

First published in:



Lisbon, May 12-15 2008

13th FIG International Symposium on Deformation Measurements and Analysis

4th IAG Symposium on Geodesy for Geotechnical and Structural Engineering

AN INNOVATIVE MATHEMATICAL SOLUTION FOR A TIME-EFFICIENT IVS REFERENCE POINT DETERMINATION

Michael LÖSLER and Maria HENNES

*Geodetic Institute of the University of Karlsruhe (TH)**Germany*

Abstract: The improvement of the local ties between different observation methods (GPS, VLBI, etc.) improves the quality of the ITRF considerably. The IVS reference point of a VLBI radio telescope is defined as the intersection between the azimuth- and elevation-axis or, if they do not intersect, the intersection of the right-angle projection from the elevation-axis onto the azimuth-axis. In the past, these axes have been estimated by fitting 3D circles, e.g. (Eschelbach et al., 2003) or (Dawson et al., 2006). The data acquisition for the determination of the circles requires that the telescope has to be moved into clearly defined positions; therefore, the basic station process (data gathering for the intrinsic telescope task) is disturbed. In this paper we present an alternative mathematical model, which computes the reference point without circle fitting. This algorithm does not need observations from predefined telescope positions and therefore the station's downtime can be reduced. The parameter estimation of this non-linear problem is implemented in two steps. At first we are using the Levenberg-Marquardt-Algorithm for a pre-evaluation to find stable approximate values (Madsen et al., 2004), which we use for the main least-square-model in a second step.

1. INTRODUCTION

The reference point of a VLBI radio telescope is defined as the intersection between the azimuth- and elevation-axis. If these axes do not intersect, the reference point is the right angle projection from the elevation-axis onto the azimuth-axis. As a rule the two axes of this telescope will be derived by 3D circle fitting and the invariant reference point will be estimated. For this the telescope rotates around one axis while the second axis is fixed and some targets on the telescope side will be observed by a theodolite or an instrument like that. This is done step by step. The trajectory of every target corresponds to a circle. The centre points of these circles are also points of the rotation axis and will be used to approximate this axis. For the determination of axis wobble, the process must be repeated for many different telescope orientations, whereas the orientation angles are not needed with high accuracy. Getting the reference point by minimization the orthogonal distance (eccentricity) between the approximated elevation- and azimuth-axis is the final step. A detailed description of this way of doing is published e.g. in (Eschelbach et al., 2003) or (Dawson et al., 2006).

Nowadays, optical tracking measuring instruments like a robot tacheometers and laser trackers enable possibilities to replace this time-consuming method, because the necessary data can be gathered while the telescope is moving and eventually doing its intrinsic task. If the conventional circle fitting method shall be applied, the circle model has to be expanded to a torus-like structure to approximate the whole unstructured data set. The dimension of the torus depends on the distance between the target and the elevation-axis; and the torus is very thin due to the small – and unknown – eccentricity between the azimuth-axis and the elevation-axis. Therefore, the results for the unknown parameters (at least the eccentricity) become uncertain.

Therefore will present an alternative method to estimate the reference point without circle (or torus) fitting in this paper. The mathematical model uses the 3D coordinates from targets on the side of the telescope as auxiliary parameters. Although the model requires the elevation angles and the azimuth angles assigned to the measuring time of the specific target to get the connection between the telescope's orientation and the local site network, the suggested algorithm does not need observations assigned to predefined telescope positions. The method can roughly be compared to solving two datasets for specific transformation parameters. To solve the non-linear-problem the use of a damped Gauß-Newton-Method called Levenberg-Marquardt-Algorithm, which is briefly described in section 3.3, provides a first reliable set of approximate values for the main least-square-model.

2. CONDITIONS AND RESTRICTIONS

2.1. Coordinate systems

There are two different coordinate systems to distinguish between. Both are defined as mathematical (right-handed), cartesian coordinate systems. Firstly there is the standard observation x_{Obs} , y_{Obs} , z_{Obs} system from the observation instrument. This one can be the local site network at the station and does not need a detailed description. The second one is the telescope system x_{Tel} , y_{Tel} , z_{Tel} . It is defined by the following:

- Origin of the coordinate system is the reference point
- The x-axis is parallel to the elevation-axis
- The z-axis corresponds to the azimuth-axis of the telescope
- The y-axis is normal to the x- and z-axis

The telescope system rotates around the z-axis relatively to the fixed geodetic observation system by the azimuth angle.

2.2. Restrictions

An ideal radio telescope is not given. Because of this, the mathematical model has to allow for some restrictions on rather corrections, which are shown in figure 1. They are parts of the unknown parameters, which are estimated, too. There are three deviations related to the construction of the telescope.

1. The elevation- and azimuth-axis do not intersect. There is an eccentricity between these axes.

2. The angle between the elevation- and azimuth-axis are not right-angled, therefore there is a tiny correction angle.
3. The azimuth-axis and the z-axis of the observation system are not parallel to each other but differ by a small angle.

In addition, the positions of the observation-targets on the side of the telescope are arbitrarily. Only the direction of rotation and the magnitude between two telescope orientations are the same (figure 4 in section 3.1). So, every angle gets a correction value for the specific target, too. To demonstrate the first and the third restriction (the second one is quite conceivable and not shown), in Figure 1 the observation coordinate systems is shift while the z-axis intersects the azimuth-axis of the telescope.

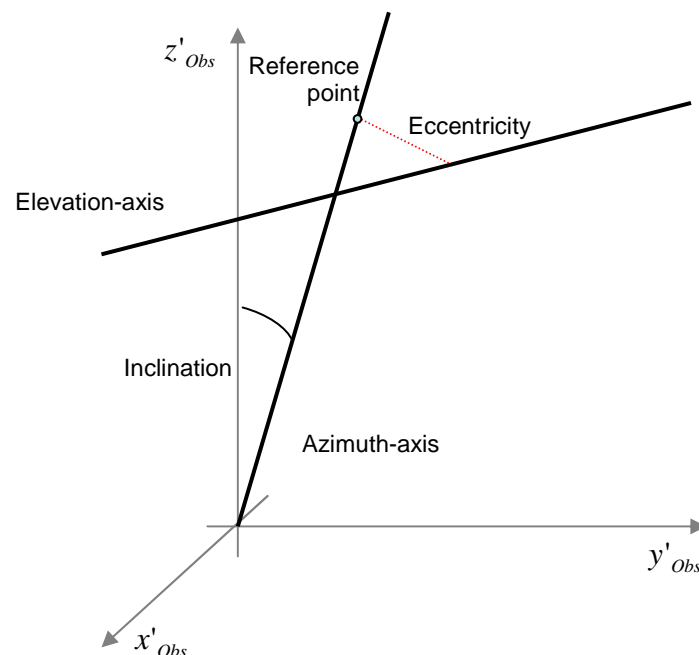


Figure 1: Restrictions (eccentricity and inclination), for clearness shown in the coordinate system x'_{Obs} , y'_{Obs} , z'_{Obs} , which emanates from the observation system x_{Obs} , y_{Obs} , z_{Obs} by translation

3. MATHEMATICAL MODEL

3.1. Derivation

Due to the restrictions 1 and 2, which result in additional parameters, all the unknown parameters can not be solved in a one-step Helmert-Transformation. Therefore, in this section we present the derivation of the new mathematical model by a step by step introduction for one target. In the end of the section we obtain three transform equations, which can be used to estimate the invariant reference point P_R in a closed mathematical model. In the following the superior index is used to denote the result of a transformation equation, in this case identical with the equation's number.

Firstly, we adopt that the two defined coordinate systems in section 2.1 are congruent to each other. So, the observation coordinate system is equal to the telescope coordinate system. In the course of the derivation the difference between these coordinate systems will be explained and the transformation formulas will be given. A general point P of the rotational solid is defined under disregard for every restriction and without any telescope twist, that means, that the telescope orientation angles are zeros, as

$$P^1 = [b \quad a \quad 0]^T, \quad (1)$$

whereas b is the distance along the x-axis and a is the shortest distance between the point P and the x-axis of the point coordinate system, which is denoted by an apostrophe and move on to the telescope system in the end (note figure 3 and 4). The z-value is set to zero, because the elevation-angle E (Epsilon) is set to zero and therefore $z_{p^1} = a \cdot \sin E = 0$. So, the point P , which is represented by the target $T_{1,1}$ in figure 2, lies within the xy-plane.

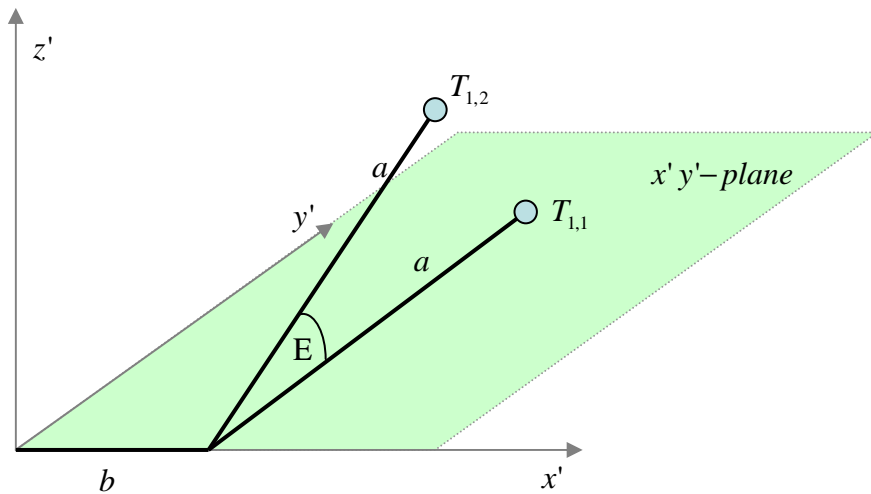


Figure 2: Point definition

If the telescope rotates around the elevation-axis by an angle E , the point P is the result of the matrix multiplication (figure 2, target $T_{1,2}$):

$$P^2 = R_x(E) \cdot P^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos E & -\sin E \\ 0 & \sin E & \cos E \end{bmatrix} \cdot \begin{bmatrix} b \\ a \\ 0 \end{bmatrix} = \begin{bmatrix} b \\ a \cdot \cos E \\ a \cdot \sin E \end{bmatrix}, \quad (2)$$

whereas $R_x(E)$ describes the rotation matrix for a rotation with the elevation-angle E around the x-axis. An eccentric distance e between the two telescope axes, see the first restriction in section 2.2, displaces the y-value of P :

$$\mathbf{P}^3 = \mathbf{Ecc} + \mathbf{P}^2 = \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix} + \begin{bmatrix} b \\ a \cdot \cos E \\ a \cdot \sin E \end{bmatrix} = \begin{bmatrix} b \\ e + a \cdot \cos E \\ a \cdot \sin E \end{bmatrix}. \quad (3)$$

The non-orthogonality between the axes of the telescope is the second restriction. It can be modelled by a rotation $\mathbf{R}_Y(\gamma)$ around the y -axis with the correction-angle γ .

$$\mathbf{P}^4 = \mathbf{R}_Y(\gamma) \cdot \mathbf{P}^3 = \begin{bmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{bmatrix} \cdot \begin{bmatrix} b \\ e + a \cdot \cos E \\ a \cdot \sin E \end{bmatrix} = \begin{bmatrix} \cos \gamma \cdot b + \sin \gamma \cdot a \cdot \sin E \\ e + a \cdot \cos E \\ -\sin \gamma \cdot b + \cos \gamma \cdot a \cdot \sin E \end{bmatrix} \quad (4)$$

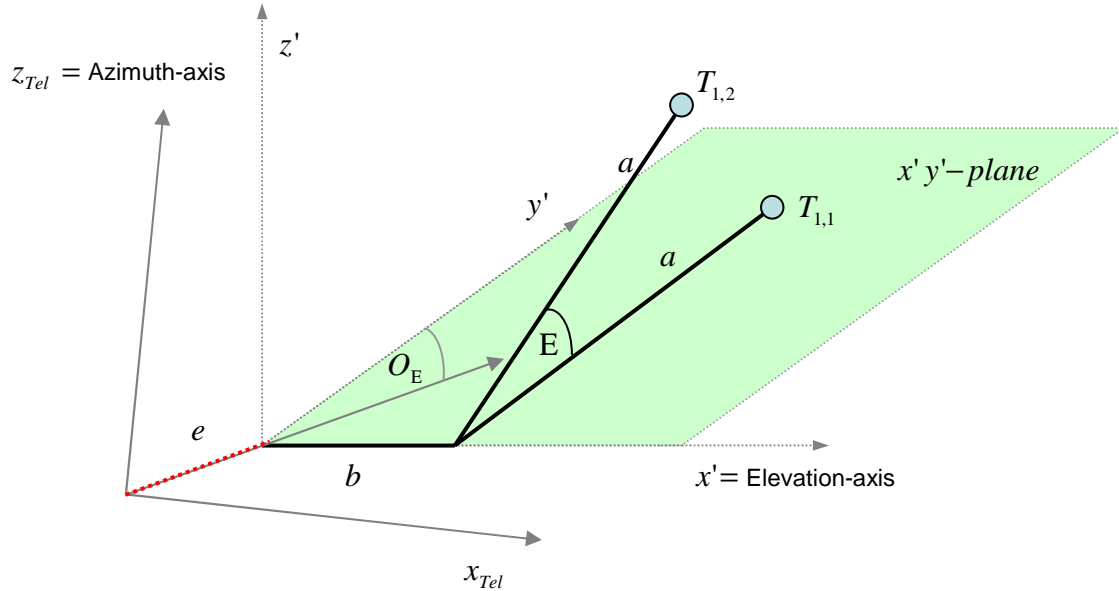


Figure 3: Connection between the point system and the telescope system

So far the two defined coordinate systems in section 2.1 are congruent to each other because there are no twists or translations. However, the target representing point \mathbf{P} rotates with the telescope around the azimuth-axis. This rotation can be described by the rotation matrix $\mathbf{R}_Z(A)$ and the azimuth angle A (Alpha) as follows:

$$\mathbf{P}^5 = \mathbf{R}_Z(A) \cdot \mathbf{P}^4 = \begin{bmatrix} \cos A & \sin A & 0 \\ -\sin A & \cos A & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \gamma \cdot b + \sin \gamma \cdot a \cdot \sin E \\ e + a \cdot \cos E \\ -\sin \gamma \cdot b + \cos \gamma \cdot a \cdot \sin E \end{bmatrix} \quad (5)$$

The third restriction in section 2.2 was the non-parallelism between the z -axis of the local network coordinate system and the azimuth-axis of the radio telescope. To model this inclination two rotations and correction-angles are needed. The rotation around the y -axis with an angle α rotates the azimuth-axis into the xz -plane. The second rotation around the x -axis with the correction-angle β is essential to get the parallelism-condition between these

two axes. The inclination correction $R_{Y,X}(\alpha, \beta)$ can be described by the matrix multiplication:

$$\begin{aligned}
 R_{Y,X}(\alpha, \beta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta \\ 0 & \sin \beta & \cos \beta \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix} \\
 &= \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ -\sin \beta \cdot -\sin \alpha & \cos \beta & -\sin \beta \cdot \cos \alpha \\ \cos \beta \cdot -\sin \alpha & \sin \beta & \cos \beta \cdot \cos \alpha \end{bmatrix}
 \end{aligned} \tag{6}$$

The point P is then described by the equation:

$$P^7 = R_X(\beta) \cdot R_Y(\alpha) \cdot P^5 = R_{Y,X}(\alpha, \beta) \cdot P^5 \tag{7}$$

Finally, a translation vector is added; that describes the connection between the origins of the two coordinate systems. This vector P_R includes the coordinates of the invariant telescope reference point. So, we get the three transformation equations – one for each coordinate-component –, which can be written as matrix addition:

$$P = P_R + P^7 \tag{8}$$

Remember the different orientations problem between the targets and the radio telescope in section 2.2. In order to use the azimuth- and elevation-angle of the telescope to transform the point P between the two coordinate systems, add the orientation corrections O_A and O_E to these angles, refer figure 4:

$$A_p = A + O_A \tag{9}$$

$$E_p = E + O_E \tag{10}$$

The elevation correction angle O_E is to estimate for every specific target T separately whereas the azimuth correction angle O_A is fixed for all targets.

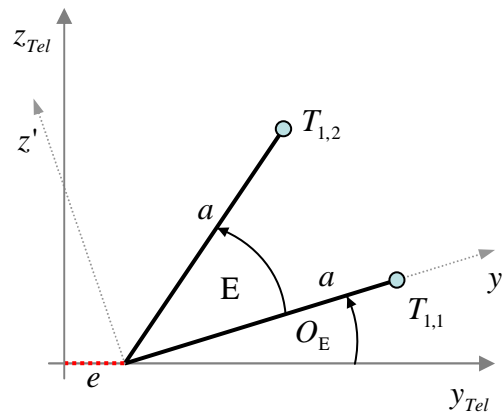


Figure 4: Target position after elevation rotation with correction angle

3.2. Least-Square-Model

The three transformation equations – one for each coordinate-component – in the section above can be used to estimate the telescope reference point \mathbf{P}_R by a least-square-adjustment called Gauß-Helmert-Model. The (error-free) observation parameters $\hat{\mathbf{L}}$ to solve the non-linear problem $F(\hat{\mathbf{L}}, \hat{\mathbf{X}})$ include the 3D coordinates $[X_{i,epo_{T_i}}, Y_{i,epo_{T_i}}, Z_{i,epo_{T_i}}]^T$ of the several targets $T_{i,epo_{T_i}}$, furthermore the telescope orientation angles $A_{i,epo_{T_i}}$ and $E_{i,epo_{T_i}}$, whereas i is the number of the specific target and epo_{T_i} the associated observation epoch. If T_1 is the first target, the observations can be written as:

$$\hat{\mathbf{L}}_1 = \mathbf{L}_1 + \mathbf{v}_1 = [X_{1,1}, Y_{1,1}, Z_{1,1}, A_{1,1}, E_{1,1}, \dots, X_{1,epo_{T_1}}, Y_{1,epo_{T_1}}, Z_{1,epo_{T_1}}, A_{1,epo_{T_1}}, E_{1,epo_{T_1}}]^T + \mathbf{v}_1 \quad (11)$$

with $\hat{\mathbf{L}}$ are the true values of the observation \mathbf{L} .

The vector of unknown parameters $\hat{\mathbf{X}}$ can be classified in two groups:

- fixed parameters $\hat{\mathbf{X}}_{const}$ and
- target-depended parameters $\hat{\mathbf{X}}_{target}$.

The eight fixed parameters are the 3D coordinates $[X_{P_R}, Y_{P_R}, Z_{P_R}]^T$ of the reference point \mathbf{P}_R , the eccentricity e between the telescope axes, the small angles α and β to correct the inclination, the angle γ to correct the non-orthogonality between the axes and the azimuth orientation correction O_A . For every target the number of unknown parameters raises up by three. These target-depended parameters are the distance values a and b along the axes with reference to the reference point \mathbf{P}_R and the telescope coordinate system and the elevation correction angle O_E . The number of unknown parameters u is:

$$u = u_{const} + u_{target} = 8 + 3 \cdot m. \quad (12)$$

It follows from the above equitation for the degree of freedom f

$$f = n - u = 3 \cdot \sum_{epo=1}^{epo_{max}} T_{i,epo} - (8 + 3 \cdot m), \quad (13)$$

whereas m is the number of targets.

The described transformation equations are non-linear; therefore they have to be linearised by a first-order Taylor expansion at first:

$$F(\hat{\mathbf{L}}, \hat{\mathbf{X}}) = F(\mathbf{L} + \mathbf{v}, \mathbf{X}_0 + \mathbf{x}) = \underbrace{F(\mathbf{L}, \mathbf{X}_0)}_{\mathbf{w}} + \underbrace{\frac{\partial F(\mathbf{L}, \mathbf{X}_0)}{\partial \mathbf{L}}}_{\mathbf{B}} \cdot \underbrace{(\hat{\mathbf{L}} - \mathbf{L})}_{\mathbf{v}} + \underbrace{\frac{\partial F(\mathbf{L}, \mathbf{X}_0)}{\partial \mathbf{X}_0}}_{\mathbf{A}} \cdot \underbrace{(\hat{\mathbf{X}} - \mathbf{X}_0)}_{\mathbf{x}} = 0 \quad (14)$$

The function of minimisation of this Gauß-Helmert-Model is given by e.g. (Niemeier, 2002) as follows:

$$\Omega = \mathbf{v}^T \mathbf{Q}_{LL}^{-1} \mathbf{v} + 2\mathbf{k}^T \cdot (\mathbf{B}\mathbf{v} + \mathbf{A}\mathbf{x} + \mathbf{w}) \rightarrow \min \quad (15)$$

with the normal-equation:

$$\begin{bmatrix} \mathbf{B} \cdot \mathbf{Q}_{LL} \cdot \mathbf{B}^T & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{k} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} -\mathbf{w} \\ \mathbf{0} \end{bmatrix} \text{ or } \begin{bmatrix} \mathbf{B} \cdot \mathbf{Q}_{LL} \cdot \mathbf{B}^T & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix}^{-1} \cdot \begin{bmatrix} -\mathbf{w} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{k} \\ \mathbf{x} \end{bmatrix}. \quad (16)$$

Whereas \mathbf{A} is the Jacobian-matrix, which contains the partial derivatives with respect to the parameters \mathbf{X} , \mathbf{B} is the design matrix of conditions, which contains the partial derivatives with respect to the observations \mathbf{L} , \mathbf{w} is the vector of contradictions, \mathbf{Q}_{LL} is the cofactor matrix of the observations \mathbf{L} , \mathbf{x} is the vector of increments and the vector \mathbf{k} consists of so called Laplace multipliers.

The Gauß-Helmert-Model needs like every optimization method approximate values \mathbf{X}_0 for the unknown parameters, which are updated by every iteration:

$$\hat{\mathbf{X}} = \mathbf{X}_0 + \mathbf{x}. \quad (17)$$

At each iteration the estimated values $\hat{\mathbf{X}}$ will be used as approximate values \mathbf{X}_0 . This determination has to be repeated until the improvements are converging towards zero. The number of iterations is depending on the quality of the approximate values \mathbf{X}_0 . To get a first reliable solution the Levenberg-Marquardt-Algorithm, which is briefly described in the next section, can be used.

3.3. Levenberg-Marquardt-Algorithm

To solve a non-linear least-square problem reliable, in 1944 Kenneth Levenberg published the suggestion to use a so called method of damped least square (Levenberg, 1944), which Donald Marquardt took up again in 1963. The Levenberg-Marquardt-Algorithm, named after its developer, is a hybrid method between the method of steepest descent (also called as: gradient descent direction) and the Gauß-Newton-Method. Both ones are able to solve a non-linear problem iteratively. The main-differences between these methods are the number of required iterations and therefore the runtime and the different convergence criteria.

The Levenberg-Marquardt-Algorithm is an iterative method and locates the minimum of a function F in respect to the unknown parameters \mathbf{X} and is a standard technique for non-linear least-square problems (Lourakis, 2005). The damped Gauß-Newton-Method is be described in (Marquardt, 1963) by the equation:

$$(\mathbf{A}^T \mathbf{A} + \mu \mathbf{I}) \mathbf{x} = -\mathbf{A}^T \mathbf{w}, \quad (18)$$

whereas \mathbf{A} is the Jacobian-matrix, which includes the first derivations of the function $F(\mathbf{X})$, and the matrix \mathbf{I} is the identity matrix. The vector \mathbf{w} contains the residuals of the function. The vector of increments \mathbf{x} is the so-called damped Newton step and μ , ($\mu \geq 0$), is the damping parameter, which influences the direction and the size of the specific step. The scalar μ has to be set one-times in dependence on the confidence of the approximation values under the condition (Marquardt, 1963):

$$\Omega_{k+1} < \Omega_k \quad (19)$$

whereas Ω is the function of minimisation at the k^{th} iteration.

For all $\mu > 0$ it ensures that \mathbf{x} converges along the direction of the minimum because the coefficient matrix \mathbf{A} is positive definite. Furthermore, a large value of μ means that the matrix is diagonal-dominated and the current solution is far from the correct one. The method works slowly because it is only a short step in the steepest descent direction, but it guarantees to converge.

$$\mathbf{x} \cong -\frac{1}{\mu} \mathbf{A}^T \mathbf{w} \quad (20)$$

On the other hand, for a very small value of μ the algorithm switches to the Gauß-Newton-Method and gets (almost) quadratic convergence (Madsen et al., 2004) because it is

$$\left(\mathbf{A}^T \mathbf{A} + \mu \mathbf{I} \right) \cong \left(\mathbf{A}^T \mathbf{A} \right) \quad (21)$$

and therefore

$$\mathbf{x} \cong -\left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \mathbf{w} . \quad (22)$$

At each iteration the error reduction will be verified (quod vide equation 19) and the damping parameter adjusted. If the current step failed to reduce the errors, the damping parameter will have to be increased. Otherwise μ will be reduced. For this reason the Levenberg-Marquardt-Algorithm is adaptive (Lourakis, 2005) and provides reliable (robust) values.

A detailed analyse of the Levenberg-Marquardt-Algorithm is published by (Madsen et al., 2004). For further information the interested reader is referred to this paper. Additionally, there is described an implementation of this algorithm. Furthermore, a short description and an improved implementation in C/C++ under the terms of the GNU General Public License are published by (Lourakis, 2005).

4. CONCLUSION

We have derived an alternative procedure to compute the invariant reference point of a VLBI radio telescope without circle fitting. The algorithm estimates the reference point and also the antenna parameters “eccentricity” and “inclination” with respect to the telescope restrictions in a closed model. It is possible to reduce the station's downtime because the mathematical model does not require observations from predefined telescope positions as it is needed for circle-fitting. Instead, the observation-data-referenced telescope-orientation is needed. This can be easily archived by combining time-stamped tachymeter (or laser tracker) data with the telescope observation protocol. Investigations proved that the determination of the reference point and the additional parameters will not be affected noteworthy by the uncertainty introduced by this method of synchronisation, if the data during a source observation is gathered, meanwhile the telescope moves very slowly. If all the data, i.e. including that gathered during the repositioning of telescope to another source, is used, a sufficient synchronisation can be achieved by using a trigger signal of the telescope's control clock, which triggers the laser tracker (Juretzko et al, 2008). In both ways of doing, a reference point determination could be carried out while the intrinsic station process is working. Our further work will focus on the economic efficiency by practical applications of different measurement equipment, culminating e.g. in active (i.e. self-orienting) reflector hubs.

In addition, we presented the damped Newton method called Levenberg-Marquardt-Algorithm for the determination of approximate values, which is an efficient technique for non-linear least-square problems because it provides reliable values. This damped Newton method is a hybrid method to solve the non-linear problem. It is a combination of the steepest descent method and the Gauß-Newton-Method.

Acknowledgments

The author is indebted to Dr. Martin Nitschke for valuable discussions during the work on the model. This research has been supported by the Deutsche Forschungsgemeinschaft (DFG).

References

- Dawson, J., Sarti P., Johnston G., Vittuari L. (2006). *Indirect approach to invariant point determination for SLR and VLBI systems: an assessment*. Journal of Geodesy, Vol 81, p.433-441.
- Eschelbach, C., Haas R. (2003). *The IVS-Reference Point at Onsala - High End Solution for a real 3D-Determination*. Schwegmann, W./Thorandt, V. (Hrsg.) Leipzig/Frankfurt a. M., Bundesamt für Kartographie und Geodäsie, Proceedings 16th Working Meeting on European VLBI for Geodesy and Astronomy, p.109-118.
- Juretzko, M., Hennes, M., Schneider, M., Fleischer, J. (2008). *Überwachung der raumzeitlichen Bewegung eines Fertigungsroboters mit Hilfe eines Lasertrackers*. Allgemeine Vermessungsnachrichten, Feb-Issue, in print.
- Levenberg, K. (1944). *A Method for the Solution of Certain Problems in Least Squares*. Quarterly of Applied Mathematics, Vol 2, p.164–168.
- Lourakis, M.I.A. (2005). *A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar*. Institute of Computer Science Foundation for Research and Technology - Hellas (FORTH), <http://www.ics.forth.gr/~lourakis/levmar/levmar.pdf> (last visited: 2008-01-08).
- Madsen, K., Nielsen H.B., Tingleff, O. (2004). *Methods for non-linear Least Square Problems*. Technical University of Denmark, 2nd Edition, http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3215/pdf/imm3215.pdf (last visited: 2008-01-08).
- Marquardt, D.W. (1963). *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*. Society for Industrial and Applied Mathematics, Vol 11, p.431–441.
- Niemeier, W. (2002). *Ausgleichsrechnung*. Berlin, 1st Edition, Walter de Gruyter.

Corresponding author contacts

Michael Lösler
michael.loesler@gik.uni-karlsruhe.de
Geodetic Institute of the University of Karlsruhe (TH)
Germany