a terrestrial reference system.
Earth tides have four main constituents:

1. solid Earth tides,
2. ocean loading,
3. atmospheric loading
4. pole tide

Solid Earth Tides

Let P be a massive body with mass M_P at the distance R_P from the mass center of the Earth. This body generates a tidal potential with the value

$$U_{tidal}(\mathbf{r}_S) = \frac{GM_P}{R_P}\left([\frac{r_S}{R_P}]^2 P_2(\cos\vartheta) + [\frac{r_s}{R_P}]^3 P_3(\cos\vartheta)\right) = U_2 + U_3 \qquad (9.24)$$

at the location $\mathbf{r}_S$ at the Earth's surface. The angle ϑ is the angle between the position vectors of $\mathbf{r}_S$ and the tidal force generating body $\mathbf{P}$.
Thee resulting displacements expressed in a topocentric system are

$$\delta = \sum_{i=2}^{3} h_i \frac{U_i}{g}\mathbf{e}_V + l_i \frac{\cos\varphi_S \frac{\partial U_i}{\partial \lambda_S}}{g}\mathbf{e}_E + l_i \frac{\frac{\partial U_i}{\partial \varphi_S}}{g}\mathbf{e}_N \qquad (9.25)$$

In this equation

- the vectors $\mathbf{e}_E, \mathbf{e}_N, \mathbf{e}_V$ are unit vectors pointing in East, North and vertical direction,
- the real numbers h_i, l_i are the vertical and horizontal Love Numbers and
- g is the gravity acceleration.

The following values have been recommended by the IERS for tidal corrections:

- $h_2 = 0.609 \quad , \quad l_2 = 0.0852 \quad , \quad h_3 = 0.292 \quad , \quad l_3 = 0.0151$
- $GM_E = 3986004.356 \cdot 10^8 m^3 s^{-2}$ (Earth)
- $GM_S = 1.32712440 \cdot 10^20 m^3 s^{-2}$ (Sun)
- $M_E/M_M = 81.300585$ (Earth/Moon mass ratio)

Ocean Loading

Ocean loading is the elastic response of the Earth to ocean tides. This effect can reach tens of millimeters for stations near the ocean shelves. Corrections for ocean tides displacements have the form

$$\delta_j = \sum_{i=1}^{N} \xi_i^j \cos(\omega_i t + V_i - \delta_i^j) \qquad (9.26)$$

The IERS standards include $N = 11$ tidal constituents. For each constituent i

- ω_i is its frequency,
- V_i is the astronomical argument and
- ξ_i^j, δ_i^j are the amplitudes and phase lags.

Atmospheric Loading

Atmospheric loading is the elastic response of the Earth's crust to varying atmospheric pressure distribution. This effect can reach several millimeters in vertical direction.

Pole Tide

The pole tide is the elastic response of the Earth's crust to the to the shift of the rotation axis. The maximum displacements of the pole tide are 10-20 mm.

Chapter 10

Conventional Reference Systems and Reference Frames

10.1 International Earth Rotation Service (IERS)

Recent reference systems are maintained through international cooperation. The International Association of Geodesy (IAG) has set up a service, the International Earth Rotation Service (IERS), which is concerned with the maintainance of the IERS Reference System. In different IERS Analysis Centers for geodetic space techniques as VLBI, SLR LLR, GPS the parameters for the Earth rotation are computed. In the Central Bureau these informations are combined and in regular updates the IERS Reference System is released.
The IERS Reference System is composed of

- IERS Standards, a set of models and parameters, which are used by the Analysis Centers.
- the IERS Celestial Reference Frame (ICRF) and
- the IERS Terrestrial Reference Frame (ITRF)

The ICRF is realized by a catalogue of compact extragalactical radio sources. The ITRF is realized by a set of terrestrial station coordinates and velocities.

10.2 Celestial Reference System

Definition 44 *The small motions of the Earth's rotation axis can be described as the sum of two components*

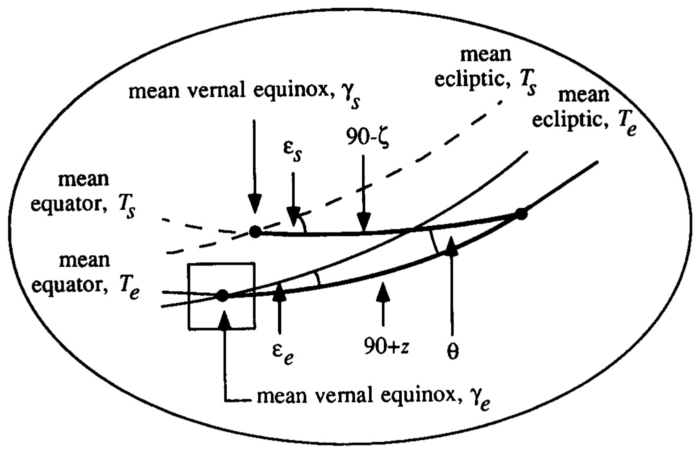
1. *astronomical nutation*

2. *polar motion*

The direction of the axis which is computed from the theory of nutation and precession is called Celestial Ephemeris Pole (CEP)
The origin of the ICRS is the barycentre of the solar system.
The axes of the ICRS are defined as the

- *the CEP,*
- *the equinox*
- *and a third axis completing the former two axis to a Cartesian coordinate system*

at the epoch J2000.0



The ICRF is a realization of the IERS consisting of catalogue of astronomical coordinates of about 200 extragalactical radio sources at the epoch J2000.0.
By adopting coordinates of quasars, implicitly a coordinate system is conventionally defined. The conventionally defined coordinate system differs in the orientation of its axis by about 0.0001 arc-seconds from the ICRS.
The transformation from the ICRF to a system with its third axis to CEP is given by the theory of nutation and precession.

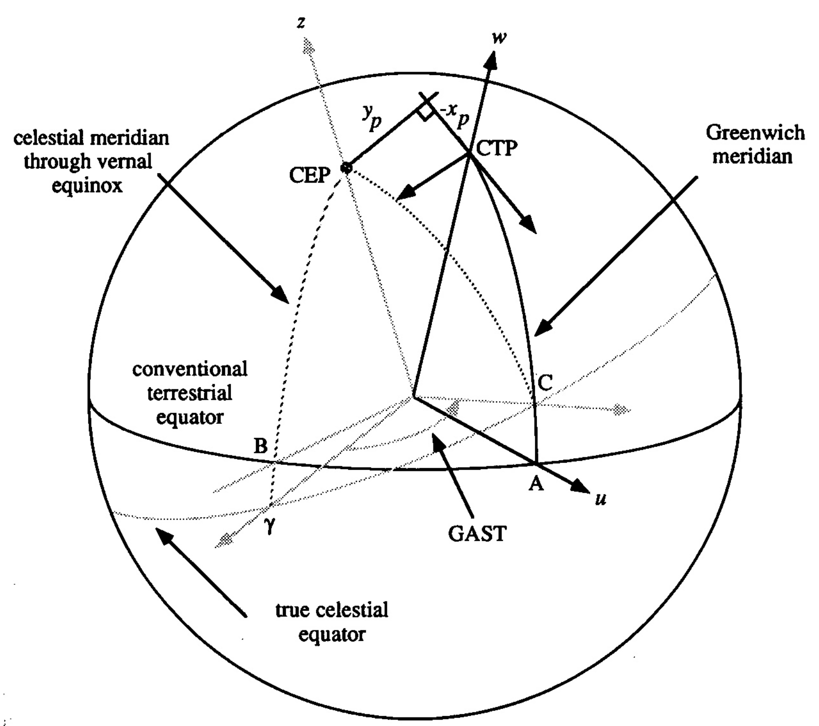
10.3 Terrestrial Reference Systems

The CEP moves with respect to the Earth's surface. In order to have a coordinate system , which is fixed with respect to the Earth the ITRS is adopted.

Definition 45 *The mean direction of the the Earth rotation axis determined by the five International Latitude Service stations in the period 1900.0 to 1906.0 is defined as the Conventional Inertial Pole (CIO) at the epoch 1903.0*

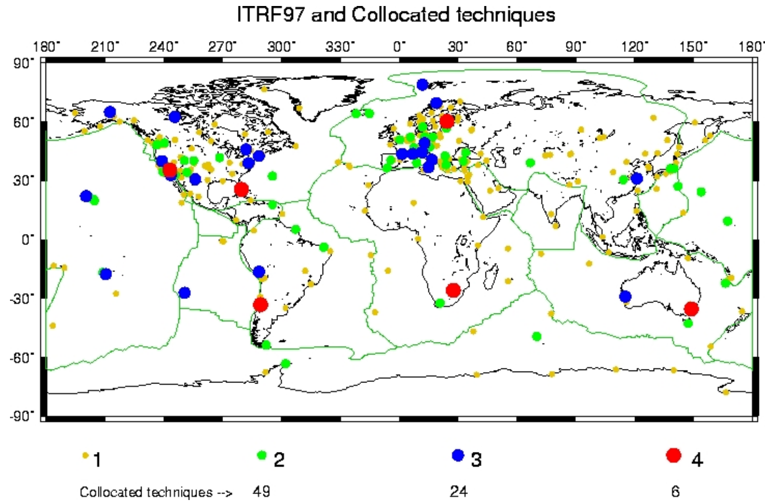
Definition 46 *The ITRS is defined with its origin at the Earth's geocenter. The axes of the ITRS are oriented in the following way:*

- *the Z axis is oriented to the CIO,*
- *the X-axis is oriented towards the 1903.0 meridian of Greenwich*
- *and the Y axis completes the former two axes to a Cartesian coordinate system.*



The ITRS is realized by the ITRF, a catalogue of Cartesian coordinates and velocities of globally distributed tracking stations.
The adopted coordinates of these stations implicitly define a new coordinate system, which differs from the ITRS by about 10mm in position and several mm/year in velocity.

The transformation between the ITRF and the ICRF is given by the pole coordinates x_P, y_P and the nutation and precession parameters $d\psi, d\varepsilon$.



10.4 WGS84

Besides the ITRF several other terrestrial reference systems are in use. Probably, the most important is the WGS84. The WGS84 is maintained by the US Department of Defense (DoD) and is the reference system of the GPS system. It is an implicitly defined system. It is defined by adopting Cartesian coordinates of the ten DoD GPS Monitoring Stations derived from Doppler observations on these sites. This results in a accuracy of the WGS84 System of about 1..2 m. In order to align the WGS84 with the more accurate ITRF the DoD has recomputed the coordinates of the ten monitoring stations using GPS observations at these sites and at at ab subset of IGS tracking stations whose ITRF coordinates were held fixed. This refined WGS84 System is called WGS84 (G730).
The WGS84 system is realized by the ephemerides of the GPS satellites. In order to compute the orbits of these satellites some additional constants have to be adopted.
An ellipsoidal coordinate system is attached to the WGS84 by locating an ellipsoid at the origin of the WGS84 system and letting the rotation axis coincide with the Z-axis of the WGS84. This means the datum parameters of the WGS84 are

Parameter	Symbol	numerical value
semi-major axis	a	6378137 m
reciprocal flattening	1/f	298.257223563
angular velocity	ω	$7.292115 \cdot 10^{-5} s^{-1}$
geocentric		
gravitational constant	GM	$398600.5 km^3 s^{-2}$
second zonal harmonic	$\bar{C}_{2,0}$	$-4884.16685 \cdot 10^{-6}$

10.5 Ellipsoidal Reference Systems

For many purposes ellipsoidal coordinate systems are more convenient than Cartesian systems. Ellipsoidal systems can be distinguished between global systems which approximate the Earth as a whole and local ellipsoidal system, which approximate the Earth's surface in a certain region.

10.5.1 The GRS80 Reference System

The most important global ellipsoidal System is the GRS System. It is defined as an ellipsoid centered at the origin of the ITRS and having its axes coinciding with the axes of the ITRS. The additional datum parameters of the GRS80 Reference System are

Parameter	Symbol	numerical value
semi-major axis	a	6378137 m
reciprocal flattening	1/f	298.257222100827
angular velocity	ω	$7.292115 \cdot 10^{-5} s^{-1}$
geocentric		
gravitational constant	GM	$398600.5 km^3 s^{-2}$
dynamical form factor	J_2	$108263 \cdot 10^{-8}$

The GRS80 Reference system is a global ellipsoidal system. It approximates the Earth as a whole. Besides global ellipsoidal systems a number of local ellipsoidal systems are in use. The approximate the Earth's surface only in their region of validity.

10.5.2 Local Ellipsoidal Systems

The Rauenberg Datum

The Rauenberg datum is the official reference system for the western part of Germany. It is an ellipsoidal System which is based on the Bessel 1841 ellipsoid. This ellipsoid has the dimensions

$$a = 6377397.155m, \quad 1/f = 299.15281285 \tag{10.1}$$

The position and orientation datum parameters are not given as the position of the origin and the orientation of the axes but in an equivalent way. First a initial point in the center of the region of validity of the reference system has to be fixed. For this initial point the following quantities are assigned

- the ellipsoidal coordinates L, B are set identical to the astronomic coordinates λ, φ,
- the geodetic azimuth A to a specific target is set identical to the astronomical azimuth a

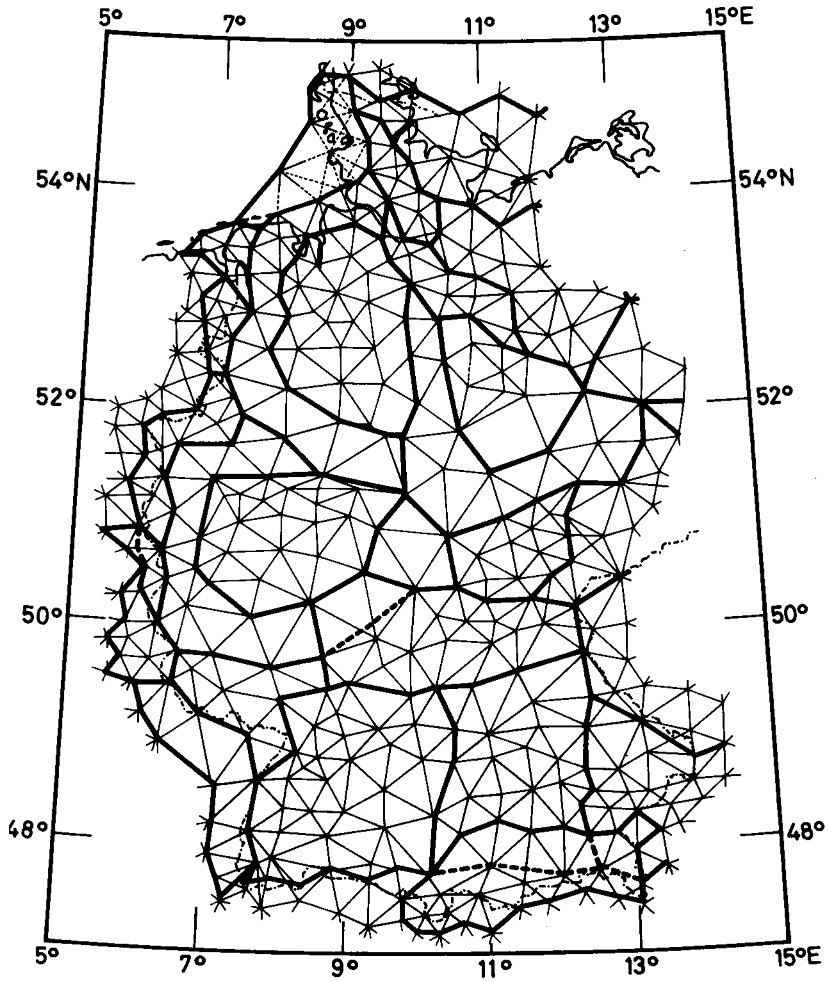
- a specific value N is adopted for the separation between the geoid and the ellipsoid in the initial point.

For the Rauenberg datum the Helmert Tower of The GeoForschungsZentrum in Postsdam was used as initial point. Its ellipsoidal coordinates were set to

$$B = 52°22'53''.9540N, \quad L = 11°34'26''.483E \tag{10.2}$$

The coordinate system was oriented by setting the geodetic azimuth of the line *Rauenberg - Marienkirche Berlin* to its astronomical azimuth.

The reference frame of the Rauenberg datum are the points of the German *Hauptdreiecksnetz* (DHDN) whose Gauß - Krüger coordinates based on a 3° zone-width are given.



The 42/83 Datum

The system 42/83 used to be the official reference system in the eastern part of Germany. It is based on the Krassovsky 1940 ellipsoid with the following dimensions

$$a = 6378245m, \quad 1/f = 298.3 \tag{10.3}$$

Its initial point is the Centra Astronomical Observatory Pulkovo (close to St. Petersburg) and its reference frame are the points of

- the *Einheitliches Astronomisch-Geodätisches Netz*(EAGN) and
- the *Staatliches Trigonometrisches Netz 1st Order* (STN1.O)

whose Gauß - Krüger coordinates based on a 6° zone-width are given.

10.6 Height Systems

So far, the horizontal position of a point is given by its ellipsoidal coordinates L, B. The vertical position, the ellipsoidal height H is given by the shortest distance of the point from the surface of the ellipsoid. Despite of this conceptual simple concept the ellipsoidal heights do not have property, which is intuitionally expected of heights:

there is no water floating between points of identical heights

In order to fulfill this requirement points of identical heights have to lie on an equipotential surface $W(\mathbf{x}) = c_0 = const$ of the gravity potential W of the Earth. Hence, a useful definition of a physical height system is to let the heights be proportional to the negative difference of the potential difference between the ocean-surface and the point

$$h_P \sim -(W_P - W_0) \tag{10.4}$$

The choice of the proportionality factor distinguished the different height systems.

10.6.1 Dynamical Heights

Definition 47 *The* dynamical height h_P^{dyn} *of a point P is given by*

$$h_P^{dyn}(P) := W_0 - W(P) \tag{10.5}$$

, where W_0 is the value of the gravity potential at a tide - gauge.

This height system has the disadvantage that the dynamical heights don't have a metric unit but the unit $m^2 s^{-2}$. For practical purposes it is more convenient to have heights given in metrical units. Therefore, a quantity with the unit $m^{-1}s^2$. Dependent of this choice orthometric or normal heights are generated.

10.6.2 Orthometric Heights

Heights Reference System

Definition 48 *The* orthometric height h_P^{orth} *of a point P is given by*

$$h_P^{orth}(P) := \frac{1}{\bar{g}}(W_0 - W(P)) \tag{10.6}$$

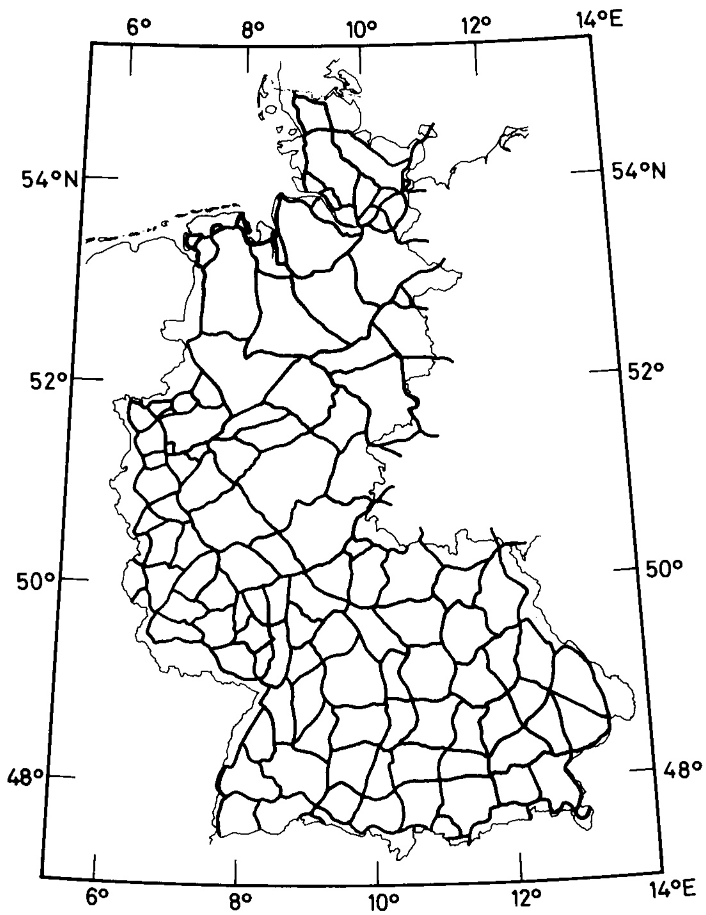
, where W_0 is the value of the gravity potential at a tide - gauge and $\bar{g}$ is the mean value of the gravity along the plumb line between the surface $W(\mathbf{x}) = W_0$ and the point P.

The orthometric heights have a nice geometric interpretation: They equal the length of the plump-line between the surface $W(\mathbf{x}) = W_0$ and the point P.

The disadvantage of the orthometric heights is that the value $\bar{g}$ cannot be measured but has to be computed from gravity measurements at the Earth surface including some hypothesizes about the density distribution inside the Earth's body. Therefore in some countries, for instance in the western part of Germany, mean value $\bar{g}$ of the real gravity is replaced by the mean value of a gravity model , the so-called normal gravity γ. The resulting heights are called *normal orthometric heights*.

Heights Reference Frame

The reference frame of the orthometric height system in the western part of Germany are the heights of the points of the *Deutsches Haupthöhennetz* (DHHN). The heights refer to the equipotential surface $W = W_0$, which passes a point 37m beneath the Berlin Astronomical Observatory.(Normalhöhenpunkt von 1879 (NH1879)). The height of the NH1879 was connected by spirit-leveling to the Amsterdam tide gauge.
Due to the demolition of the Berlin Astronomical Observatory the NH1879 was replaced by a NH1912 in Hoppegarten 40 km eastward of Berlin. Hence, the normal orthometric heights are more or less the vertical distance from an equipotential surface passing the tide gauge in Amsterdam.



10.6.3 Normal Heights

Normal Heights Reference System

Definition 49 *The* normal height h_P^n *of a point* P *is given by*

$$h_P^n(P) := \frac{1}{\gamma(P)}(W_0 - W(P)) \tag{10.7}$$

, where W_0 *is the value of the gravity potential at a tide - gauge and* $\gamma(P)$ *is the mean value of the normal gravity at the latitude of the point* P*.*

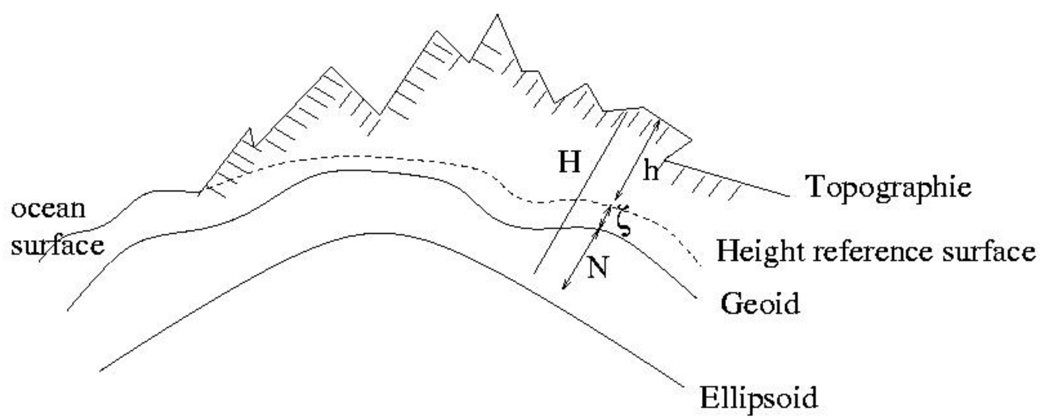
The normal heights do not have an obvious geometric interpretation but the can be derived without any geophysical hypothesis.

Normal Heights Reference Frame

The heights of the points of the *Staatliches Nivellementsnetz 1st Ordnung* (STNN1O) form the reference frame of the normal height Reference System in the eastern part of Germany. They refer to an equipotential surface passing the tide gauge in Kronstadt (close to St. Petersburg).
Due to oceanographic effects the tide gauges in Amsterdam and Kronstadt do not belong to the same equipotential surface of the gravity potential W of the Earth. This has the consequence that the orthometric and the normal height of the same point differ by about 15 cm.

10.6.4 Conversion between geometrical and physical heights

Since due to GPS a direct access to geometric heights is possible and since physical heighst can be measured by a combination of spirit leveling and gravity measurements, the question of a conversion of the two types of height systems arises. The height systems are in a conceptally simple relationship



$$H = h + N + \zeta \tag{10.8}$$

with

- H the ellipsoidal height,
- h the physical height,
- N the so called geoid undulation,
- ζ the deviation of tide gauge from the geid, the so called sea-surface topography

The central concept, connecting geometrical with physical heights is the concept of the geoid.

Definition 50 *The equipotential surface*

$$W(\mathbf{x}) = W_0 = const \tag{10.9}$$

, which coincides with the undisturbed surface of the oceans is called the geoid.

If the Earth were a regular body with homogeneous mass distribution being in a hydrostatic equilibrium, its shape would be an ellipsoid an its gravity potential W would coincide with the normal potential U. Due to the deviation from this model assumptions the gravity potential W and the normal potential differ from each other. Let be U_0 the value of the normal potential at the surface of the ellipsoid, than the geoid is the equipotential surface $W = U_0$, and the separation between ellipsoid and geoid is called geoid undulation N. It can be determined from gravity measurements along the surface of the Earth. The determinationof the geoid is one central topic of Physical geoid.
Assuming that the tide gauge would be exactly located at the surface of the geoid , the following simple relation between physical and geometrical heights were true

$$H = h + N \tag{10.10}$$

In practice it is impossible, to locate a tide gauge at the geoid, it is alway located at the mean surface of the ocean insteadt. The mean surface of the oceans differs by some dm up to some meters from the geoid. this small difference is the cause of stationary ocean circulations like the gulf stream. It it is called sea-surface topography and can be determined by oceanographic measurements and satellite altimetry.
Due to the sea-surface topography the height reference surface is not the geoid but another equipotential surface passing thrugh the tide gauge. Both surfaces differ in height by the amount ζ.