LMV-rapport 2010:1

Rapportserie: Geodesi och Geografiska informationssystem

On geodetic transformations

Bo-Gunnar Reit

Gävle 2009

LANTMÄTERIET

Author Bo-Gunnar Reit

Typography and layout Rainer Hertel

Total number of pages 57

LMV-Report 2010:1 – ISSN 280-5731

On geodetic transformations

Bo-Gunnar Reit

Gävle

December 2009

Preface

The purpose of the report is to give an exhaustive description of the methods developed by the author during the years 1995 – 2004 in the process of establishing transformations between RT 90, the municipal systems and SWEREF 93/99. It is assumed that the reader is aware of the fundamental concepts and the geodetic systems that are used in Sweden.

The author greatly thanks Jonas Ågren and other colleagues at the geodetic development division for contributing with valuable comments on the work. Special thanks to Lars E Engberg for transferring the text to the document standard of Lantmäteriet* .

December 2009

Bo-Gunnar Reit
bo-gunnar.reit@telia.com

Lantmäteriet thanks Maria Andersson for her work with the English translation.

* Lantmäteriet – the Swedish mapping, cadastral and land registration authority

Table of Contents

1 Problem description

With the break-through of GPS technology, the need arose to transform coordinates between SWEREF 99 (initially SWEREF 93) and RT 90 as well as various local systems.

2 Involved systems

2.1 SWEREF 99

SWEREF 99 differs from the other systems by being a true 3-dimensional system with global connection. The positions of the reference points are determined in a Cartesian coordinate system (X, Y, Z), the origin of which nearly coincides with the centre of gravity of the earth. The reference ellipsoid GRS 80 is tied to the system. The centre of the ellipsoid coincides with the origin of the Cartesian system. The relation between the Cartesian coordinates of a point (X, Y, Z) and the geodetic coordinates of the point, latitude, longitude and height above the ellipsoid, (φ, λ, h), can be written by the formula (see also figure 1)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} (N+h)\cos\varphi\cos\lambda \\ (N+h)\cos\varphi\sin\lambda \\ (N(1-e^2)+h)\sin\varphi \end{bmatrix} \qquad (2\text{-}1)$$

where $N = a/\sqrt{1-e^2\sin^2\varphi}$ and a is the semi major axis of the ellipsoid, e^2 is the first eccentricity squared and N is the radius of curvature of the prime vertical.

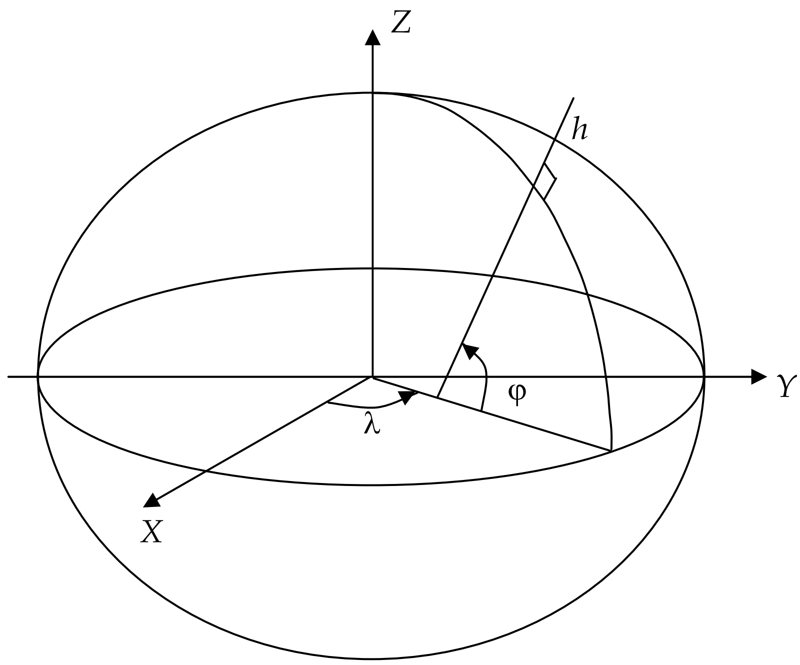

***Figure 1**: Geocentric Cartesian system and geodetic system.*

As shown in figure 1, the X- Y- and Z coordinates refer to a system with its origin in the centre of the ellipsoid. *Let us, somewhat improperly, call this type of coordinates geocentric coordinates.*

The transformation from (X, Y, Z) to (φ, λ, h) is here left out but can be performed completely without loss of accuracy, see for example Bowring (1976).

2.2 RT 90

RT 90 is a 2-dimensional system where positions are given as latitude and longitude, (φ, λ), relative to the reference ellipsoid Bessel 1841. For most triangulation points in the national network, there are heights above the sea level in the RH 70 system, yet their quality is varying. Worse still, the geoid corrections required for transforming the RH 70 heights to heights over the Bessel ellipsoid are of even lower quality, with errors at the level of 1-2 m. This contaminates the geocentric coordinates (X, Y, Z) which can only be obtained from (φ, λ, h) through transformation using equation (2-1).

2.3 RR 92

The National Reference System of 1992. An "untrue" three-dimensional system based on Bessel's ellipsoid. It is simply a joining together of the horizontal system RT 90, the geoid height system RN 92 and the height system RH 70.

The origin that is the centre point of the reference ellipsoid, of RR 92 was placed about a kilometre from the Earth's centre of gravity. It was placed there to obtain a good national fit to the geoid. Globally however, this placement as well as the dimensions of the ellipsoid is of poor accordance.

RN 92

The geoid heights in RN 92 refer to Bessel's ellipsoid, oriented in such a way as to make the geoid heights roughly coincide with those of the older Swedish geoid height system RAK 70. RN 92 was thereby intended to be of use for three-dimensional computations, for example in GPS surveying, as well as for height reduction of conventionally surveyed distances.

2.4 Municipal systems

The municipal systems are 2-dimensional Cartesian grid systems (x, y). The manner in which the systems are defined varies among the municipalities. Most systems are connected to the older national system RT 38 or one of the so-called regional systems. Because of the inferior geometrical quality of RT 38, the system has in some cases been tied to just one point in combination with orienting the system with the help of some additional point, in order to avoid the defects of RT 38 being propagated in the local system. Municipal systems, defined completely separately from the national systems here mentioned, also occur. By applying projection corrections for Gauss-Krüger's projection, in accordance with current regulations (VF/TFA), the municipal systems have received geometrical properties corresponding to this projection. In most cases, there is no method, given à priori, to transform the grid coordinates (x, y) to geodetic coordinates, and, consequently, neither to geocentric coordinates.

2.5 Conventional systems

From now on, the term conventional systems is used for all systems that, similar to RT 90 and the municipal systems, have been created with the help of conventional distance- and angle measurements.

3 Transformation methodology

Transformation methodology refers to the methods that are applied when two or more horizontal systems are used in the same geographic area and one wishes to transfer the coordinates of points within the area from one system to another.

The most common method for transforming coordinates between globally connected systems and national reference frames of an older type, in our case between SWEREF 99 and RT 90, is to use a similarity transformation in three dimensions (3D Helmert transformation). It is assumed that one has access to coordinates of good quality in both systems for a number of points, henceforth called common points. The points should preferably be evenly distributed within the area in which the set of transformation parameters is to be used. The procedure is to first perform a fit based on the common points, where the seven parameters that compose the transformation are estimated: three translations, three rotations and one scale correction. The estimated parameters are then used to transform the remaining points in the area. Even though three common points are sufficient to determine the parameters, the number of points should be no less than 10 and may well be more, depending on the circumstances. In order to avoid ambiguousness, only one set of parameters should be determined for each area.

A detailed run through of how the 3D Helmert transformation has been implemented is given in the following section.

The question of transformation between SWEREF 99 and the municipal systems is more complex and this will be treated in section 8 Projection fit.

4 Similarity transformation in 3 dimensions

By the name *Similarity transformation in 3 dimensions* – or *3D Helmert transformation* – it is made clear that this transformation preserves the shape of objects. In vector form the mathematic relation can be written as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_B = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + (1+\delta)\mathbf{R}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_A \tag{4-1}$$

where the vectors indexed A and B, respectively, symbolize coordinates of the two systems that the transformation is performed between, where $(\Delta X, \Delta Y, \Delta Z)^T$ constitutes the translation vector between the origins of the systems, δ the scale correction and where the rotation matrix $\mathbf{R}$ is defined as

$$\mathbf{R} = \mathbf{R}_Z\mathbf{R}_Y\mathbf{R}_X = \begin{pmatrix} \cos\Omega_Z & \sin\Omega_Z & 0 \\ -\sin\Omega_Z & \cos\Omega_Z & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \cos\Omega_Y & 0 & -\sin\Omega_Y \\ 0 & 1 & 0 \\ \sin\Omega_Y & 0 & \cos\Omega_Y \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Omega_X & \sin\Omega_X \\ 0 & -\sin\Omega_X & \cos\Omega_X \end{pmatrix} \tag{4-2}$$

and Ω_X, Ω_Y and Ω_Z is the rotation around each axis.

(4-1) can be written compactly as

$$\mathbf{X}_B = \Delta\mathbf{X} + (1+\delta)\mathbf{R}\mathbf{X}_A \tag{4-3}$$

A coordinate transformation can be interpreted in two ways: either one studies the changes and movements of an object within a single coordinate system or one studies the same object seen in two separate coordinate systems. In geodesy, we deal with the latter issue, that is, the two reference systems involved are seen as two different models that describe reality.

The three-dimensional Helmert transformation became more generally spread within geodesy with the use of satellite technology for positioning. Initially, neither the measurement techniques nor the systems were very accurate, which is why linearised versions of formulas (4-1) and (4-2) were widely spread, not least because of the publication by the Defence Mapping Agency in the USA of a report, DMA TECHNICAL REPORT tr8350.2-a, which contained a linearised formula. The problem with the simplified formulas is that they do not fulfil the current consistency requirements of the computations. It is also awkward to produce the strict inverse of the linearised versions. The strict inverse transformation of formula (4-1) can very easily be computed thanks to the fact that the inverse of matrix $\mathbf{R}$ is identical to its transpose ($\mathbf{R}^{-1}=\mathbf{R}^T$). There is also no reason in terms of efficiency to linearise formula (4-1) since the nine elements of matrix $\mathbf{R}$ are only computed once, which means that the transformation of a certain number of points takes the same time either the complete formula or a linearised version is used.

It should be noted that even though the transformation is in three dimensions, in practice, it is only the horizontal component that one is interested in transforming. The height component can be regarded as an unfortunate necessity. As we shall see later, it causes quite a lot of trouble both at the computation of the parameters and when transforming points.

4.1 Procedure for determining ΔX, ΔY, ΔZ, ΩX, ΩY, ΩZ and δ

To be able to compute numerical values for the parameters, access to common points, whose coordinates are known in both systems is required. As previously pointed out, the common points should be evenly distributed over the area within which one wishes to use the parameters. By inserting the known coordinates for systems *A* and *B* into formulas (4-1) and (4-2), each point yields three equations, one for each of the coordinates *X*, *Y* and *Z*, contributing to determine the constants ΔX, ΔY, ΔZ, Ω_X, Ω_Y, Ω_Z and δ. Since three or more common points are used, the system of equations will be over-determined, which makes the method of least squares suitable for solving it. Since the equations are not linear with respect to the rotations and scale, some manual work is required to solve for the unknown parameters. More is said about this in section 5 3D Helmert fit between two topocentric systems.

In the immediate following section, the procedure for computation of the older set of transformation parameters between WGS 84 and RT 90 and between SWEREF 93 and RT 90 (RR 92), published by the National Land Survey, is presented.

4.2 The WGS 84 – RT 90 transformation parameters

With the break-through of GPS technology, a need to transform coordinates determined by GPS to RT 90 immediately arose. In 1989, the National Land Survey, Hedling & Reit(1989), produced a set of transformation parameters. The WGS 84 coordinates were based on two Scandinavian Doppler campaigns (SCANDOC) which contained seven Swedish points. The fit was performed using the module Helmer in the so-called Bernese software. The accuracy was quite modest, with a residual of 2.4 metres per coordinate. However, this was of no great consequence since the transformation parameters were meant for applications in cartography and navigation.

4.3 The SWEREF 93 – RT 90 transformation parameters

In 1993, a surveying campaign was carried out where 22 Swedish stations were surveyed, most of which are among the present SWEPOS stations, with massive support from the Onsala Space Observatory. The solution was computed by Jan Johansson at Onsala in ITRF 91, epoch 1992.5, and was fitted into EUREF 89 using 11 points in Northern Europe with known EUREF 89 coordinates. The coordinates obtained thereof defines SWEREF 93. The internal accuracy for the horizontal component (1σ, 2D) was estimated to 2 cm. The corresponding accuracy in RT 90 was, according to experience, 1-2 cm between adjacent points. Because of the strength of the network, where the observations were adjusted along with the triangulation nets of the neighbouring countries, no great deformations of RT 90 over the country as a whole was expected, apart from a possible difference in scale. The residuals should presumably be around 5-10 cm expressed as rms (2D).

On this occasion too, the computation of the parameters was performed in the software module Helmer, in which the Helmert formula is implemented entirely in conformity with equations (4-1) and (4-2). An analysis showed that rms of the horizontal residuals was over 13 cm, with a maximum error of 35 cm in Kiruna, which was considerably worse than expected. A graphic presentation showed apparent systematic tendencies for the residual vectors. Further investigations yielded that part of the error originated in flaws in the geodetic definition of RT 90.

The foundation of RT 90 was an adjustment on Hayford's ellipsoid of all distance- and angle measurements, performed with original coordinates in ED 87. The reason for this was that a reliable geoid model for distance reduction was accessible only for Hayford's ellipsoid from 1910. When RT 90 was introduced, there was a demand from cartographers that the RT 90 coordinates should differ as little as possible from the corresponding RT 38 coordinates. With regard to the tight schedule for the construction of the digital map, a change from Bessel 1841 to Hayford 1910 was not possible. Therefore, the creation of RT 90 somewhat meant pulling oneself up by the hair. However, compared to many other countries, the size of the residuals for the Swedish transformation was on a rather modest level. By the end of 1994 the new set of transformation parameters were made public.

5 3D Helmert fit between two topocentric systems

5.1 Problem definition

The issue of the poor fit remained. Most likely, the discrepancy was caused by flaws in the geoid model and that the difference in the radius of curvature between the Bessel and Hayford ellipsoids in some way affected the definition of RT 90. Both phenomenons were clearly height related. Software was needed, in which to perform the fit, where the height constraint could be eliminated. The module Helmer could not manage this and neither could any other software, which gave cause for the development of an own programme. In this software, the RT 90 heights were introduced as unknown entities. This was not difficult to accomplish, which can be realized when studying equations (2-1) and (4-1). The results of this experiment proved very successful and the rms value dropped from 13 to 5 cm.

Even though the results were satisfactory, the formulation was not sufficiently universally applicable. A more general approach would be to reformulate the problem so that in the fit, the heights are weighted according to their expected accuracy. Another disadvantage with both Helmer and this software was that the rotation parameters referred to the geocentric coordinate axes and the translations to the shift between the centres of the ellipsoids. If the fit is performed between two global systems, with common points distributed over several continents, this type of parameters is well suited to describe the relation between the systems, but in remaining cases the area covered by the fit is a relatively small part of the earth surface. When looking at a globe, one realizes this holds true also for continental systems such as ED 50 and NAD 83.

For systems covering a minor part of the earth surface, introducing a topocentric system for each ellipsoid, see figure 2, and performing the fit between these systems, has distinct advantages and, which will be shown later, facilitates the understanding of the procedure of the fit. Provided the coordinates for the origins of the topocentric systems are chosen correctly, the rotation around the z-axis will correspond to the azimuthal rotation between the systems while the rotations around the other two axes describes the local tilt between the ellipsoid surfaces in the area. Finally, the translation (Δz) along the topocentric z-axis gives the distance between the ellipsoid surfaces around the topocentre. As will be seen later on, the translation (Δz) and the scale correction δ are closely connected.

Among the advantages is also an improved numeric precision in the computations, compare to centre of gravity reduction at 2D Helmert fit.

5.2 Relation between topocentric systems

The following is a presentation of how the relation between the topocentric systems can be derived and how the parameters obtained from a fit can be converted to the corresponding geocentric parameters.

From now on, we will use capital letters for geocentric coordinates and lower-case letters for topocentric. The origins of the topocentric systems are placed at the point (φ_0, λ_0, 0) on the surface of each ellipsoid, with the z-axis coinciding with the outward-directed ellipsoid normal, the x-axis in the meridian plane and the y-axis oriented so as to make up a left-oriented system. The topocentric xy-plane is consequently a tangential plane of the ellipsoid, see figure 2.

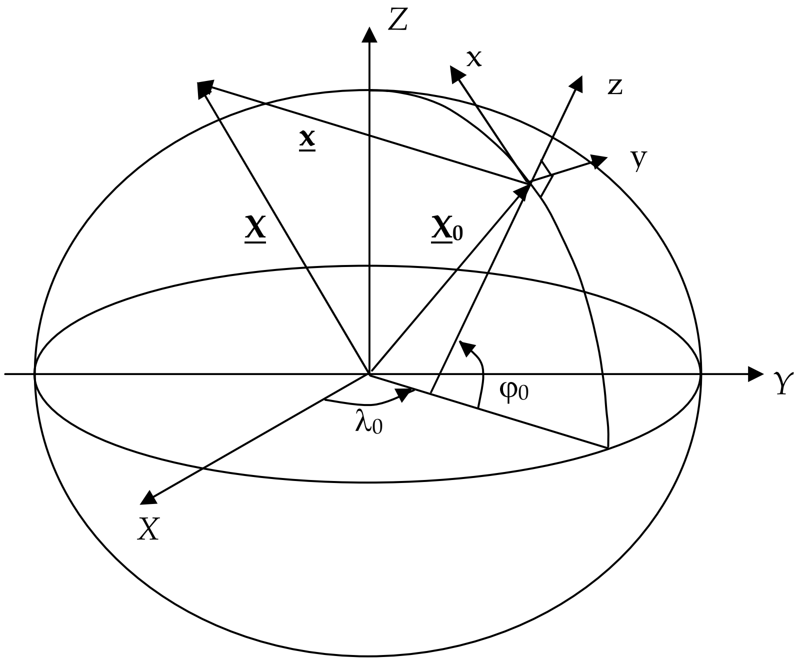

Figure 2: The topocentric and geocentric coordinate systems

From the figure we can derive the following relation between the vectors $\underline{\mathbf{X}}$, $\underline{\mathbf{X}}_0$ and $\underline{\mathbf{x}}$ and between the unity vectors in the geocentric and the topocentric system.

$$\underline{\mathbf{X}} = \underline{\mathbf{X}}_0 + \underline{\mathbf{x}} \tag{5-1}$$

$$\begin{aligned}
\underline{\boldsymbol{e}}_x &= -\sin\varphi_0 \cos\lambda_0 \underline{\boldsymbol{e}}_X - \sin\varphi_0 \sin\lambda_0 \underline{\boldsymbol{e}}_Y + \cos\varphi_0 \underline{\boldsymbol{e}}_Z \\
\underline{\boldsymbol{e}}_y &= -\sin\lambda_0 \underline{\boldsymbol{e}}_X + \cos\lambda_0 \underline{\boldsymbol{e}}_Y \\
\underline{\boldsymbol{e}}_z &= \cos\varphi_0 \cos\lambda_0 \underline{\boldsymbol{e}}_X + \cos\varphi_0 \sin\lambda_0 \underline{\boldsymbol{e}}_Y + \sin\varphi_0 \underline{\boldsymbol{e}}_Z
\end{aligned} \tag{5-2}$$

With equations (5-1), (5-2) we can now write

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \begin{pmatrix} -\sin\varphi_0 \cos\lambda_0 & -\sin\lambda_0 & \cos\varphi_0 \cos\lambda_0 \\ -\sin\varphi_0 \sin\lambda_0 & \cos\lambda_0 & \cos\varphi_0 \sin\lambda_0 \\ \cos\varphi_0 & 0 & \sin\varphi_0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{5-3}$$

We introduce the term $\mathbf{M}_0$ for the matrix in equation (5-3) that transfers the topocentric vector components to the corresponding geocentric ones. As can be seen, the equations (4-1) and (5-3) are similar. However, observe that in equation (5-3), the two systems involved are of different orientation. Similar to matrix $\mathbf{R}$, the inverse of $\mathbf{M}_0$ is equal to its transpose ($\mathbf{M}_0^{-1}=\mathbf{M}_0^{T}$).

For conversion of geocentric coordinates to topocentric the formula is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{M}_0^{-1} \left[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} \right] \tag{5-4}$$

Before the geocentric coordinates can be converted to topocentric, φ_0 and λ_0 must be assigned appropriate values. Theoretically, one can chose arbitrary values for each ellipsoid (nor does h need be equal to 0), but for the conversion to topocentric systems to be effective *both systems should be assigned the same numerical values (φ_0, λ_0, 0), whereby the axes of the topocentric system are given the same orientation relative to the geocentric system in both system A and system B,* and the point chosen should be in the centre of the area concerned in the fit, for example the average of the horizontal coordinates of the common points in the system deemed to have the most reliable coordinates.

We thus have

$$\mathbf{X}_A = \mathbf{X}_{0A} + \mathbf{M}_0 \mathbf{x}_A \quad \text{and} \quad \mathbf{x}_A = \mathbf{M}_0^{-1}(\mathbf{X}_A - \mathbf{X}_{0A}), \tag{5-5}$$

$$\mathbf{X}_B = \mathbf{X}_{0B} + \mathbf{M}_0 \mathbf{x}_B \quad \text{and} \quad \mathbf{x}_B = \mathbf{M}_0^{-1}(\mathbf{X}_B - \mathbf{X}_{0B}) \tag{5-6}$$

respectively.

Note that, even though the same numerical values (φ_0, λ_0, 0) are used in system A and system B, the vector $\mathbf{X}_{0A}$ will differ from the vector $\mathbf{X}_{0B}$, since different ellipsoids were used in the computation.

After the conversion to topocentric systems, the fit is performed. Indicate the parameters for the translations and rotations between the topocentric systems by Δx, Δy, Δz and ω_x, ω_y, ω_z, respectively. The scale correction δ is the same irrespective of whether the fit is performed between the geocentric systems or the topocentric.

Analogous to the geocentric case, the similarity transformation between the topocentric systems can be written as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_B = \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} + (1+\delta)\,\mathbf{R}_{Topo} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_A \quad \text{or, alternatively} \quad \mathbf{x}_B = \Delta\mathbf{x} + (1+\delta)\mathbf{R}_{Topo}\mathbf{x}_A \tag{5-7}$$

where

$$\mathbf{R}_{Topo} = \mathbf{R}_z(\omega_z)\mathbf{R}_y(\omega_y)\mathbf{R}_x(\omega_x) =$$

$$= \begin{pmatrix} \cos\omega_z & \sin\omega_z & 0 \\ -\sin\omega_z & \cos\omega_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\omega_y & 0 & -\sin\omega_y \\ 0 & 1 & 0 \\ \sin\omega_y & 0 & \cos\omega_y \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_x & \sin\omega_x \\ 0 & -\sin\omega_x & \cos\omega_x \end{pmatrix} \tag{5-8}$$

As previously pointed out, the choice of identical values (φ_0, λ_0, 0) in the definition of the origin for system A and system B makes the interpretation of the rotation parameters easier. For example, the rotation around the topocentric z-axis is equivalent to an azimuthal rotation between the systems. As is shown in equation (5-7), Δz corresponds to the separation of the ellipsoid surfaces in the area of the topocentric origins.

By inserting the known coordinates in equation (5-7), three equations for each common point are obtained. If the number of points is ≥ 3, we have more constraints than unknown parameters and the system of equations is solved by the method of least squares, which means that for each equation a so called residual is added and the chosen solution is the one that minimizes the sum of the squares of the residuals. As mentioned in the beginning, the accuracy of the heights of the common points is normally lower than that of the planar components. This is particularly true for the conventional systems, partly because the heights of the common points are often not obtained by levelling and partly because of defects in the geoid model. To avoid the poor fit in height spilling over to the fit in the horizontal components it is necessary to down-weight the height fit. For small areas it is sufficient to down-weight the equation for the z component of system *B*. Because of the curvature of the earth the orientation of the axes, north, east and up, for points far from the topocentre will differ from the axes of the topocentric system. Thus, the weighting will not be entirely correct. As a first step to solving this problem we bring back equation (5-7) to the geocentric axial orientations of system *B* by multiplying with the matrix $\mathbf{M}_0$, which is obtained by inserting the latitude- and longitude values of the topocentre. We also reverse the left-hand and right-hand sides of the equation. For the i^{th} common point we get

$$\mathbf{M}_0(\Delta\mathbf{x} + (1+\delta)\mathbf{R}_{Topo}\mathbf{x}_{iA}) = \mathbf{M}_0\mathbf{x}_{iB} \tag{5-9}$$

We now want to transfer equation (5-9) to the axial orientations (north, east, up) of the i^{th} point. We achieve this by multiplying equation (5-9) with the matrix obtained by inserting the latitude- and longitude values of the i^{th} point for system *B* into the expression for the inverse of matrix **M**. Finally, by adding the residual vector $\mathbf{v}_i$, the three observation equations for the i^{th} point can be written as

$$\mathbf{M}_{iB}^{-1}\mathbf{M}_0(\Delta\mathbf{x} + (1+\delta)\mathbf{R}_{Topo}\mathbf{x}_{iA}) = \mathbf{M}_{iB}^{-1}\mathbf{M}_0\mathbf{x}_{iB} + \mathbf{v}_i \tag{5-10}$$

Now the equations can be assigned weights corresponding to the quality of the entered coordinates. We can affirm that with our way of formulating the observation equations there is no correlation between the observations, with the possible exception of that caused by the manner in which the coordinates were once established.

In the case of conventional systems, we can assume that errors in the planar components and the height component are uncorrelated. Since the planar coordinates most likely have been obtained in a process of adjustment, there probably exists correlation between the errors of the coordinates of different points. However, it is less likely that these correlations are accessible. The same is true for the two planar components of each point. Because of the added geoid correction, it is also likely that the errors in height show a certain correlation between different points. To try to take that into consideration is difficult and hardly of value for the problem of the fit.

In the case of the SWEREF systems and similar global systems, the error level is so low that the coordinates in this context can be regarded as without error. Therefore, in the fit, it feels suitable to always choose these systems as the system one transforms from (system A), thus the residuals will be added to the coordinates of the conventional systems.

The conclusion of the above discussion is that one can, without major limitations, regard the variance-covariance matrix as diagonal, which means that when forming the observation equations one only needs to divide each equation with its à priori standard deviation.

5.3 Linearization

The next problem to handle is the fact that equation (5-10) is not linear with respect to the unknown entities Δx, Δy, Δz, ω_x, ω_y, ω_z and δ. This is customarily solved by linearization combined with iteration. The method is favourable for this problem since the sought rotations and scale correction normally are small entities, but works excellently also with arbitrarily large rotations and scale changes.

The procedure can shortly be described in the following way. The expression within parentheses in equation (5-10) is denominated by $\mathbf{F}(\Delta x,\Delta y,\Delta z,\omega_x,\omega_y,\omega_z,\delta)$. We then get

$$\mathbf{F}(\Delta x,\Delta y,\Delta z,\omega_x,\omega_y,\omega_z,\delta)=\Delta\mathbf{x}+(1+\delta)\mathbf{R}_{Topo}\mathbf{x}_{iA} \tag{5-11}$$

We now perform a Taylor series expansion around the approximate values

$(\Delta x)_0$, $(\Delta y)_0$, $(\Delta z)_0$, $(\omega_x)_0$, $(\omega_y)_0$, $(\omega_z)_0$ and $(\delta)_0$

$$\begin{aligned}\mathbf{F}=\mathrm{F}_0&+(\frac{\partial\mathbf{F}}{\partial\Delta x})_0\,d\Delta x+(\frac{\partial\mathbf{F}}{\partial\Delta y})_0\,d\Delta y+(\frac{\partial\mathbf{F}}{\partial\Delta z})_0\,d\Delta z+\\&+(\frac{\partial\mathbf{F}}{\partial\omega_x})_0\,d\omega_x+(\frac{\partial\mathbf{F}}{\partial\omega_y})_0\,d\omega_y+(\frac{\partial\mathbf{F}}{\partial\omega_z})_0\,d\omega_z+(\frac{\partial\mathbf{F}}{\partial\delta})_0\,d\delta\end{aligned} \tag{5-12}$$

where the corrections $d\Delta x$, $d\Delta y$, $d\Delta z$, $d\omega_x$,, $d\omega_y$,, $d\omega_z$ and $d\delta$ are the unknowns and where the approximate value $\mathbf{F}_0$ is defined as

$$\mathbf{F}_0=\mathbf{F}((\Delta x)_0,(\Delta y)_0,(\Delta z)_0,(\omega_x)_0,(\omega_y)_0,(\omega_z)_0,(\delta)_0)$$

For the partial derivatives we get

$$\frac{\partial\mathbf{F}}{\partial\Delta x}=\begin{pmatrix}0\\1\\0\end{pmatrix};\quad\frac{\partial\mathbf{F}}{\partial\Delta y}=\begin{pmatrix}0\\1\\0\end{pmatrix};\quad\frac{\partial\mathbf{F}}{\partial\Delta z}=\begin{pmatrix}0\\0\\1\end{pmatrix};\quad\frac{\partial\mathbf{F}}{\partial\delta}=\mathbf{R}_{Topo}\mathbf{x}_A\;;$$

$$\frac{\partial \mathbf{F}}{\partial \omega_{\mathbf{x}}} = (1+\delta)\mathbf{R}(\omega_z)\mathbf{R}(\omega_y)\frac{\partial \mathbf{R}(\omega_x)}{\partial \omega_x}\mathbf{x}_A \text{ where } \frac{\partial \mathbf{R}(\omega_x)}{\partial \omega_x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin\omega_x & \cos\omega_x \\ 0 & -\cos\omega_x & -\sin\omega_x \end{pmatrix}$$

$$\frac{\partial \mathbf{F}}{\partial \omega_y} = (1+\delta)\mathbf{R}(\omega_z)\frac{\partial \mathbf{R}(\omega_y)}{\partial \omega_y}\mathbf{R}(\omega_x)\mathbf{x}_A \text{ where } \frac{\partial \mathbf{R}(\omega_y)}{\partial \omega_y} = \begin{pmatrix} -\sin\omega_y & 0 & -\cos\omega_y \\ 0 & 0 & 0 \\ \cos\omega_y & 0 & -\sin\omega_y \end{pmatrix}$$

$$\frac{\partial \mathbf{F}}{\partial \omega_z} = (1+\delta)\frac{\partial \mathbf{R}(\omega_z)}{\partial \omega_z}\mathbf{R}(\omega_y)\mathbf{R}(\omega_x)\mathbf{x}_A \text{ where } \frac{\partial \mathbf{R}(\omega_z)}{\partial \omega_z} = \begin{pmatrix} -\sin\omega_z & \cos\omega_z & 0 \\ -\cos\omega_z & -\sin\omega_z & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Inserting the expression on the right-hand side of the equals sign in equation (5-11) into equation (5-10) gives us the final observation equations. After estimating the unknown corrections with the method of least squares, these values are added to the approximate values and the entire procedure is repeated as long as the corrections make a significant contribution to the estimated parameters. It can be shown that having all approximate values set to zero in the first iteration step works excellently.

Comment: The formulation of the observation equations were based on the topocentric systems. If formulating the equations for geocentric parameters is preferred, one shall instead multiply equation (4-1) with the inverse (transpose) of matrix $\mathbf{M}_{iB}$. Then the three observation equations for the point become

$$\mathbf{M}_{iB}^{-1}(\Delta\mathbf{X} + (1+\delta)\mathbf{R}\mathbf{X}_{iA}) = \mathbf{M}_{iB}^{-1}\mathbf{X}_{iB} + \mathbf{v}_i \qquad (5\text{-}13)$$

The disadvantage of this approach is that one deprives oneself of the possibility to analyze the result of the fit that is offered by the topocentric approach, see detailed studies in a subsequent section (7 Detailed Study of 3D Helmert fit with real data).

6 Computation of transformation parameters

The parameters obtained by the procedure previously described can be used together with equation (5-7) to transform topocentric coordinates, which is not particularly useful. Instead we shall derive a method for computation of parameters for the transformation between the geocentric systems from the topocentric parameters.

6.1 Computation of geocentric transformation parameters from topocentric

We begin by replacing $\mathbf{x}_A$ and $\mathbf{x}_B$ in equation (5-7) with the expressions from (5-5) and (5-6). For anyone who doubts that the scale correction must be the same in

(4-1) and (5-7) we denote the topocentric scale δ_{Topo}. After reducing the expression in accordance with (5-7) we get

$$\mathbf{X}_B = \mathbf{X}_{0B} + \mathbf{M}_0 \Delta\mathbf{x} - (1+\delta_{Topo})\mathbf{M}_0 \mathbf{R}_{Topo} \mathbf{M}_0^{-1} \mathbf{X}_{0A} + (1+\delta_{Topo})\mathbf{M}_0 \mathbf{R}_{Topo} \mathbf{M}_0^{-1} \mathbf{X}_A \tag{6-1}$$

Similar to (4-1), equation (6-1) shall hold true for all points. By comparing equation (6-1) and (4-1) it is evident that

$$\delta = \delta_{Topo} \tag{6-2}$$

$$\Delta\mathbf{X} = \mathbf{X}_{0B} + \mathbf{M}_0 \Delta\mathbf{x} - (1+\delta_{Topo})\mathbf{M}_0 \mathbf{R}_{Topo} \mathbf{M}_0^{-1} \mathbf{X}_{0A} \tag{6-3}$$

$$\mathbf{R} = \mathbf{M}_0 \mathbf{R}_{Topo} \mathbf{M}_0^{-1} \tag{6-4}$$

The translation vector $\Delta\mathbf{X}$ can, with the help of equation (6-4), be further simplified to

$$\Delta\mathbf{X} = \mathbf{X}_{0B} + \mathbf{M}_0 \Delta\mathbf{x} - (1+\delta)\mathbf{R}\mathbf{X}_{0A} \tag{6-5}$$

Comment: One can easily be misled into thinking that if one transforms the coordinates for $\mathbf{X}_{0A}$ one gets the coordinates for $\mathbf{X}_{0B}$, but this is not the case since the two topocentres refer to different points in the three-dimensional space.

It is appropriate to first solve for the rotations Ω_X , Ω_Y and Ω_Z, since they are needed to compute the translations according to equation (6-5).

We begin by multiplying the algebraic expressions for the three matrices $\mathbf{R}_X$, $\mathbf{R}_Y$ and $\mathbf{R}_Z$ on the left-hand side of equation (6-4), compare to equation (4-1). We similarly compute numeric values for the nine matrix elements on the right-hand side. We then get

$$LHS =$$

$$\begin{pmatrix} \cos\Omega_Y \cos\Omega_Z & \cos\Omega_X \sin\Omega_Z + \sin\Omega_X \sin\Omega_Y \cos\Omega_Z & -\cos\Omega_X \sin\Omega_Y \cos\Omega_Z + \sin\Omega_X \sin\Omega_Z \\ -\cos\Omega_Y \sin\Omega_Z & \cos\Omega_X \cos\Omega_Z - \sin\Omega_X \sin\Omega_Y \sin\Omega_Z & \sin\Omega_X \cos\Omega_Z + \cos\Omega_X \sin\Omega_Y \sin\Omega_Z \\ \sin\Omega_Y & -\sin\Omega_X \cos\Omega_Y & \cos\Omega_X \cos\Omega_Y \end{pmatrix} \tag{6-6}$$

$$RHS = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \tag{6-7}$$

A comparison of the left-hand and the right-hand sides gives $\Omega_Y = \arcsin R_{31}$. With Ω_Y known, Ω_X can be computed from R_{32} and Ω_Z from R_{21}. The fact that arc sinus is ambiguous is a complication. Normally, the rotation angles are very small, which means that the values within the range $-\pi/2$ to $\pi/2$ can be chosen. If a solution valid for arbitrarily large rotations is required, it gets considerably more difficult since there are eight combinations to choose from, of which not all recreate the values on the right-hand side. For certain angles, e g if $\Omega_Y \approx \pm\pi/2$, one also needs to choose a different approach than the one suggested here.

Finally, we compute the numeric values for the translation vector with equation (6-5).

6.2 Computation of inverse parameters

There is sometimes need to transform coordinates in the opposite direction of that intended by the computed parameters. Three alternative ways of solving this problem can be considered.

1. *Use the same parameters but take the inverse of equation. (4-1). That is, use the formula*

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{\mathrm{A}} = \frac{\mathbf{R}^{-1}}{(1+\delta)}\left[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{\mathrm{B}} - \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}\right] \tag{6-8}$$

As previously mentioned, the inverse of the rotation matrix is $\mathbf{R}^{-1} = \mathbf{R}^{T} = (\mathbf{R}_Z\,\mathbf{R}_Y\,\mathbf{R}_X)^T = (\mathbf{R}_X)^T\,(\mathbf{R}_Y)^T\,(\mathbf{R}_Z)^T$ which gives

$$\mathbf{R}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Omega_X & -\sin\Omega_X \\ 0 & \sin\Omega_X & \cos\Omega_X \end{pmatrix}\begin{pmatrix} \cos\Omega_Y & 0 & \sin\Omega_Y \\ 0 & 1 & 0 \\ -\sin\Omega_Y & 0 & \cos\Omega_Y \end{pmatrix}\begin{pmatrix} \cos\Omega_Z & -\sin\Omega_Z & 0 \\ \sin\Omega_Z & \cos\Omega_Z & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6-9}$$

2. *Use equation(4-1) but compute the inverse parameters by performing the fit in the opposite direction, that is reverse A and B in equation (4-1).*
3. *Use equation (4-1) but compute the inverse parameters from the matemathic/numeric inverse.*

In the third alternative, the same method is applied as in the computation of geocentric parameters from the topocentric, with the only difference that in the computation of the inverse rotation parameters the matrix with the numerically computed elements is first transposed, se equation (6-7). The inverse translation vector is obtained from the original translations with formula

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}_{\mathrm{invers}} = -\frac{\mathbf{R}_{invers}}{(1+\delta)}\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} \tag{6-10}$$

where δ is the original scale correction. Finally, the inverse scale correction is obtained from

$$\delta_{invers} = -\frac{\delta}{(1+\delta)} \tag{6-11}$$

With the exception of GTRANS, there are probably exceedingly few software that can handle the inverse transformation according to the first alternative. Generally, one is therefore obliged to resort to one of the other two alternatives. The problem with the second alternative is that, because of the two least squares-solutions not being completely symmetric, the parameters obtained from the inverse fit will not agree with the strict inverse. Thus, the conclusion is that one should choose the third alternative if consistency in the computation is desired. This leads to the next question at issue.

6.3 Computational consistency for 3D Helmert

Coordinates based on geodetic measurements are always marred by errors. The size of the errors varies and can have many different causes. Safe to say is that one always aims at the best possible accuracy with regard to costs and other prerequisites. Geodesists, as suppliers of the fundamental geodetic infrastructure, must strive to offer methods and products that satisfy the accuracy needs of *all* customers. There is no reason to let errors caused by simplified formulas or other flaws in the numerical handling be added to the error budget.

Commonly in geodesy, coordinates are stored and presented with as many numbers needed to achieve consistency within the millimetre. This means latitude and longitude should be stored with at least five decimals in the arc second part or equivalent. In certain studies, for example based on SWEPOS data, there is reason to present tenths of millimetres, which requires six decimals in the arc second part. If a transformation of coordinates is performed in several steps, where intermediate results are stored as an edited print in a text file, these coordinates should be given at least one extra decimal.

It is crucial that the methods offered by geodesists to users, and the computer software used when handling coordinates, have a computational consistency that corresponds to the highest expected accuracy. With the power of computers, it is long since there were any computational motives to waive this.

The 3D Helmert transformation is one of the most abused procedures in geodesy. The algorithm is the object of a number of different simplifications. Moreover, there is some arbitrariness on how to define the order in which to perform the rotations as well as which direction of rotation should be considered positive. By applying different conventions and approximations in different contexts the numeric results run the risk of becoming inconsistent.

Usually when performing a 3D Helmert transformation, one uses previously computed and published parameters. In case the software one uses to transform the coordinates does not apply the same conventions as the software that was used to estimate the parameters, there is a risk the results will not be entirely correct. We shall investigate some of the most common hazards.

Do the parameters describe transformation from system A to system B or the opposite? It is not too uncommon that one is mistaken on this point. If the results seem wrong but appears OK if the sign is changed for all parameters, this can be an indication, but do not stop here – investigate the details and figure out the true cause. Simply using parameters with changed signs can cause errors of varying size in the coordinates. Cheating like that when using the official transformation parameters for SWEREF 99 ↔ RR 92 will cause an error of roughly 1 cm, but if the rotations are larger the error grows rapidly (quadratically).

Another obscurity that occurs is whether the rotations are considered positive clockwise or anti-clockwise: both options exist.

As mentioned in a previous section, a common simplification of the computational algorithm is neglecting higher-order terms in the rotation matrix. One then gets

$$\mathbf{R} = \begin{pmatrix} 1 & \Omega_Z & -\Omega_Y \\ -\Omega_Z & 1 & \Omega_X \\ \Omega_Y & -\Omega_X & 1 \end{pmatrix} \tag{6-12}$$

Sometimes further "simplification" has been done by placing the scale factor 1+δ on the diagonal of the matrix. The benefit of these modifications is hard to realize. As pointed out earlier, no time is gained since the nine elements of the matrix are only computed once regardless of the number of transformed points. Formula (6-12) looks deceivingly simple, but anybody interested can try to derive the strict inverse of this matrix. The error caused by linearization grows with the square of the size of the rotations. For the official set of transformation parameters SWEREF 99 ↔ RR 92, the error is 2-3 mm. As shall be seen when we get to transformation parameters between SWEREF 99 and the municipal systems, some systems has an azimuthal rotation amounting to several degrees. For a rotation of 100 arc seconds (approximately 30 mgon), the error is 0,3 m and for 1 gon the size is in the order of 300 m.

The next problem concerns the definition of rotation matrix **R.** So far, we have assumed that the multiplication order of the three partial matrices is $\mathbf{R}_Z\mathbf{R}_Y\mathbf{R}_X$. Theoretically, there are six different ways of multiplying the matrices. At least the following two variations have occurred in different international contexts, $\mathbf{R}_X\mathbf{R}_Y\mathbf{R}_Z$ and $\mathbf{R}_Y\mathbf{R}_X\mathbf{R}_Z$. The effect of using the wrong order of rotation is small. As long as the rotations remain small (<10 arc seconds) it amounts to some millimetres, but here as well the error grows quadratically with the size of the rotations. For rotations at the level of 1 degree it is a question of hundreds of metres.

The complete lack of generally accepted conventions for 3D Helmert leads us to the following code of conduct:

Never give out parameters without attaching a number of points with coordinates in the system from which transformation is to be performed as well as coordinates obtained by transformation using the parameters. Similarly, one should obviously demand a number of points for control of consistency between the parameters and software from those who give out the parameters. The valid geographic area for the parameters should *also be stated, in principle the area that is represented by the common points. For example, parameters can be determined for a project with very limited geographic extension such as the construction of a road, but from the stated system names one can easily get the impression that they are valid generally between the systems if not otherwise stated.*

Finally, an issue that one is not too uncommonly affected by, that the points to be transformed lack heights or that the available heights are heights above the sea level. That this can affect the accuracy of the horizontal position is due to the fact that the two involved ellipsoid surfaces are most often not completely parallel in the area in question. The size of the error is linearly dependent on the tilt between the ellipsoid surfaces in the area and on how erroneous the height is. An example of this is that if one transforms the 20 SWEPOS stations used in the detailed study in the next section with all heights set to zero, the maximum error is 11 mm. At a guess, a neglected geoid height gives an error of <1 mm.

7 Detailed study of 3D Helmert fit with real data

In this section we shall see how the modified approach for computation of 3D Helmert parameters works in practice.

From what can be concluded of articles in journals and other literature, in Sweden as well as internationally, a deeper understanding of how the computation of transformation parameters for 3D Helmert works in a geodetic context seems to be surprisingly rare. A simple example of how the different parameters influence the result of the fit is illustrated below.

Step by step the effect of the different parameters on the result of the fit is presented. As foundation for the study, the computation of parameters for transformation between SWEREF 99 and RR 92 is chosen, based on the 20 fundamental points in the SWEPOS network. The example shows the gradual improvement of the residuals at the introduction of each new parameter. As a last step, the consequence of removing the constraint in height is studied.

In order to make everything as concrete as possible, the following simple mechanical model is used. We regard the GRS 80- and Bessel ellipsoids as two completely separate models, both claiming to, as best as possible, describe reality. For each common point we imagine that on the surface of each ellipsoid an antenna is mounted, whose position coincides with the geodetic coordinates (φ, λ) and whose antenna height corresponds to the point's height above the ellipsoid. For GRS 80 we use the SWEREF 99 coordinates and for Bessel the RR 92 coordinates.

Performing a fit means that we try to place the ellipsoids relative to each other in such a way as to minimize the sum of the squares of the distances between the antennae tips for each common point, obviously with regard to the limitation of movement implied by the choice of transformation parameters. For example, if we do not estimate any rotations in the fit we must at all times keep the axes of the ellipsoids parallel.

7.1 0-parameter fit

As a first step we imagine a fit where all seven parameters are set to zero, that is the ellipsoids are placed concentrically with coinciding directions of axes and with no difference in scale. As a transformation interpreted this case means that we simply take the SWEREF 99 coordinates and regard them as RR 92 coordinates. The North, East and Up components (in short N, E, U) in table 1 represent the vector running from the tip of the antenna on Bessel's ellipsoid to the corresponding antenna tip on the GRS 80 ellipsoid. The axial directions for (N, E, U) is defined by the RR 92 coordinates on Bessel's ellipsoid.

According to theory of errors, (N, E, U) is a vector of observational corrections, compare to equation (5-10), but in the fit application there is ambiguity as to what should be improved. In our case one can for example ask oneself whether it is the transformed SWEREF 99 coordinates that should be improved or whether it is the RR 92 coordinates that needs correcting. In many cases one calls the vector components residuals without changing their sign.

As seen in table 1, the size of the Up component seem reasonable since we know that the equatorial radius of the Bessel ellipsoid is approximately 740 m less than that of GRS 80.

***Table 1**: Residuals for 0 parameter fit (concentric ellipsoids) (unit: metre).*

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-200.963	-172.885	707.434	265.095
KIRU.0	-215.782	-191.545	702.039	288.534
OVER.0	-193.129	-210.784	705.912	285.883
SKEL.0	-178.458	-201.779	710.361	269.373
VILH.0	-182.570	-165.654	712.106	246.522
BORA.0	-97.265	-159.096	723.630	186.473
JONK.0	-96.893	-168.657	723.573	194.509
SUND.0	-150.196	-183.189	717.170	236.890
HASS.0	-75.938	-171.236	724.671	187.319
NORR.0	-105.719	-183.842	722.483	212.072
ONSA.0	-93.692	-152.141	723.700	178.676
VANE.0	-110.142	-148.822	722.518	185.147
KARL.0	-118.826	-158.470	722.096	198.072
LEKS.0	-134.007	-165.428	720.212	212.896
LOVO.0	-113.338	-194.270	721.344	224.914
MART.6	-129.985	-185.447	719.883	226.465
OSKA.0	-86.534	-186.813	723.872	205.882
OSTE.0	-168.486	-155.984	714.992	229.605
SVEG.0	-150.578	-159.589	717.914	219.413
UMEA.0	-164.617	-193.711	714.041	254.209
R.m.s.	**144.249**	**176.313**	**717.524**	**227.803**

7.2 1-parameter fit

We further realize from table 1 that if we move the Bessel ellipsoid along the normal defined by the barycentre of the common points, so that it approaches the surface of the GRS 80 ellipsoid, the Up component will decrease and, thus, the quadratic sum of the residuals. We achieve this by performing a fit where the topocentric parameter Δz is set free. The results of this can be seen in table 2.

As expected, the Up component is radically improved, but an improvement is also made in North. Let us also take a look at the transformation parameters, which in this case become:

Topocentric parameters:

Translation x:	0.0000000000	(fixed)
Translation y:	0.0000000000	(fixed)
Translation z:	-30.2854252412	
Rotation x:	0.0000000000	(fixed)
Rotation y:	0.0000000000	(fixed)
Rotation z:	0.0000000000	(fixed)
Scale correction:	0.00000000	(fixed)

Geocentric parameters:

Translation X:	-379.4375609537
Translation Y:	-109.3308257612
Translation Z:	-603.5274797678
Rotation X:	0.0000000000
Rotation Y:	0.0000000000
Rotation Z:	0.0000000000
Scale correction:	0.00000000

Units: metre, arc second and ppm

Table 2: Residuals for 1 parameter fit with free translation in topocentric dz (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-81.970	-158.725	-3.776	178.641
KIRU.0	-78.627	-157.191	-5.201	175.759
OVER.0	-76.379	-164.685	-4.316	181.534
SKEL.0	-78.482	-167.509	-3.089	184.983
VILH.0	-83.518	-162.280	-2.286	182.510
BORA.0	-86.311	-181.000	2.814	200.525
JONK.0	-85.247	-182.516	2.569	201.442
SUND.0	-82.091	-172.235	-.757	190.798
HASS.0	-85.175	-187.448	3.681	205.892
NORR.0	-83.236	-182.631	1.602	200.704
ONSA.0	-87.118	-180.688	3.064	200.593
VANE.0	-87.200	-176.616	2.188	196.969
KARL.0	-85.942	-176.145	1.831	195.992
LEKS.0	-84.793	-173.655	.708	193.251
LOVO.0	-81.609	-182.152	.911	199.598
MART.6	-82.365	-177.259	.270	195.460
OSKA.0	-83.240	-187.322	2.648	204.985
OSTE.0	-85.173	-164.340	-1.364	185.100
SVEG.0	-85.134	-169.034	-.282	189.263
UMEA.0	-80.170	-170.019	-1.841	187.972
R.m.s.	**83.240**	**173.903**	**2.619**	**192.799**

The value of the estimated topocentric parameter dz is -30.285 m, which means that the surface of the GRS 80 ellipsoid as a result of the fit lies around 30 metres below that of the Bessel ellipsoid in the area of the topocentre (the barycentre of the common points) of the parameters. The geocentric translations compose the vector between the origins of the geocentric systems.

7.3 3-parameter fit

As a next step we perform a 3 parameter fit with free translations. We then get:

Topocentric parameters:

Translation x:	83.5665813030	
Translation y:	173.1980509464	
Translation z:	-30.2854252412	
Rotation x:	0.0000000000	(fixed)
Rotation y:	0.0000000000	(fixed)
Rotation z:	0.0000000000	(fixed)
Scale correction:	0.00000000	(fixed)

Geocentric parameters:

Translation X:	-497.8058422939
Translation Y:	36.8071586681
Translation Z:	-563.3581152987
Rotation X:	0.0000000000
Rotation Y:	0.0000000000
Rotation Z:	0.0000000000
Scale correction:	0.00000000

Units: metre, arc second and ppm

Table 3:Residuals for 3 parameter fit with free translations (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-4.458	16.989	6.093	17.564
KIRU.0	-9.832	21.725	10.196	23.846
OVER.0	-12.113	15.884	11.361	19.976
SKEL.0	-8.941	11.395	8.672	14.485
VILH.0	-1.441	11.538	3.346	11.628
BORA.0	5.122	-12.133	-7.438	13.170
JONK.0	3.265	-11.997	-5.788	12.433
SUND.0	-2.813	2.929	2.899	4.061
HASS.0	3.901	-17.404	-7.792	17.836
NORR.0	-.214	-9.209	-2.027	9.211
ONSA.0	6.644	-13.241	-9.226	14.814
VANE.0	6.543	-9.005	-7.809	11.131
KARL.0	4.195	-6.400	-4.734	7.652
LEKS.0	1.903	-2.021	-1.847	2.776
LOVO.0	-2.691	-6.787	.823	7.301
MART.6	-1.947	-2.578	1.058	3.231
OSKA.0	.289	-14.219	-3.600	14.222
OSTE.0	1.597	7.269	.175	7.442
SVEG.0	2.066	2.361	-1.124	3.137
UMEA.0	-6.096	7.264	6.209	9.483
R.m.s.	5.332	11.469	6.104	12.648

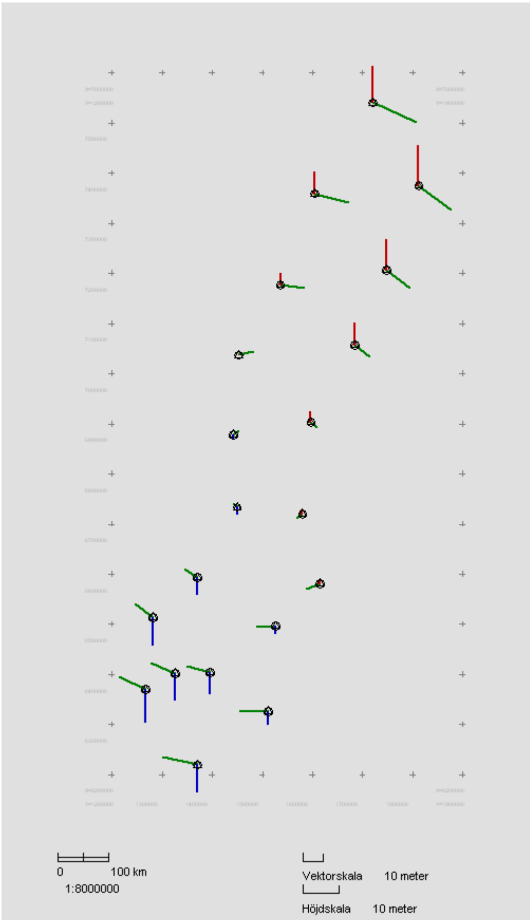

Figure 3: Residuals for 3 parameter fit. Green vectors for horizontal residuals and red/blue for vertical.

As can be seen, the topocentric translations in x and y match the rms values for North and East in table 2 well, which is not wholly unexpected. The size of the translation in topocentric z is exactly equal in the 1- and 3 parameter fits.

If we study the residuals graphically, figure 3, we can see from the green vectors that show the horizontal residuals, that a clear azimuthal rotation between the systems remains that was not modelled by the 3 parameter fit. One can also guess that the topocentre of the parameters, that is the barycentre of the common points, is somewhere to the southeast of Sveg.

Note that the vector scale in the different figures presented in this study varies between the figures, horizontally as well as vertically. Therefore, one cannot simply compare the size of the residuals using the figures. The figures are solely there to help see possible patters indicating systematic differences between the systems, not handled by the transformation model.

7.4 4-parameter fit

To come to terms with the azimuthal rotation we perform a new fit where we allow one rotation, around the topocentric z-axis, as well as the three translations. The result of this operation is shown in table 4 and figure 4.

Table 4: Residuals for 4 parameter fit with free translations and rotation around the topocentric z-axis (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-1.102	-.098	6.093	1.106
KIRU.0	-1.681	-.366	10.197	1.721
OVER.0	-1.174	-.670	11.362	1.352
SKEL.0	-.811	-.547	8.673	.979
VILH.0	-.645	-.106	3.346	.654
BORA.0	-.083	.060	-7.437	.103
JONK.0	-.031	.020	-5.787	.037
SUND.0	-.217	-.311	2.899	.380
HASS.0	.047	.228	-7.791	.232
NORR.0	.068	-.107	-2.027	.127
ONSA.0	-.138	.120	-9.225	.183
VANE.0	-.060	-.054	-7.808	.081
KARL.0	-.006	-.116	-4.734	.116
LEKS.0	-.055	-.142	-1.847	.152
LOVO.0	.180	-.185	.823	.258
MART.6	-.008	-.268	1.058	.268
OSKA.0	.162	.052	-3.599	.171
OSTE.0	-.392	-.098	.174	.404
SVEG.0	-.182	-.157	-1.124	.240
UMEA.0	-.477	-.430	6.210	.642
R.m.s.	.594	.269	6.104	.652

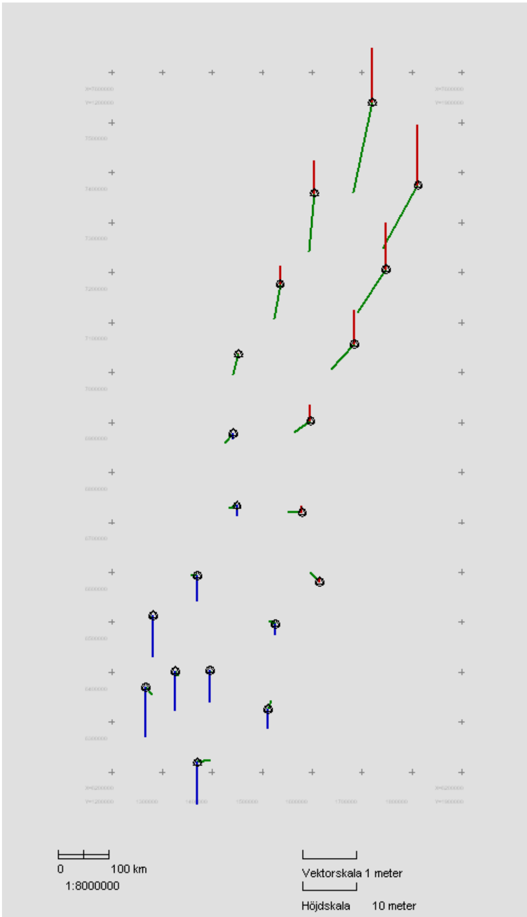

Figure 4: Residuals from fit with 3 translations and rotation around the topocentric z-axis. Green vectors for horizontal residuals and red/blue for vertical.

As can be seen, the residuals decrease considerably in North and East while the Up component is unaffected.

As can be seen, the topocentric translations do not change compared to the 3 parameter fit. The geocentric translations on the other hand are somewhat changed, which is completely in order considering that the rotation matrix $\mathbf{R}_{Topo}$ in equation (6-3) ceases to be an identity matrix with the rotation around the topocentric z-axis

Topocentric parameters:

Translation x:	83.5665813031
Translation y:	173.1980509464
Translation z:	-30.2854252412
Rotation x:	0.0000000000(fixed)
Rotation y:	0.0000000000(fixed)
Rotation z:	6.2909707405
Scale correction:	0.00000000 (fixed)

Units: metre, arc second and ppm

Geocentric parameters:

Translation X:	-497.6534618051
Translation Y:	36.2783418898
Translation Z:	-563.3581193341
Rotation X:	-2.9057790373
Rotation Y:	-0.8372266913
Rotation Z:	-5.5165095465
Scale correction:	0.00000000

From figure 4 we see that the remaining horizontal residuals in northern Sweden are significantly larger than those in the southern part of the country. Furthermore, there seems to be a north-to-south tilt between the ellipsoid surfaces, since the Up components south of the barycentre have the opposite sign of those to the north.

7.5 5-parameter fit

The tilt visible in figure 4 can be rectified by also allowing a rotation around the topocentric y-axis. The next step is therefore to perform a fit that includes rotation around the topocentric y-axis.

Table 5: Residuals for fit with three translations and rotation around the topocentric z- and y-axes (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-.547	.010	-2.124	.547
KIRU.0	-.915	.015	-.639	.916
OVER.0	-.725	-.254	2.942	.768
SKEL.0	-.492	-.321	2.679	.587
VILH.0	-.278	-.102	-2.224	.296
BORA.0	.246	.234	-1.733	.339
JONK.0	.322	.136	-.092	.349
SUND.0	-.014	-.291	1.320	.292
HASS.0	.594	.429	.569	.733
NORR.0	.364	-.098	2.325	.377
ONSA.0	.193	.362	-3.051	.411
VANE.0	.160	.097	-3.729	.187
KARL.0	.206	-.047	-1.810	.212
LEKS.0	.141	-.134	-.966	.195
LOVO.0	.421	-.217	3.942	.474
MART.6	.191	-.272	2.146	.332
OSKA.0	.608	.080	3.224	.613
OSTE.0	-.132	-.149	-3.365	.199
SVEG.0	.017	-.179	-2.350	.179
UMEA.0	-.234	-.328	2.393	.403
R.m.s.	.413	.221	2.416	.469

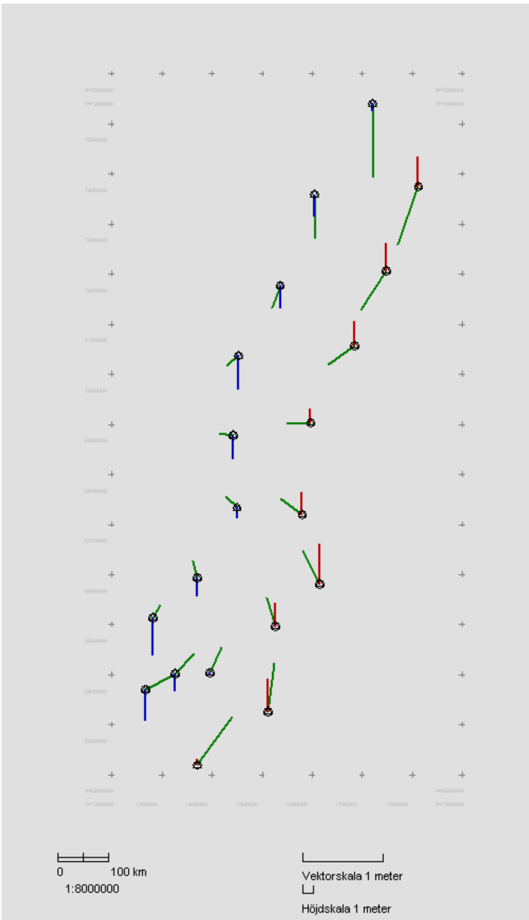

Figure 5: Residuals from fit with 3 translations and rotations around the topocentric z- and y-axis. Green vectors for horizontal residuals and red/blue for vertical.

As expected, the residuals in the Up components are considerably reduced. In northern Sweden from just over 10 m to 2-3 m. Rms decreases to a third. The East components are the ones that are the least altered, which is natural since the y-rotation is around an axis oriented in the east-west direction, see figure 2. The rms-value of the North components is reduced from 6 dm to about 4 dm.

Moreover, column 2D of the table as well as figure 5 show that the horizontal residuals in northern Sweden have decreased at the expense of those in the south. Visible in the figure is also a trend that the horizontal error vectors in the south point to the north and those in the north to the south. The length of the vectors grow with the distance from the barycentre. This indicates a scale difference that has not been modelled.

Another trend that is now more clearly seen is that the Up components in the west have the opposite sign of those in the east. This suggests a tilt between the ellipsoid surfaces in the east-west direction, which means that a rotation around the topocentric x-axis is needed.

We shall also have a look at the parameters in this case.

Topocentric parameters:			**Geocentric parameters:**	
Translation x:	83.7598521439		Translation X:	-419.6034892653
Translation y:	173.1980450421		Translation Y:	58.7696070963
Translation z:	-30.2854266534		Translation Z:	-608.1837657355
Rotation x:	0.0000000000	(fixed)	Rotation X:	-3.7451134820
Rotation y:	-3.0142533647		Rotation Y:	2.0578150035
Rotation z:	6.3012847571		Rotation Z:	-5.5255071173
Scale correction:	0.00000000	(fixed)	Scale correction:	0.00000000

Units: metre, arc second and ppm

The change in the topocentric parameters compared to the 4 parameter fit is minute. The geocentric parameters, on the other hand, change considerably. The change in the rotations is no surprise since the newly added rotation around the topocentric y-axis naturally is divided among all three axial rotations in the geocentric parameters. The geocentric translations are changed with 20-80 m. Studying equation (6-3) that is not unreasonable considering a rotation of 1 arc second moves the points up to 30 m.

We now continue with a fit with three translations and rotations around all three axes.

7.6 6-parameter fit

We begin by looking at the residuals.

The rotation around the topocentric x-axis causes a decrease of the error in the Up component of almost 20 times, with the rms dropping from 2.42 m to 0.13 m. One may guess that this rotation is related to defects in the handling of the geoid in RT 90, since we know from other studies that the geoid has a clear tilt in the east-west direction.

In the horizontal components there is only a marginal improvement.

***Table 6**: Residuals for fit with three translations and rotation around the topocentric z-, y- and x-axes (unit: metre).*

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-.647	-.162	-.204	.667
KIRU.0	-.724	-.594	-.236	.937
OVER.0	-.367	-.405	-.054	.547
SKEL.0	-.358	-.209	.012	.414
VILH.0	-.436	.040	.023	.438
BORA.0	.406	.142	-.028	.430
JONK.0	.349	.044	-.066	.352
SUND.0	-.110	.036	.124	.116
HASS.0	.658	-.079	-.103	.662
NORR.0	.201	-.020	-.110	.202
ONSA.0	.484	.211	-.143	.528
VANE.0	.364	.216	-.132	.423
KARL.0	.253	.152	.235	.295
LEKS.0	.078	.180	.168	.197
LOVO.0	.190	.000	-.042	.190
MART.6	.052	.031	.030	.060
OSKA.0	.393	-.135	-.102	.415
OSTE.0	-.299	.161	-.116	.340
SVEG.0	-.082	.180	.005	.198
UMEA.0	-.244	-.065	.197	.252
R.m.s.	.384	.206	.129	.436

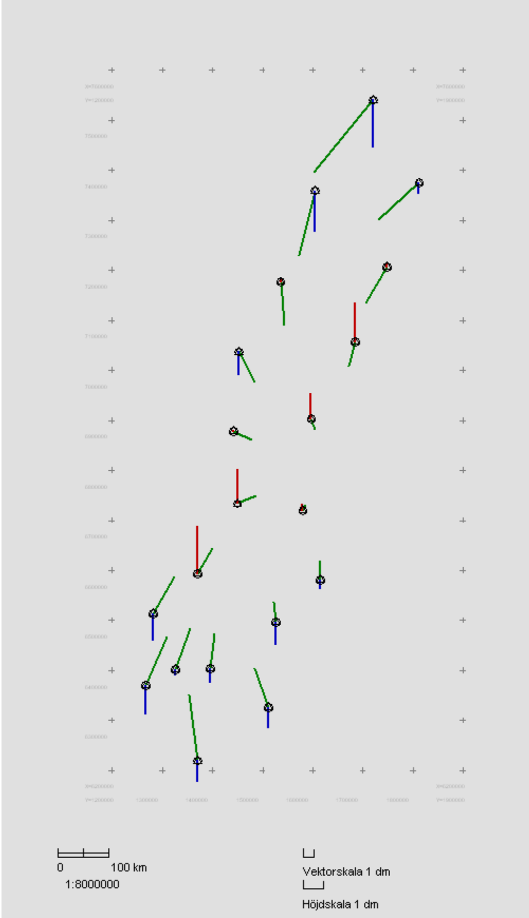

***Figure 6**: Residuals from fit with 3 translations and rotations around the topocentric z- , y- and x-axis. Green vectors for horizontal residuals and red/blue for vertical.*

Topocentric parameters:

Translation x:	83.6624623462	
Translation y:	173.5181699139	
Translation z:	-30.2854294630	
Rotation x:	4.9926229146	
Rotation y:	-1.4952101782	
Rotation z:	*6.2479875359*	
Scale correction:	0.00000000	(fixed)

Units: metre, arc second and ppm

Geocentric parameters:

Translation X:	-416.3877417780
Translation Y:	-100.2165160226
Translation Z:	-585.5944844559
Rotation X:	0.9069425078
Rotation Y:	1.8174190607
Rotation Z:	-7.8786778950
Scale correction:	0.00000000

Again, the change of the topocentric translations compared to the previous fit is only slight. Nor is the rotation around the topocentric z-axis much affected. The possibility of rotation around the topocentric x-axis has effect also on the rotation around the y-axis.

As can be seen, the geocentric parameters have once again been the subject of a dramatic rearrangement.

Looking at the residuals graphically, figure 6, there is no longer a significant unmodelled tilt between the ellipsoid surfaces. However, a certain systematic is discerned in the vertical residuals. The scale effect in the horizontal components naturally remains, and we continue with the next step of our study.

7.7 7-parameter fit

In this case it can be briefly mentioned that the translations and rotations change only minutely compared to the 6 parameter fit, regarding both the topocentric parameters as well as the geocentric. The value of the estimated scale correction is 1.01032050 ppm.

As will be shown in a later section there are in certain circumstances a very strong correlation between the scale correction and the topocentric shift dz.

Table 7: Residuals from fit with 3 translations, 3 rotations and scale correction (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-.083	-.052	-.164	.097
KIRU.0	.006	-.325	-.175	.325
OVER.0	.180	-.044	-.007	.185
SKEL.0	.037	.060	.044	.070
VILH.0	-.051	.066	.048	.084
BORA.0	.003	-.030	.000	.030
JONK.0	-.048	-.065	-.039	.081
SUND.0	-.003	.122	.139	.122
HASS.0	.075	-.207	-.062	.220
NORR.0	-.100	-.010	-.090	.101
ONSA.0	.042	-.013	-.110	.044
VANE.0	.068	-.002	-.108	.068
KARL.0	.045	.013	.253	.047
LEKS.0	.016	.116	.182	.117
LOVO.0	-.028	.095	-.024	.099
MART.6	-.025	.095	.044	.098
OSKA.0	-.079	-.139	-.071	.160
OSTE.0	-.056	.096	-.097	.111
SVEG.0	.001	.106	.020	.106
UMEA.0	.011	.121	.218	.121
R.m.s.	.064	.116	.119	.132

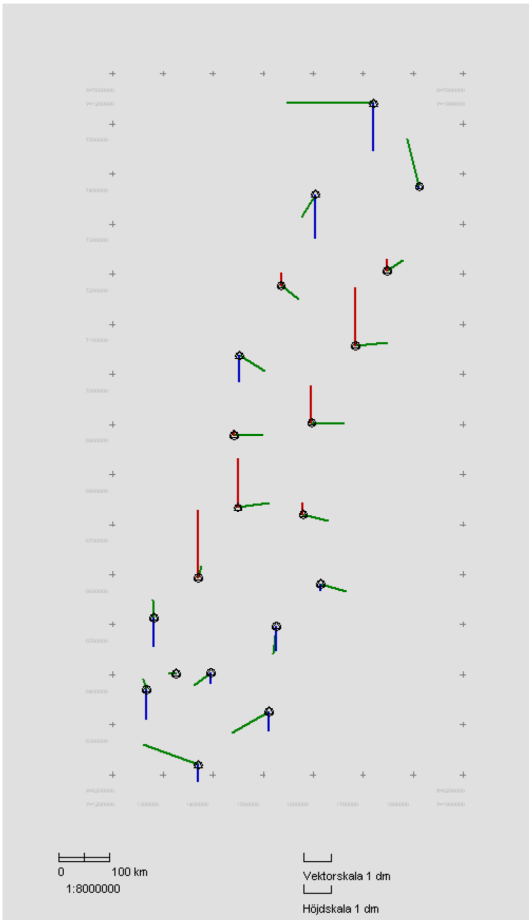

Figure 7: Residuals from fit with 3 translations, 3 rotations and scale correction. Green vectors for horizontal residuals and red/blue for vertical.

Studying the size of the residuals in table 7, we once again see a substantial decrease in the horizontal components. Rms for North drops from 0.384 m to 0.064 m and for East from 0.206 m to 0.116 m. The effect on the Up component is smaller.

When looking at the graphic image of the horizontal residuals, one may notice two whirls, one anti-clockwise for the northern points and one clockwise for the southern ones.

So far, all fits have been made with equal weights on all included components. Studying table 7, there is nothing directly indicating that for example the Up component would pose a problem, on the contrary, the residuals in this component are surprisingly small. The only disturbing factor is that the horizontal accordance still does not match the expected accuracy of RT 90 considering the strength of the adjusted triangulation net, together with the nets of the surrounding Nordic countries, including about 3800 points within Sweden in a homogenous net with sides of 10 km. Moreover, considering that the points in the network are connected by around 15000 distance- and 1500 direction measurements and where the expected accuracy between adjacent points was estimated to 1-2 cm, it is surprising that the fit to the SWEREF systems show discrepancies of over 30 cm.

As implied in a previous section there were certain problems in the geodetic definition of RT 90. Evidently, the standard version of the 7 parameter fit cannot cope with modelling the deformation, possibly caused by the flaws in the geodetic definition. An analysis of the problem points primarily in the direction of the cause being systematic errors in the geoid model used. As a next step we shall therefore perform a fit where the constraint in height is removed. We accomplish this by assigning a higher à priori standard deviation to the height components. With the height constraint gone we introduce a strong correlation between the topocentric translation dz and the scale correction δ. To avoid an ill-conditioned equation system we therefore fix δ to the value 0. We will take a closer look in the relation between scale and height translation in a later section.

7.8 Weighted fit without height constraint

(Appendix 1 contains the complete result file of the fit)

Table 8: Residuals for fit with three translations and three rotations, but without height constraint (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	-.0357	.0232	-7.9041	.0426
KIRU.0	-.0389	-.0723	-7.5773	.0820
OVER.0	.0508	.0431	-5.9735	.0667
SKEL.0	-.0188	.0369	-5.8436	.0414
VILH.0	.0036	.0199	-7.5933	.0202
BORA.0	-.0505	.0267	-6.5201	.0571
JONK.0	-.0517	-.0144	-5.9488	.0537
SUND.0	.0191	.0077	-5.9215	.0206
HASS.0	.0764	.0030	-5.4994	.0765
NORR.0	-.0403	-.0331	-5.2133	.0522
ONSA.0	-.0580	.0719	-7.0300	.0924
VANE.0	-.0114	-.0233	-7.4501	.0259
KARL.0	.0209	-.0453	-6.6183	.0499
LEKS.0	.0286	.0088	-6.5250	.0299
LOVO.0	.0506	.0168	-4.6851	.0533
MART.6	.0155	-.0115	-5.4631	.0193
OSKA.0	.0141	-.0534	-4.6677	.0552
OSTE.0	-.0046	-.0144	-7.9372	.0151
SVEG.0	.0245	-.0193	-7.3052	.0312
UMEA.0	.0026	.0348	-5.6617	.0349
R.m.s.	**.0369**	**.0349**	**6.4462**	**.0508**

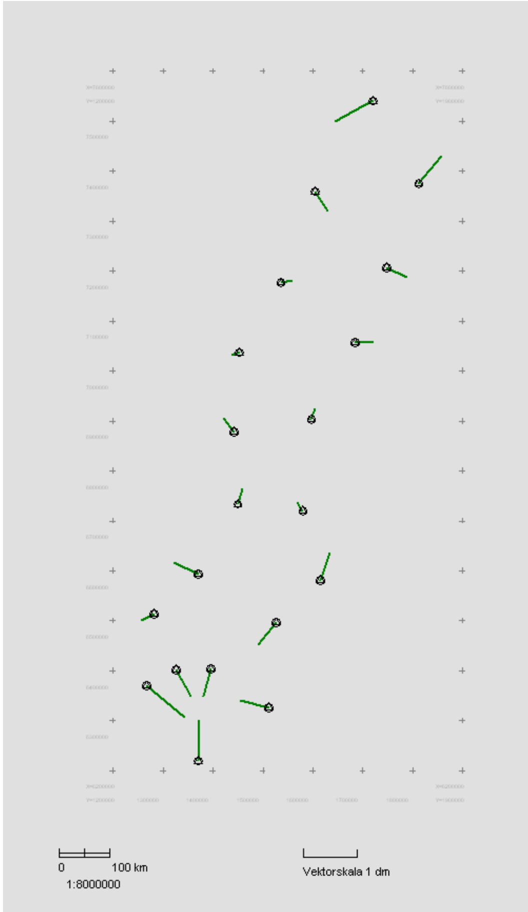

Figure 8: Residuals from fit with 3 translations and 3 rotations but without height constraints. The vectors show the horizontal residuals.

Topocentric parameters:

Translation x:	83.6859793085	
Translation y:	173.4068423468	
Translation z:	-36.6385863800	
Rotation x:	3.1751605455	
Rotation y:	-2.2943202986	
Rotation z:	6.2681584553	
Scale correction:	0.00000000	(fixed)

Units: metre, arc second and ppm

Geocentric parameters:

Translation X:	-414.0978562888
Translation Y:	-41.3381702518
Translation Z:	-603.0627127551
Rotation X:	-0.8550428002
Rotation Y:	2.1413464567
Rotation Z:	-7.0227212665
Scale correction:	0.00000000

We see that it is primarily the parameters for the translation in topocentric z and the rotations around the topocentric x- and y-axes that have changed. The change in the z-translation is about 6 m, which compensates for the scale correction of 1 ppm observed before.

From table 8 we can see that the residuals are now approaching a level more closely coinciding with the expected accuracy of RT 90. Of course, this is at the expense of the Up component, but as we shall see from the discussion further ahead, this is also not a problem. A small cosmetic flaw in this context is that the largest horizontal residual is in Onsala, the most renowned of the geodetic stations in the country.

An inherent weakness of all comparisons done on the rms values is that they do not take into consideration that the number of over-determinations decreases at the same rate as the number of estimated parameters grows. Comparing rms from the standard fit of 7 parameters with the one without height constraint, in the first case 60 equations are used to estimate 7 parameters while in the latter case 40 equations determines 6 parameters. Even with regard to that, the fit without height constraint led to a significant improvement of the horizontal accordance.

In figure 8, a prominent vector pattern for the 6 southernmost stations indicates a diverging scale for the south part of the country. This is not entirely implausible since the south part of the country is primarily surveyed with the microwave instrument Tellurometer, while from the valley of Mälaren and north Geodimeter has been used, an instrument that measures with visible light. It is a known fact that there is a scale difference between these instrument types. Obviously the attempt, made in connection with the adjustment that formed the basis of RT 90, to correct for this scale difference was not entirely successful. In a minor study, not presented in any detail here, a fit without height constraint based on the 6 southern stations gives a standard deviation of 18 mm (2D, 1σ).

7.9 The effect of the scale factor in 3 dimensions

In this section we will investigate the influence of scale on our coordinates. We begin by studying what happens in 3-dimensional space.

We start with the RR 92 coordinates for the 20 SWEPOS stations. As a first step we perform a 3D Helmert transformation with all parameters set to 0 except for the scale correction which is set to 1 ppm. Then we subtract the original coordinates from the transformed. The result is shown in table 9.

The 3D re-scaling means all points move away from each other. Seen from the origin of the Cartesian system, which is in the centre of the ellipsoid, all stations are moved outwards radially. Since the distance of the points to the origin is roughly 6360 km they will end up roughly 6.36 m above the ellipsoid surface. Because of flattening, the radial direction somewhat diverges from the direction of the normal in each point, which explains the difference of 15-20 mm in latitude.

Table 9: Difference between 3D coordinates that has a scale changed by 1 ppm minus the original coordinates (unit: metre).

Station	Topocentric components		
	North	East	Up
ARJE.0	-.016	.000	6.360
KIRU.0	-.015	.000	6.360
OVER.0	-.016	.000	6.360
SKEL.0	-.016	.000	6.360
VILH.0	-.017	.000	6.360
BORA.0	-.019	.000	6.362
JONK.0	-.019	.000	6.362
SUND.0	-.018	.000	6.361
HASS.0	-.020	.000	6.363
NORR.0	-.019	.000	6.362
ONSA.0	-.019	.000	6.362
VANE.0	-.019	.000	6.362
KARL.0	-.019	.000	6.362
LEKS.0	-.018	.000	6.362
LOVO.0	-.019	.000	6.362
MART.6	-.018	.000	6.361
OSKA.0	-.019	.000	6.362
OSTE.0	-.017	.000	6.361
SVEG.0	-.018	.000	6.361
UMEA.0	-.017	.000	6.360

Thus, 3D re-scaling primarily changes the heights of the points above the ellipsoid. The latitude is minutely changed and the longitude remains completely unaltered. The conclusion is that the scale in a system is affected by the height above the ellipsoid assigned to the points. Consequently, an erroneous geoid model can cause errors in scale.

It should be pointed out that the global systems, such as SWEREF 99, ITRF and so on, in this context can be regarded as error-free. It is the conventional systems, such as RT 90 and ED 87 among others, that can possibly be afflicted with scale errors and defects in the geoid model.

When performing a 3D Helmert fit between a global system and a conventionally defined system using the same weight on all observation equations, a possible scale difference is modelled by a scale correction.

When performing a 3D Helmert fit without height constraint one can choose to fix the scale correction to 0. Firstly, one gets the best possible accordance of the horizontal coordinates and secondly, the ellipsoid is placed relative to the earth surface in such a way as to eliminate the scale error. Then, the height residuals can be used to modify an incorrect geoid model. Note that this is true either the scale error is caused by an incorrect geoid model or by faults in distance measuring instruments or in the facilities used for calibration of distance measurement techniques.

In other words, the system has been given a more correct definition without having to change the coordinates. With the improved definition, distances, for example, can be reduced to the ellipsoid without the need to apply a scale correction.

7.10 The effect of the scale factor in 2 dimensions

We start off from the geodetic RR 92 coordinates of the SWEPOS stations. As a first step they are projected with Gauss-Krüger's projection to 2.5 gon V, but instead of the x/y additions (0/1500000) we use (-6500000/0). This places the origin of the projected coordinates on the central meridian some ten kilometres south of Finspång. The next step is to re-scale the planar coordinates and then change them back to latitude, longitude and height above the Bessel ellipsoid again. As in the 3D case, we take the difference between the geodetic coordinates with changed scale and the original ones, see table 10.

Table 10: Difference between coordinates with a scale change of 1 ppm in the projection plane and corresponding original coordinates (unit: metre).

Station	Topocentric components			
	North	East	Up	2D
ARJE.0	.856	.136	.000	.866
KIRU.0	1.018	.308	.000	1.064
OVER.0	.834	.407	.000	.928
SKEL.0	.685	.306	.000	.750
VILH.0	.677	.044	.000	.678
BORA.0	-.104	-.170	.000	.199
JONK.0	-.099	-.101	.000	.142
SUND.0	.401	.108	.000	.415
HASS.0	-.284	-.121	.000	.309
NORR.0	-.004	.026	.000	.026
ONSA.0	-.143	-.225	.000	.267
VANE.0	.002	-.219	.000	.219
KARL.0	.089	-.133	.000	.160
LEKS.0	.233	-.054	.000	.240
LOVO.0	.077	.117	.000	.141
MART.6	.219	.085	.000	.235
OSKA.0	-.173	.012	.000	.173
OSTE.0	.536	-.055	.000	.539
SVEG.0	.378	-.064	.000	.383
UMEA.0	.546	.216	.000	.587

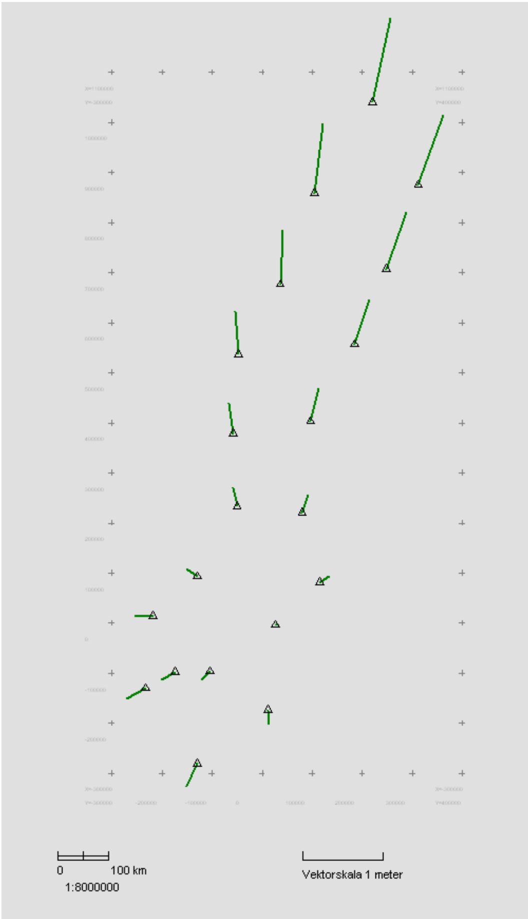

Figure 9: Difference between planar coordinates with the origin just south of Finspång, with a scale change of 1 ppm, and corresponding original coordinates.

The table shows how much the points have been moved relative to the point south of Finspång. Obviously, the heights are not affected.

If we perform a 3D Helmert fit between the coordinates re-scaled in the projection plane and the original coordinates, the scale change will be modelled by the transformation so that the discrepancy < 1 mm in the horizontal components (latitude and longitude) where the scale change is 1 ppm. However, because of different metrics in the projection plane compared to 3-dimensional space, the 3D Helmert fit will not be able to fully model the difference in scale. The divergence grows with the size of the scale difference and amounts to 0-5 mm at a scale difference of 10 ppm.

7.11 How does the fit work without height constraint?

Before we proceed to discuss and conclude the results of the different studies of the Helmert fit carried out, we shall for a moment return to our antennae-equipped ellipsoids.

Each antenna tip is in a position given by the geodetic coordinates (φ, λ, h) for each ellipsoid. Performing a fit with equal weight for all observation equations means that the positions of the ellipsoids are modified relative to each other so that the sum of the squares of the distances between the antennae tips is minimized. In the examples, the fit has always meant computation of parameters for transformation from SWEREF 99 to RR 92. Studying equation (5-10), representing the three observation equations for the i^{th} point, we see that the vector of observational corrections (residual vector) v is expressed in the topocentric system on Bessel's ellipsoid whose origin has the RT 90 coordinates (φ_i, λ_i). The vector has its starting point in the antenna tip on Bessel's ellipsoid and its end point, after the fit is performed, coincides with the corresponding antenna tip on GRS 80, see figure 10a.

What happens when we let go of the height constraint? As before, the vector of observational corrections runs from antenna tip to antenna tip, but now the length of the vector is insignificant; what is minimized is the sum of the squares of the horizontal components of the vectors or, differently put, for each antenna tip on the GRS 80 ellipsoid we measure the distance to the line that coincides with the antenna on Bessel's ellipsoid. A fit with no height constraint means that we minimize the sum of the squares of all these distances, se figure 10b.

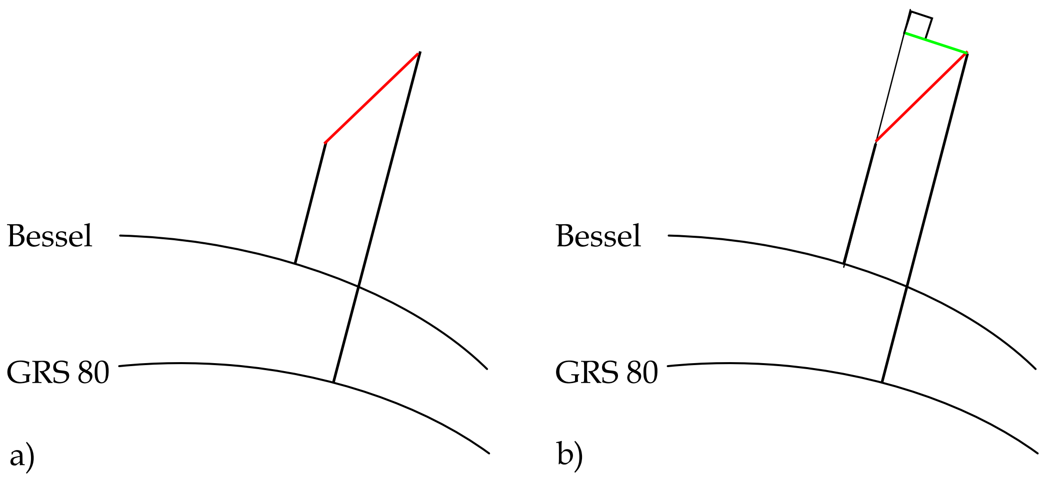

Figure 10: Relation with height constraint a) and without height constraint b), respectively.

7.12 Discussion of the results

As mentioned in one of the previous sections, we are primarily interested in transforming the horizontal coordinates. The disadvantage of the prevalent procedure for computing the seven parameters is that all observation equations are given the same weight. The fit in height comes at the cost of the horizontal fit. Re-formulating the observation equation according to equation (5-10) solves this problem.

The method of estimating topocentric parameters increase our understanding of how the 3D Helmert works. For example, we see that the previously estimated parameters change moderately when more parameters are gradually added in the estimate. The geocentric parameters, on the other hand, are massively changed. Intuitively it is easy to see that a small rotation around the topocentric x- or y-axis leads to a substantial change in the relative position of the ellipsoid centres, which is equivalent to a change in the geocentric translations. A rotation of one arc second changes the position with around 30 m. It is common when presenting the parameters to also state the standard deviation for each parameter. The author's opinion is that this information is of no interest and often quite misleading. If repeated fits are performed within the same area, but based on different common points, the divergence of the parameters may greatly exceed the stated standard deviations. On some occasion this led to the results of a surveying campaign being wrongly questioned.

As stated earlier, erroneous entered heights can in a coordinate transformation cause errors in the horizontal position, because of the ellipsoids not being parallel. We have also confirmed that the transformation should not be used outside the area defined by the common points. We primarily had the horizontal distribution of the points in view, but the same principle is of course valid vertically. The vertical side brings an interesting issue. The common points are generally triangulation points and are therefore usually situated on mountain tops while the points one wishes to transform are situated down in the communities. As an extreme case one can imagine an alpine village at 1000 m height surrounded by 2000-3000 m high Alps. A fit without height constraint would in this case give parameters that minimize the sum of the squares of the residuals at the level of the Alps, compare to figure 10b. The conclusion of this is that one should perform the fit with the heights of the global system (the from-system) set to 1000 m. A more general approach is to always set the heights of the global system to 0. In principle, this procedure can be applied for all fits without height constraint. One then gets parameters that minimize the residuals at the surface of the GRS 80 ellipsoid. Also when transforming remaining points, the heights of the global system should be set to 0. If one wishes to perform the transformation in the opposite direction, the local heights should really be set to the distance between the ellipsoid surfaces in each point, but since these distances generally are small, amounting to some tens of metres, setting the entering heights to 0 will suffice here as well.

A fit without height constraint minimizes the horizontal residuals. Does that mean that this method should always be used? The answer is no. Simply put, one may say that the method should only be used if it significantly reduces the horizontal residuals. What primarily happens when the height constraint is removed is that one allows larger rotations around the topocentric x- and y-axes. This causes a tilt between the ellipsoid surfaces compensating for systematic height errors, for example induced by defects in the geoid model. The change in the horizontal coordinates caused by such systematic errors in the height model is moderate. The problem when performing a fit without height constraint is that the method of least squares cannot distinguish if the residuals are caused by systematic errors or by pure measurement errors in the triangulation net. If the measurement noise is at a higher level than the systematic errors, there is a risk that the topocentric rotations are trying to compensate for errors in the coordinates of individual points caused by measurement errors. This is particularly the case if the number of common points is small and/or the area of the fit is small. For national systems, such as RT 90 and the Finnish KKJ, with a large coverage area with many common points of good quality, the probability of the fit modelling possible systematic errors related to the definition of the system is high.

We now leave the 3D Helmert transformation and instead occupy ourselves with the following question.

8 Projection fit

8.1 Background

As mentioned in the introduction, the municipalities in Sweden use grid systems. How the systems were established varies, but very few have a geodetic definition that make it possible to directly convert the x/y grid coordinates to latitude and longitude, which causes considerable problems when creating a transformation to SWEREF 99.

One of the reasons for the RIX 95 project was to estimate transformation parameters between the municipal systems and the national systems RT 90 and SWEREF 99. A special group was formed solely to handle the question of how these sets of transformation parameters were to be produced. The conclusions of the group were that the transformations should be as accurate as possible, involve as few steps as possible and use standard methods implemented in most software for geodesy- and GIS applications, something not entirely easy to achieve.

An investigation of the market showed that most software could do Transverse Mercator projection in both directions as well as 3D Helmert transformation. Surprisingly, 2D Helmert transformation was for the most part lacking.

A way to establish transformation parameters without involving 2D Helmert transformation was to use a method, developed a few years previously, based on the Transverse Mercator (TM) projection that was called projection fit. The idea of this approach emanated from Ilmar Ussisoo. What he did was to project (φ, λ) in ED 50 to (x, y) on Hayford's ellipsoid with central meridian 15°48'23.0", after which he fitted the results to RT 38 2.5 gon V with a 2D Helmert transformation.

Since the TM projection contains the possibility of re-scaling and false northing and easting, which in principle corresponds to scale and translations in 2D Helmert transformation, it gave inspiration to try to find projection parameters, directly giving grid coordinates, that as well as possible coincided with RT 90 2.5 gon V, without applying a 2D Helmert transformation. In the beginning this was done by trial and error; one guessed an appropriate value for the central meridian, transformed and performed a Helmert fit. Then the scale and translation parameters from the fit were used as projection parameters in a new computation round where the longitude of the central meridian was adjusted. This procedure was repeated until the results were probably good. Compared to Ilmar Ussisoo's method the result was worse. This led to the abandonment of the half-manual computing and instead a Fortran programme was written to solve for all projection parameters at the same time according to the method of least squares and thus, the method of projection fit was born. That Ussisoo's approach was not sufficient depended on the fact that 2D Helmert transformation was not implemented in some GPS equipment and one wanted the possibility to offer a transformation for all hand-held GPS receivers, so that one could easily get a position that could be found on the map.

We shall now look at this method more closely.

8.2 Representation $(\varphi,\lambda) \rightarrow (x,y)$ based on Transverse Mercator projection according to the formulas of Gauss-Krüger

The idea is, therefore, to transform coordinates between two geodetic reference frames with the help of a projection computation, where the parameters of the projection is decided by an iterative procedure analogous to that used for 3D Helmert fit. The method is implemented for Transverse Mercator projection, but should be possible to apply also for other projections. We begin by describing in detail the representation from the (φ, λ) of the ellipsoid surface to the (x, y) of the projection plane according to the method of Gauss-Krüger.

Symbols and definitions:

- a semi-major axis of the ellipsoid
- f flattening of the ellipsoid
- e^2 first eccentricity squared
- φ geodetic latitude, positive direction North
- λ geodetic longitude, positive direction East
- x grid coordinate, positive direction North
- y grid coordinate, positive direction East
- λ_0 longitude of the central meridian
- k_0 scale factor along the central meridian
- x_0 false northing
- y_0 false easting

All angles (latitude, longitude and so on) should be expressed in radians. Note that the x-axis points to the North and the y-axis to the East.

From the ellipsoid parameters *a* and *f* the following entities are computed:

$$e^2 = f(2-f)$$

$$n = \frac{f}{(2-f)}$$

$$\hat{a} = \frac{a}{(1+n)}\left(1+\frac{1}{4}n^2+\frac{1}{64}n^4+\ldots\right)$$

Compute the conform* latitude φ^*

$$\varphi^* = \varphi - \sin\varphi\cos\varphi\left(A + B\sin^2\varphi + C\sin^4\varphi + D\sin^6\varphi + \ldots\right) \tag{8-1}$$

where the coefficients *A*, *B*, *C*, and *D* are obtained from the formulas:

$$A = e^2$$

$$B = \frac{1}{6}\left(5e^4 - e^6\right)$$

$$C = \frac{1}{120}\left(104e^6 - 45e^8 + \ldots\right)$$

$$D = \frac{1}{1260}\left(1237e^8 + \ldots\right)$$

Define ξ' and η' as

$$\xi' = \arctan(\tan\varphi^* / \cos(\lambda-\lambda_0)) \tag{8-2}$$

$$\eta' = \operatorname{arctanh}(\cos\varphi^* \sin(\lambda-\lambda_0)) \tag{8-3}$$

One then obtains

$$\begin{aligned} x = k_0\hat{a}(\xi' &+ \beta_1\sin 2\xi'\cosh 2\eta' + \beta_2\sin 4\xi'\cosh 4\eta' + \\ &+ \beta_3\sin 6\xi'\cosh 6\eta' + \beta_4\sin 8\xi'\cosh 8\eta' + \ldots) + x_0 \end{aligned} \tag{8-4}$$

$$\begin{aligned} y = k_0\hat{a}(\eta' &+ \beta_1\cos 2\xi'\sinh 2\eta' + \beta_2\cos 4\xi'\sinh 4\eta' + \\ &+ \beta_3\cos 6\xi'\sinh 6\eta' + \beta_4\cos 8\xi'\sinh 8\eta' + \ldots) + y_0 \end{aligned} \tag{8-5}$$

---

* Older Swedish literature calls this quantity isometric latitude (isometrisk latitud). Today, the term isometric latitude is used for the quantity $\psi = \ln\{\tan(\pi/4+\varphi/2)[(1-e\sin\varphi)/(1+e\sin\varphi)]^{e/2}\}$. The isometric latitude is computed from the conform latitude according to the formula $\psi = \ln\tan(\pi/4+\varphi^*/2)$. Compare with John P. Snyder: Map Projections - A Working Manual, U.S. Geological Survey Professional Paper 1395.

where the coefficients $\beta_1, \beta_2, \beta_3$ och β_4 are computed as

$$\beta_1 = \frac{1}{2}n - \frac{2}{3}n^2 + \frac{5}{16}n^3 + \frac{41}{180}n^4 + \ldots$$

$$\beta_2 = \frac{13}{48}n^2 - \frac{3}{5}n^3 + \frac{557}{1440}n^4 + \ldots$$

$$\beta_3 = \frac{61}{240}n^3 - \frac{103}{140}n^4 + \ldots$$

$$\beta_4 = \frac{49561}{161280}n^4 + \ldots$$

8.3 Projection fit based on Transverse Mercator projection with the formulas of Gauss-Krüger

Given: *A number of points with known geodetic coordinates (φ, λ). We also know the coordinates (x, y) in a grid system.*

Sought: *A Transverse Mercator projection (or, shorter, TM projection) that converts the given (φ, λ) values into grid coordinates (x, y) that coincides with the given (x, y) values.*

To perform a TM projection one needs to specify the semi-major axis (a) and flattening (f) of the ellipsoid used, the longitude of the central meridian (λ_0), the scale along the central meridian (k_0) and the false Northing and Easting (x_0) and (y_0). We assume that the ellipsoid parameters a and f are known.

Note that the ellipsoid parameters are always got from the system with the given (φ, λ) values.

We regard x and y as functions of the projection parameters according to the following x=x(λ_0, k_0, x_0, y_0) and y=y(λ_0, k_0, x_0, y_0). As usual we do a Taylor series expansion around the approximate values (λ_0), (k_0), (x_0), (y_0). The observation equations then become

$$x + v_x = \mathrm{x}((\lambda_0),(k_0),(x_0),(y_0)) + (\frac{\partial x}{\partial \lambda_0})_0 \Delta\lambda_0 + (\frac{\partial x}{\partial k_0})_0 \Delta k_0 + (\frac{\partial x}{\partial x_0})_0 \Delta x_0 + (\frac{\partial x}{\partial y_0})_0 \Delta y_0 \qquad (8\text{-}6)$$

$$y + v_y = \mathrm{y}((\lambda_0),(k_0),(x_0),(y_0)) + (\frac{\partial y}{\partial \lambda_0})_0 \Delta\lambda_0 + (\frac{\partial y}{\partial k_0})_0 \Delta k_0 + (\frac{\partial y}{\partial x_0})_0 \Delta x_0 + (\frac{\partial y}{\partial y_0})_0 \Delta y_0 \qquad (8\text{-}7)$$

where $\Delta\lambda_0$, Δk_0, Δx_0 and Δy_0 are unknown corrections to the approximate values and v_x and v_y are the residuals of the observed (known) values x and y.

We shall now derive expressions for the partial derivatives. We use the formulas of Gauss-Krüger according to equations (8-4) and (8-5) above and get

$$x = k_0\ \hat{a}\ \mathrm{f}(\xi'(\lambda_0), \eta'(\lambda_0)) + x_0 \qquad (8\text{-}8)$$

$$y = k_0 \; \hat{a} \; \mathrm{g}(\xi'(\lambda_0), \eta'(\lambda_0)) + y_0 \tag{8-9}$$

The partial derivatives are

$$\frac{\partial \mathrm{x}}{\partial k_0} = \hat{a} \; \mathrm{f} \qquad \frac{\partial \mathrm{x}}{\partial x_0} = 1 \qquad \frac{\partial \mathrm{x}}{\partial y_0} = 0 \tag{8-10}$$

$$\frac{\partial \mathrm{y}}{\partial k_0} = \hat{a} \; \mathrm{g} \qquad \frac{\partial \mathrm{y}}{\partial x_0} = 0 \qquad \frac{\partial \mathrm{y}}{\partial y_0} = 1 \tag{8-11}$$

$$\frac{\partial \mathrm{x}}{\partial \lambda_0} = k_0 \hat{a} \left\{ \frac{\partial \mathrm{f}}{\partial \xi'} \frac{\partial \xi'}{\partial \lambda_0} + \frac{\partial \mathrm{f}}{\partial \eta'} \frac{\partial \eta'}{\partial \lambda_0} \right\} \tag{8-12}$$

$$\frac{\partial \mathrm{y}}{\partial \lambda_0} = \mathrm{k}_0 \hat{\mathrm{a}} \left\{ \frac{\partial \mathrm{g}}{\partial \xi'} \frac{\partial \xi'}{\partial \lambda_0} + \frac{\partial \mathrm{g}}{\partial \eta'} \frac{\partial \eta'}{\partial \lambda_0} \right\} \tag{8-13}$$

According to equations (8-4), (8-5), (8-8) and (8-9) we get

$$\mathrm{f}(\xi', \eta') = \xi' + \sum_{i=1}^{4} \beta_i \sin 2i \; \xi' \cosh 2i \; \eta' + \ldots \tag{8-14}$$

$$\mathrm{g}(\xi', \eta') = \eta' + \sum_{i=1}^{4} \beta_i \cos 2i \; \xi' \sinh 2i \; \eta' + \ldots \tag{8-15}$$

Four terms in the series expansion is more than enough for millimetre precision.

From equations (8-14) and (8-15) we get

$$\frac{\partial \mathrm{f}}{\partial \xi'} = 1 + \sum_{i=1}^{4} 2i\beta_i \cos 2i\xi' \cosh 2i\eta' + \ldots \tag{8-16}$$

$$\frac{\partial \mathrm{f}}{\partial \eta'} = \sum_{i=1}^{4} 2i\beta_i \sin 2i\xi' \sinh 2i\eta' + \ldots \tag{8-17}$$

$$\frac{\partial \mathrm{g}}{\partial \xi'} = -\sum_{i=1}^{4} 2i\beta_i \sin 2i\xi' \sinh 2i\eta' + \ldots \tag{8-18}$$

$$\frac{\partial \mathrm{g}}{\partial \eta'} = 1 + \sum_{i=1}^{4} 2i\beta_i \cos 2i\xi' \cosh 2i\eta' + \ldots \tag{8-19}$$

Comment: If we compare equations (8-16) and (8-17) to (8-18) and (8-19), we see that (8-16) and (8-19) are identical. The same is true for (8-17) and (8-18) except for the sign, that is

$$\frac{\partial f}{\partial \xi'} = \frac{\partial g}{\partial \eta'} \quad \text{and} \quad \frac{\partial f}{\partial \eta'} = -\frac{\partial g}{\partial \xi'}$$

This is a general relation that is valid for all conformal representations and is commonly called the differential equations of Cauchy-Riemann.

Finally, from equations (8-2) and (8-3) after some manipulation of the formulas we get the numerically well-behaved formulas

$$\frac{\partial \xi'}{\partial \lambda_0} = -\frac{\sin\varphi^*\cos\varphi^*\sin(\lambda\text{-}\lambda_0)}{\sin^2\varphi^* + \cos^2\varphi^*\cos^2(\lambda\text{-}\lambda_0)} \tag{8-20}$$

$$\frac{\partial \eta'}{\partial \lambda_0} = -\frac{\cos\varphi^*\cos(\lambda\text{-}\lambda_0)}{\sin^2\varphi^* + \cos^2\varphi^*\cos^2(\lambda\text{-}\lambda_0)} \tag{8-21}$$

With one exception, we now have all the information needed to form the observation equations. What is missing is approximate values for the unknowns prior to the first iteration. Since the false northing and easting, x_0 and y_0, are linear parts of equations (8-8) and (8-9), the approximate values for them can be set to 0, but tests show that 0 is good enough also for λ_0. If one wishes to improve the convergence λ_0 can be set to the average of the smallest and the largest longitude of the common points. For k_0, a suitable choice is 1.

The corrections of the parameters are solved for in the over-determined linear equation system with the method of least squares, after which they are added to the approximate values before the next iteration. Normally, the procedure has a rapid convergence.

8.4 Discussion of the usability of the method

There are mainly two factors that limit usability.

The grid coordinates must originate from a TM projection. Different projections deform the representation in the projection plane differently. In the TM projection the deformations grow with the distance to the central meridian, while in, for example, Lambert's projection they grow with the distance to the standard parallels. For very small areas (a few kilometres at most) a grid system with Lambert geometry can be approximated by a TM projection but, as said, the errors grow very rapidly with the size of the area.

The x-axis of the local system must be parallel with the representation of the central meridian. The latter means, for example, that if the local system is regarded as originating from 5 gon V, then the local x-axis should be parallel with the x-axis in 5 gon V. Studies carried out show that bad orientation of the local system makes a good fit between projected and local coordinates impossible. The residuals are proportional to the rotation and grow linearly with the size of the coverage area of the local system. See table 11.

Table 11: Errors in metre caused by rotation of the local system.

	1 mgon	**2 mgon**	**10 mgon**	**100 mgon**	**1000 mgon**	**10000 mgon**
1*1 km²	0,000	0,000	0,000	0,000	0,000	0,004
10*10 km²	0,000	0,000	0,000	0,003	0,031	0,305
50*50 km²	0,001	0,002	0,008	0,077	0,770	7,700
100*100 km²	0,003	0,006	0,031	0,308	3,080	30,080

Apart from errors caused by the local system not having TM geometry or by it being rotated, there are discrepancies caused by scale differences in the definition of the reference frames (the curvature of the ellipsoids etc). These effects grow with the size of the area. Current studies indicate that the effect on Sweden as a whole does not exceed 1-2 dm. For smaller areas, (50 * 50 kilometres), the error is probably less than 1 millimetre.

In conclusion, it can be said that, concerning the municipal systems, what primarily causes problems are defects in the orientation. The threshold is somewhere around 1-5 mgon. Unfortunately, it appears that some municipal/local systems have a rotation amounting to several gon. For these systems it is necessary to also consider the rotation. We then need to combine the projection fit with either a planar or a 3-dimensional Helmert transformation, (2D/3D Helmert). If we use a 2D Helmert transformation, the estimate of the Helmert parameters can be done in the same least squares fit as the projection parameters. In the next section we run through how the observation equations are set up in this case.

9 Projection fit combined with a planar Helmert transformation

Since both the TM projection and the 2D Helmert transformation include a scale factor and shifts in the x- and y-coordinates, certain confusion arises when naming the entering variables. For the TM projection we keep all names of variables from the previous section, with the exception of renaming the coordinates obtained from the left-hand side of equations (8-14) and (8-15) x' and y'. These are the coordinates that are to be further transformed with the planar Helmert transformation. Like 3D Helmert transformation, 2D Helmert transformation is a similarity transformation, but because it works in two dimensions we only get four parameters; a scale factor (s_H), a rotation (α), and two translations (x_{0H} and y_{0H}). We use the following formula to describe the planar Helmert transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_{0H} \\ y_{0H} \end{pmatrix} + s_H \mathbf{R} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{9-1}$$

where **R** is a rotation matrix defined as the following

$$\mathbf{R} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \tag{9-2}$$

The final coordinates x and y are now functions of eight parameters, $x=x(\lambda_0, k_0, x_0, y_0, s_H, \alpha, x_{0H}, y_{0H})$ and $y=y(\lambda_0, k_0, x_0, y_0, s_H, \alpha, x_{0H}, y_{0H})$.

The observation equations are set up analogous to equations (8-6) and (8-7), but with the difference of adding the terms for Δs_H, $\Delta\alpha$, Δx_{0H} and Δy_{0H}. The partial derivatives are also somewhat modified

$$\begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial k_0} \\ \dfrac{\partial \mathrm{y}}{\partial k_0} \end{pmatrix} = s_H \mathbf{R} \begin{pmatrix} \dfrac{\partial \mathrm{x}'}{\partial k_0} \\ \dfrac{\partial \mathrm{y}'}{\partial k_0} \end{pmatrix} \text{and} \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial \lambda_0} \\ \dfrac{\partial \mathrm{y}}{\partial \lambda_0} \end{pmatrix} = s_H \mathbf{R} \begin{pmatrix} \dfrac{\partial \mathrm{x}'}{\partial \lambda_0} \\ \dfrac{\partial \mathrm{y}'}{\partial \lambda_0} \end{pmatrix}, \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial x_0} \\ \dfrac{\partial \mathrm{y}}{\partial x_0} \end{pmatrix} = s_H \mathbf{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{and} \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial y_0} \\ \dfrac{\partial \mathrm{y}}{\partial y_0} \end{pmatrix} = s_H \mathbf{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The partial derivatives for x' and y' with respect to k_0 and λ_0 above are obtained by equations (8-10), (8-11) and (8-12), (8-13).

For the new parameters we get

$$\begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial s_H} \\ \dfrac{\partial \mathrm{y}}{\partial s_H} \end{pmatrix} = \mathbf{R} \begin{pmatrix} x' \\ y' \end{pmatrix} \text{and} \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial \alpha} \\ \dfrac{\partial \mathrm{y}}{\partial \alpha} \end{pmatrix} = s_H \frac{\partial \mathbf{R}}{\partial \alpha} \begin{pmatrix} x' \\ y' \end{pmatrix}, \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial x_{0H}} \\ \dfrac{\partial \mathrm{y}}{\partial x_{0H}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{and} \begin{pmatrix} \dfrac{\partial \mathrm{x}}{\partial y_{0H}} \\ \dfrac{\partial \mathrm{y}}{\partial y_{0H}} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

As before, the parameters are computed with the method of least squares by solving for corrections to approximate values of the sought parameters. The procedure is iterated by adding the estimated corrections from the previous iteration step to the approximate values. As starting values for the iteration the scale, s_H, can be set to 1. For the remaining three, (α, x_{0H} och y_{0H}), 0 is a suitable starting value.

Since both the TM projection and the 2D Helmert transformation contain an unknown scale factor as well as translations in x and y, a complication arises. If one tries to estimate all parameters in one single fit, the linear equation system becomes singular. Consequently, one has to fix one of the scale factors and one of the translations in x and y, respectively, to values previously determined. For example, one can set the TM scale to 1 and the false northing and easting to 0 and 1500000, respectively.

As mentioned earlier, in the initial phase of the RIX 95 project many software on the market lacked the possibility to perform a 2D Helmert transformation, but in most of them a 3D Helmert was implemented. The next section deals with the issue of how to combine the TM projection with a 3D Helmert transformation.

10 Projection fit combined with a 3D Helmert transformation

The transformation chain for conversion from (φ, λ) to (x, y), when the parameters once are determined, is in this case to first do a 3D Helmert followed by a TM projection.

The basic problem when determining the parameters of the transformation, is that we do not have any given way of converting the grid coordinates (x, y) to geodetic coordinates (φ, λ). If we had, it would be a trivial matter to compute the parameters for the 3D Helmert transformation. Unfortunately, no observation equations that make it possible to estimate the parameter sets for both the 3D Helmert transformation and the TM projection in the same fit has been formulated. Instead, a procedure founded on the following reasoning has been chosen.

The task of the Helmert transformation, either it takes place in two or three dimensions, is to handle the rotation between the systems since it cannot be modelled by the TM projection. It is therefore reasonable to assume that the same TM projection can be used with both a 2D and a 3D Helmert transformation.

The approach is then, as a first step, to perform a combined TM- and 2D Helmert fit. In the next step the projection parameters obtained are used to transform the grid coordinates of the common points to fictitious latitude- and longitude values. As a last step, a 3D Helmert fit is done between the global system, SWEREF 99 in our case, and the fictitious system. The procedure has the advantage of the same projection parameters being used whether or not the projection is combined with a 2D or 3D Helmert transformation.

Since the municipal systems cover relatively small areas, the horizontal accordance is not significantly improved by performing the fit without height constraint; and there is a risk of getting unrealistically large rotations around the topocentric x- and y-axes in this case. One would rather want the rotations to be as small as possible, to make the transformation agree with the 2D Helmert variant. When working with RIX 95, the 3D fit is therefore made with height constraint, but with the heights set to zero in both systems.

Normally the accuracy in a problem involving a fit is reduced if the computation is divided into two separate fits. Because of the nature of the problem, the chosen approach gives residuals at the same level whether you combine the TM projection with a 2D or a 3D Helmert transformation. The difference in the individual x/y coordinates is however noticeable, but generally does not exceed 5-10 mm. For small rotations the effect is merely 1-2 mm in some occasional point.

11 Implementation in RIX 95

Finally, a short presentation of how the transformation methods described in the previous sections are implemented in the computation routine used to determine the sets of transformation parameters in RIX 95 is given.

As mentioned initially, a goal for the work of producing the transformation parameters was that they should be as accurate as possible, have as few transformation steps as possible and be based on standard methods implemented in most software for geodetic applications as well as for GIS applications. Undoubtedly, transformation parameters produced by projection fit meet these requirements. However, some manual work is needed, especially when the projection fit is combined with a 2D or 3D Helmert transformation. At the time of writing the program code for the transformation methods described in previous sections, there were commercial software for TM projection in which only relatively few digits could be entered for the longitude of the central meridian (λ_0). If the software presented λ_0 with more digits than what the user's program could accept, the user had to round it off, which could give a systematic error in the transformed coordinates. For example, rounding the second part off to four decimals gives an error of 1.5 mm in the y coordinate. If λ_0 is given in degrees, this corresponds to eight decimals. In order to, as far as possible, reduce the errors of the rounding, round off of λ_0 and k_0 was included in these programs. The computation is done gradually. In step one, a regular combined fit with the TM projection and a 2D Helmert transformation is performed. In step two the parameters estimated for λ_0 and k_0 in step one are rounded off to an appropriate number of decimals, after which they are fixed to the rounded values and the fit is redone. The point of redoing the fit with fixed values is that the other estimated parameters compensate for the round off. It is important that the rounding of λ_0 is done in degrees, since the decimal part of the degrees can always be converted to a finite decimal fraction in minutes or in minutes and seconds, while the opposite is not true. The above description of how the rounding is implemented is somewhat simplified, but a detailed description would hardly be readable.

To improve the work of RIX 95, the programmes have been adjusted to the rest of the production environment. For example, apart from a text file of the representation of the fit, tf/tfi files (transformation files) and a GPLOT file to be used in GTRANS are given. In the tf/tfi files there are also geodetic coordinates and their transformed equivalents for the four extreme points that encircle the area of the fit. Appendix 2 contains examples of a text file and a transformation file.

In conclusion, it can be mentioned that in Stockholm, Gothenburg and Malmö, the same grid system has been applied in the neighbouring municipalities, with local accents arising as a result. Often the local accent is of good quality within each municipality, but contradictions arise when working across the municipal borders. This means that a mutual set of transformation parameters for e g Greater Stockholm would give too poor an accordance in all municipalities. A compromise is to estimate parameters for each municipality and at the same time add artificial points along the municipal borders that, with appropriate weighting, stop the contradictions on the borders from becoming too large. The method of multiple fit, which is similar to photogrammetric block triangulation, has been implemented both for 2D Helmert and for TM projection.

Another facility to be mentioned is that all the programmes for fitting can handle the problem of points having different identities in different coordinate files. The solution is a so called key file, where one states what identities are valid for each point. The multifit programmes can cope with three identities per point. In these programmes there is no need to gather all geodetic coordinates or grid coordinates in the same file, one can use a meta file with, among other things, the names of the files containing the coordinates.

References

Bjerhammar A (1967), *Geodesi.* Almqvist & Wiksell. Stockholm 1967.

Bowring B R (1976), *Transformation from spatial to geodetic coordinates.* Survey Review 181:323-327.

Hedling G & Reit B-G (1989), *Transformation WGS 84 till RT 90 (RT38).* Kartavdelningen Informerar: Lantmäteriverktet, Gävle

Krüger L (1912), *Konforme Abbildung des Erdellipsoids in der Ebene.* eröffentlichung des Könlichen Preuszischen Geodätischen Institutes, 52

Reit B-G (1998), *The 7-Parameter Transformation To A Horizontal Geodetic Datum.* Survey Review Vol. 34 No. 268 (April 1998)

Reit B-G (1999), *Improving A Horizontal Datum Without Changing The Coordinates.* Survey Review Vol. 35 No. 272 (April 1999)

Ussisoo I (1977) *Kartprojektioner* Lantmäteriverkets tekniska skrifter, 1977/6

Appendices

Appendix 1: 3D Helmert fit without height constraint

```
/*Program WOPTFIT Computation made Ferbuary 11, 2009.
From system: SWEREF 99 lat long ellh
To   system: RR 92
From input file: Sw99_LatLongh.k
To   input file: RR92.k
Number of points found in "From"-file:   20
Number of points found in "To  "-file:   21
Number of common points used in the fitting:   20
The matching of points is based on common identities of the two
input files.
Number of least squares iteration steps:    6
Units: length - meter, arc - arcsec, scale - ppm (mm per km)
Ellipsoid of FROM-system:  a=6378137.000 1:f=298.257222101 (GRS 80)
Ellipsoid of   TO-system:  a=6377397.155 1:f=299.152812800 (Bessel
1841)
The "topocentres" are oriented by the following geodetic coordinates
FROM-system:    61  16  11.003050    16   4  25.549145         .0000
  TO-system:    61  16  11.003050    16   4  25.549145         .0000
Geocentric coordinates of the origins of the "topocentres"
FROM-system:   2953641.3560    851059.7834   5569851.9261
  TO-system:   2953275.9071    850954.4833   5569274.9557
Transformation parameters between topocentric systems (RZ*RY*RX)
(Note, the topocentric systems are left-handed.)
Translation  of topocentric x:       83.6859793085 m
Translation  of topocentric y:      173.4068423468 m
Translation  of topocentric z:      -36.6385863800 m
Rotation around topocentric x:        3.1751605455 arcsec
Rotation around topocentric y:       -2.2943202986 arcsec
Rotation around topocentric z:        6.2681584553 arcsec
             Scale correction:           .00000000 ppm (mm per km)  (fixed)
                       (Scale:    1.00000000000000)  (fixed)
Transformation parameters between geocentric systems (RZ*RY*RX)
Translation  of  geocentric X:     -414.0978562888 m
Translation  of  geocentric Y:      -41.3381702518 m
Translation  of  geocentric Z:     -603.0627127551 m
Rotation around  geocentric X:        -.8550428002 arcsec
Rotation around  geocentric Y:        2.1413464567 arcsec
Rotation around  geocentric Z:       -7.0227212665 arcsec
             Scale correction:           .00000000 ppm (mm per km)
                       (Scale:    1.00000000000000)
```

Sign of residuals: transformed minus original

	Residuals				A priori st.dev.		
Station	North	East	Up	2D	North	East	Up
ARJE.0	-.0357	.0232	-7.9041	.0426	.0500	.0500	999.0000
KIRU.0	-.0389	-.0723	-7.5773	.0820	.0500	.0500	999.0000
OVER.0	.0508	.0431	-5.9735	.0667	.0500	.0500	999.0000
SKEL.0	-.0188	.0369	-5.8436	.0414	.0500	.0500	999.0000
VILH.0	.0036	.0199	-7.5933	.0202	.0500	.0500	999.0000
BORA.0	-.0505	.0267	-6.5201	.0571	.0500	.0500	999.0000
JONK.0	-.0517	-.0144	-5.9488	.0537	.0500	.0500	999.0000
SUND.0	.0191	.0077	-5.9215	.0206	.0500	.0500	999.0000
HASS.0	.0764	.0030	-5.4994	.0765	.0500	.0500	999.0000
NORR.0	-.0403	-.0331	-5.2133	.0522	.0500	.0500	999.0000
ONSA.0	-.0580	.0719	-7.0300	.0924	.0500	.0500	999.0000
VANE.0	-.0114	-.0233	-7.4501	.0259	.0500	.0500	999.0000
KARL.0	.0209	-.0453	-6.6183	.0499	.0500	.0500	999.0000
LEKS.0	.0286	.0088	-6.5250	.0299	.0500	.0500	999.0000
LOVO.0	.0506	.0168	-4.6851	.0533	.0500	.0500	999.0000
MART.6	.0155	-.0115	-5.4631	.0193	.0500	.0500	999.0000
OSKA.0	.0141	-.0534	-4.6677	.0552	.0500	.0500	999.0000
OSTE.0	-.0046	-.0144	-7.9372	.0151	.0500	.0500	999.0000
SVEG.0	.0245	-.0193	-7.3052	.0312	.0500	.0500	999.0000
UMEA.0	.0026	.0348	-5.6617	.0349	.0500	.0500	999.0000
R.m.s.	.0369	.0349	6.4462	.0508			

```
Residual having largest absolute value:
Topocentric  x-component    .0764 at point HASS.0
Topocentric  y-component   -.0723 at point KIRU.0
Topocentric  z-component  -7.9372 at point OSTE.0
Topocentric 2D-component    .0924 at point ONSA.0
```

Appendix 2: Projection fit combined with a 2D Helmert fit

Example of a text file presenting the results of a fit

```
PROJFIT compiled May 04, 2004
The parameters are based on a least squares fit using a combined
Transverse Mercator projection and 2D similarity transformation.
The result is based on coordinates from the files
Sweref99_latlong_h=0.k and Rotstad.k
Geodetic coordinates: SWEREF 99 lat long
Grid     coordinates: Rotstad lokala
The matching of points is based on the identities of the key file
key.txt.
In total  119 points of file Rotstad.k was not matched with a point
in file Sweref99_latlongh=0.k
Number of common points:   12
--------------------------------------------------------------------
Minimum and maximum coordinate values in degrees and minutes:
Latitude                   55  54            56  14
Longitude                  12  34            12  57
--------------------------------------------------------------------
/
PROJECTION PARAMETERS
PROJECTION Transverse Mercator
REFERENCE FRAME SWEREF 99 lat long/
GRID SYSTEM Fictive x y /
ELLIPSOID GRS 1980
    6378137.000   298.2572221010 /
CENTRAL MERIDIAN
    13  31  42.4560000000 /
SCALE
     .999972040000 /
FALSE NORTHING
  -6203871.2490 /
FALSE EASTING
     61645.0200 /
LATITUDE OF ORIGIN
          .0000 /
END OF PROJECTION PARAMETERS

/*------------------------------------------------------------------

HELMERT PARAMETERS /
FSYSTEM Fictive x y /
TSYSTEM Rotstad lokala/
AREA          -5250.          1700.         29582.         25040.
/
FORMEL PLAN
6-PAR HELMERT
     -646.511370993850300   9.989597174353925E-001
4.560132414182313E-002
      604.239294856388700  -4.560132414182313E-002
9.989597174353925E-001
/
END OF HELMERT PARAMETERS

HELMERT INVERSE PARAMETERS /
FSYSTEM Rotstad lokala/
```

```
AREA          -5250.          1700.         29582.         25040.
/
TSYSTEM Fictive x y /
FORMEL PLAN
6-PAR HELMERT
      673.392929897349200   9.989597196110401E-001  -
4.560132424113886E-002
     -574.128941913437000   4.560132424113886E-002
9.989597196110401E-001
/
END OF INVERSE HELMERT PARAMETERS

/*------------------------------------------------------------------

RESIDUALS
/* Sign of residuals: transformed minus original grid coordinates
                           North        East      Radial
  732311   1               -.107        .008        .107
  732631   29               .001        .014        .014
  732121   57               .058       -.054        .079
  732041   58               .014       -.049        .051
 7327190   60               .001       -.031        .031
732531.2   598              .039        .084        .093
 7323190   10111           -.095        .009        .095
 7324090   10113           -.074       -.046        .087
 7315990   10116            .057       -.009        .058
732122.2   732122.2         .046        .002        .046
  732241   732241           .057        .030        .065
  732611   732611           .002        .043        .043

R.m.s.                      .057        .040        .070
Max deviation              -.107        .084        .107 at   732311
1

Number of common points       12

/*------------------------------------------------------------------

DESCRIPTION OF THE 2-DIMENSIONAL HELMERT TRANSFORMATION

The parameters should be used with the algebraic formula:

xt = dx + Scale * ( xf * cos(Rot.) - yf * sin(Rot.) )
yt = dy + Scale * ( xf * sin(Rot.) + yf * cos(Rot.) )

Let a = Scale * cos(Rot.) and b = Scale * sin(Rot.) then

xt = dx + a * xf - b * yf
yt = dy + b * xf + a * yf

/*------------------------------------------------------------------

Parameters for the direction Fictive x y   ->  Rotstad lokala

Rot.  =    -2.904077551862 gon (  -2.613669796676 degr.)
Scale =  .9999999989110433
dx    =         -646.51137 m.  (translation along South-North axis)
dy    =          604.23929 m.  (translation along West - East axis)
a     =  .9989597174353924
b     = -.0456013241418231

/*------------------------------------------------------------------
```

```
Parameters for the direction Rotstad lokala  ->  Fictive x y

Rot.  =     2.904077551862 gon (   2.613669796676 degr.)
Scale = 1.0000000010889570
dx    =          673.39293 m.  (translation along South-North axis)
dy    =         -574.12894 m.  (translation along West - East axis)
a     =  .9989597196110401
b     =  .0456013242411389

/*------------------------------------------------------------------
```

Example of a transformation file (tf-file) ready to use in GTRANS

```
/* PROJFIT compiled May 04, 2004

TRANSFORMATION details

The parameters are based on a least squares fit using a combined
Transverse Mercator projection and 2D similarity transformation.

Ellipsoid: GRS 1980

TRANSVERSE MERCATOR PARAMETERS
Central meridian                13  31  42.456000 degr. min. sec.
Scale along central merdian         .999972040000
False northing                      -6203871.2490 m
False easting                          61645.0200 m

/
FSYSTEM SWEREF 99 lat long/
LATLONG DEG/
TSYSTEM Rotstad lokala/
AREA          -5250.          1700.         29582.         25040./
ELLIPSOID GRS 1980/
PROJ Gauss
G     13  31 42.4560000000 DEG
  -6203871.2490
     61645.0200
     .999972040000 /
SYSTEM Fictive x y /
FORMEL PLAN
6-PAR HELMERT
     -646.511370993850300   9.989597174353925E-001
4.560132414182313E-002
      604.239294856388700  -4.560132414182313E-002
9.989597174353925E-001
/
GRUNDMEDELFEL            .055     19 /
STOP /
Number of common points          12

/*-------------------------------------------------------------------

DESCRIPTION OF THE 2-DIMENSIONAL HELMERT TRANSFORMATION

The parameters should be used with the algebraic formula:

xt = dx + Scale * ( xf * cos(Rot.) - yf * sin(Rot.) )
yt = dy + Scale * ( xf * sin(Rot.) + yf * cos(Rot.) )

Let a = Scale * cos(Rot.) and b = Scale * sin(Rot.) then

xt = dx + a * xf - b * yf
yt = dy + b * xf + a * yf

/*-------------------------------------------------------------------

Parameters for the direction Fictive x y   ->  Rotstad lokala

Rot.  =    -2.904077551862 gon (  -2.613669796676 degr.)
Scale =  .9999999989110433
dx    =         -646.51137 m.  (translation along South-North axis)
```

```
dy    =          604.23929 m.  (translation along West - East axis)
a     =  .9989597174353924
b     = -.0456013241418231

/*------------------------------------------------------------------

The coordinates used in the fitting process are taken from the
files:
Geodetic   coordinates: Sweref99_latlong_h=0.k
Plane grid coordinates: Rotstad.k

Geodetic system name: SWEREF 99 lat long
Grid     system name: Rotstad lokala_lokalt

Further details of the computation are given in the file
Rix_chk_S2_1.
/*------------------------------------------------------------------
Worked example, (4 corners surrounding the valid area)
                      Latitude             Longitude
Northing(m)    Easting(m)
South-west   55°  54'   .00000"    12°  34'   .00000"        -
6769.862      2369.249
South-east   55°  54'   .00000"    12°  57'   .00000"        -
5943.070     26333.935
North-west   56°  14'   .00000"    12°  34'   .00000"
30326.446      1193.302
North-east   56°  14'   .00000"    12°  57'   .00000"
31145.096     24952.114
/*------------------------------------------------------------------
End of information.
```

Rapporter i geodesi och geografiska informationssystem från Lantmäteriet

2006:4 Klang Dan: KRIS-GIS® projekt i Eskilstuna. Kvalitet i höjdmodeller.

2006:5 von Malmborg Helena: Jämförelse av Epos och nätverks-DGPS.

2006:8 Wennström Hans-Fredrik (ed.): Struve Geodetic Arc 2006 International Conference – the Struve arc and extensions in space and time.

2006:9 Shah Assad: Systematiska effekter inom den tredje riksavvägningen.

2007:1 Johnsson Fredrik & Wallerström Mattias: En nätverks-RTKjämförelse mellan GPS och GPS/GLONASS.

2007:4 Ågren Jonas & Svensson Runar: Postglacial land uplift model and system definition for the new Swedish height system RH 2000.

2007:8 Halvardsson Daniel & Johansson Joakim: Jämförelse av distributionskanaler för projektanpassad nätverks-RTK.

2007:10 Lidberg Martin & Lilje Mikael: Evaluation of monument stability in the SWEPOS GNSS network using terrestrial geodetic methods - up to 2003.

2007:11 Lilje Christina, Engfeldt Andreas, Jivall Lotti: Introduktion till GNSS.

2007:12 Ivarsson Jesper: Test and evaluation of SWEPOS Automated Processing Service.

2007:14 Lilje Mikael, Eriksson Per-Ola, Olsson Per-Anders, Svensson Runar, Ågren Jonas: RH 2000 och riksavvägningen.

2008:4 Johansson S Daniel & Persson Sören: Kommunikationsalternativ för nätverks-RTK – virtuell referensstation kontra nätverksmeddelande.

2009:1 Ågren Jonas: Beskrivning av de nationella geoidmodellerna SWEN08_RH2000 och SWEN08_RH70.

2009:2 Odolinski Robert & Sunna Johan: Detaljmätning med nätverks-RTK – en noggrannhetsundersökning.

2009:4 Fridén A & Persson A-K: Realtidsuppdaterad etablering av fri station – ett fälttest med radioutsänd projektanpassad nätverks-RTK.

2009:5 Bosrup, Susanna & Illerstam, Jenny. Restfelshantering med Natural Neighbour och TRIAD vid byte av koordinatsystem i plan och höjd.

L A N T M Ä T E R I E T

Vaktmästeriet 801 82 GÄVLE Tfn 026 - 65 29 15 Fax 026 - 68 75 94
Internet: www.lantmateriet.se